Molecular anthropology:
Touching the past through ancient DNA retrieval.
Methodological aspects

Henryk W. Witas

Molecular Biology Unit, Institute of Pediatrics, Medical University of Łódź,
Sporna 36/50, 91-736 Łódź, Poland; E-mail: witas@alef.am.lodz.pl

ABSTRACT   The revolution which introduced new techniques of molecular biology applied to DNA analysis enormously accelerated the progress in most areas of medicine and biology. Techniques such as *polymerase chain reaction* (PCR), restriction analysis and sequencing are widely used for diagnosis of a number of diseases, for genetic screening, phylogenetic analysis and population studies. Moreover, it became possible to study genetic relationships of extinct to contemporary organisms and even to follow evolutionary events. Variation in DNA sequences, especially that of humans, is fascinating not only for our own sake, but also because of the inferences that can be drawn from it about our recent evolution, demography and movements. Selected problems arising during ancient DNA (aDNA) isolation and analysis are discussed. Environment and time related factors altering the structure of nucleic acids as well as contamination of isolated material are among methodological problems that arise during the procedure of isolation and processing of aDNA. Resolving them is of great importance for the authentication of the identified sequences. Most common informative targets of aDNA are presented and among them mtDNA, and the sequences localized within nuclear DNA. The first, as well as the most important findings in the field are mentioned.

KEY WORDS   molecular biology, PCR, restriction analysis, sequencing, aDNA, mtDNA, Neandertals

Introduction

No other organism's evolution is as interesting as our own, and there are those of us who desire to get to know the origin and history of the genus we belong to. Will it be possible to accomplish this in the near or only in the distant future? Probably not in every detail: the 'life variant' matter making up the universe seems to be too fragile to remain as traces of passed generations.

Until recently all records of human and few other species' past presence on Earth were identified by means of the comparative morphology of fossils, the remains left and retained due to the specific conditions of the environment.

'Trapping the past' is the term which describes what molecular evolutionists as well as molecular anthropologists attempt to do, and what paleontologists do every day. The difference between them consists in the fact that they are interested in different kinds of details within fossilized remains. The kind of information the paleontologist is looking for is reduced to morphological details in describing appearance, and consequently behavior or even habits of a studied organism. By contrast, representatives of a completely new branch of science follow the primary structure of the most informative macromolecule, i.e., deoxyribonucleic acid (DNA). The advent of molecular genetic characterization of prehistoric and historic individuals, and even populations, was made possible by two contemporaneous discoveries. The first one was authored by Mullis in 1983 (see MULLIS and FALOONA [1987]). He proposed an elegant method, *polymerase chain reaction*, which enzymatically amplifies a low number of DNA copies into millions, available for further molecular analysis. The other, by PÄÄBO [1986], showed that nucleic acids found in remains of ancient specimens were frequently preserved well enough to be recovered and analyzed. Almost at the same time, phylogenetically informative sequences of mitochondrial DNA were isolated from a museum specimen of *Equus quagga* and retrieved for the first time by HIGUCHI *et al.* [1984].

What happens to the nucleic acid after the death of an organism?

The term 'ancient nucleic acid' (aDNA or aRNA) covers any trace or bulk of molecules originating from dead organisms or their parts or extracorporeally encountered DNA or RNA. Probably all biological macromolecules that form living organisms begin to decompose spontaneously after death and the rate of the process is dependent on environmental characteristics and time. At neutral pH and 100°C egg-white lysozyme has half-life of only a few minutes, mainly due to deamidation of asparagine residues [AHERN and KLIBANOV 1985]. The lower the temperature, the longer the half-life; with every 10°C-drop, a threefold decrease in reaction rate is observed. Nucleic acids, together with other macromolecules, undergo spontaneous degradation and RNA, in particular, is very susceptible to it. DNA is much more resistant due to the lack of 2'-OH group in the sugar. In single-stranded DNA, each cytosine is converted to uracil with a half-life of about 200 years. A double strand protects extremely well against deamination and the rate of hydrolytic cytosine alteration was described as only 0.5% of that when present in a single strand. This extends the half-life of the compound under the same defined circumstances to at least 30 kyr [FREDERICO *et al.* 1990]. The sites of hydrolytic and oxidative attack against DNA's structural integrity are shown in Fig. 1.

After an organism dies the digestion of its macromolecules by endogenous nucleases begins. This results in partial degradation of DNA which is a mecha-
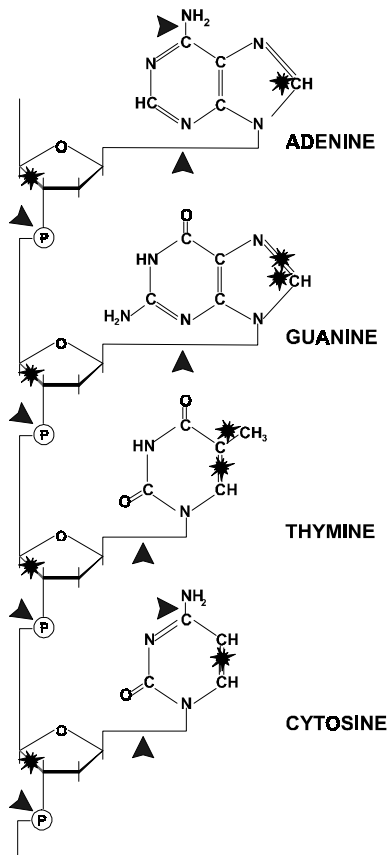
Fig. 1. The scheme of the DNA structure and the localization of possible alterations. Sites susceptible to hydrolytic attack are marked with *arrows* and those susceptible to oxidation are marked with *asterisks*. The main site of destruction is located between sugar and purine.

nism leading to the formation of basic structural units, i.e., mononucleotides. The time interval between the presence of native DNA molecules and their complete destruction depends on the specific physical and chemical nature of the environment they interact with. Some specimens may stand the test of time for centuries or even millennia. The information enclosed in aDNA may be inaccessible due to particular obstacles preventing the process of ancient nucleic acid retrieval. The faster the rate of dehydration of the specimen and the lower the temperature and/or the higher the salt concentration in its close environment, the longer is the retention of the original shape and condition of aDNA. Such factors may restrict or even inhibit the destructive activity of endonucleases. Also, during isolation, amplification, restriction analysis or sequencing (if they are not carefully enough carried out) the structure of aDNA may be altered (Fig. 2), which is frequently observed [KOLMAN and TUROSS 2000].

Depurination, deamination and oxidation (as shown in Fig. 1) occur and disrupt the process of aDNA retrieval. Apurinic sites are lesions that block polymerization of a newly synthesized strand. Besides, the factors which affect and destroy the molecule, and thus dramatically limit the number of fragments to be amplified, are also the factors influencing PCR, e.g., hydantoins (Fig. 2). Even if an adequate number of long enough fragments to be amplified is present in the sample and lack of inhibition of PCR allows amplification of the template, some limitations of the polymerization reaction are still observed. Deaminated nucleotides, e.g., uracil, a product of cytosine modification, may affect the sequence expected since it is being inserted into the synthesized amplicon structure [PÄÄBO 1989].

The presence of oxygen and water is a source of oxygen free radicals. Even their traces are destructive to DNA's molecular integrity, and the initiation of the disintegration process is only a function of environmental features and time. Probably one of the main reasons why the nucleus (as an organelle) was
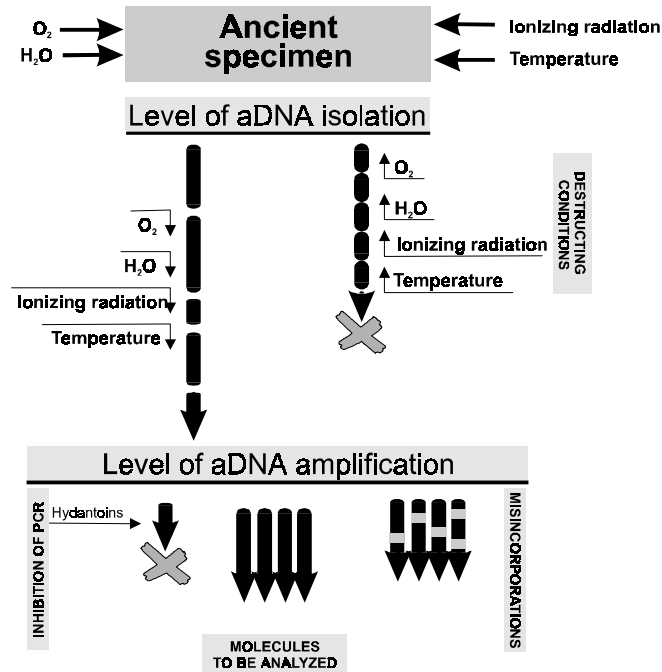
Fig. 2.   Possible features influencing ancient DNA retrieval.

selected early in the evolutionary process was the need to protect molecules located inside against attack of oxygen (mostly in the form of free radicals). The shift of oxygen metabolism to mitochondria resulted in very poor oxygenation of the nucleus, which lacks a detectable O_2 metabolism, thus giving protection to nucleic acids. When an organism dies, the fragile structure of the formerly separated macromolecules becomes exposed to a high concentration of oxygen.

A number of oxidation products of aDNA are known [LINDAHL 1993; HÖSS *et al*. 1996], and among them, 8-hydroxyguanine base-pairs with adenine misincorporated, which generates false results during sequencing of amplified fragments [KASAI and NISHIMURA 1984].

How to get results from retrieving aDNA: The procedure

Research into ancient DNA, i.e., its recovery and genetic study of the primary structure obtained, employs deceptively simple methods. In general, a three-step procedure is applied: *The first* step involves extraction of the remaining fragments of nucleic acids usually from ancient bones, teeth or well preserved soft tissues, e.g., from mummified bodies. Three types of cells can be a source of DNA isolated from bones. These are mainly bone-eroding cells (i.e., osteocytes, osteoblasts, osteoclasts) and sparse blood cells and epithelial cells from the Haversian system [HERRMANN and HUMMEL 1993]. Recently fecal material (coprolite) was used as a source
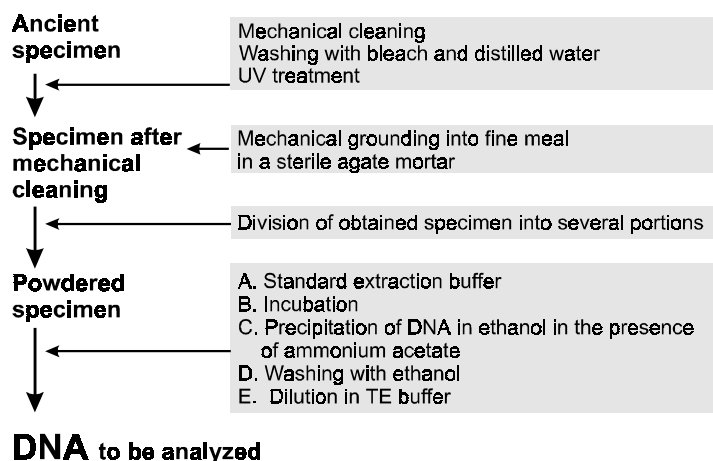
Fig. 3.   Main steps of standard isolation procedure of ancient DNA.

of DNA from both the defecator and its diet [HOFREITER *et al*. 2000]. *The second* step involves amplification of the target sequence chosen. PCR, as the most powerful method, allows collection of a large enough number of copies of the sequence to be analyzed and identified. *The third* step concludes the whole procedure and should guarantee identification of the primary structure through restriction analysis or direct sequencing. The superficial simplicity of the procedure belies its true complexity and the risks for researchers who carry out identification of aDNA sequences. Fig. 3 shows the standard isolation procedure. Principles involved in the three methods applied during DNA retrieval are shown in Fig. 4.

Methods used to verify the macromolecular content of the remains

Some ancient specimens are extremely well protected against the effect of dam-aging environmental factors, but most are not. Thus, only a minute portion of aDNA withdrawn from the remains may be amplified by PCR. Most of the molecules isolated from ancient biological specimens range between 100 and 500 base pairs (bp) in size, although longer fragments can sometimes be identified [PÄÄBO 1989; HANDT *et al*. 1994]. Several protocols which allow screening of samples to determine which of them is well-enough preserved to undergo DNA extraction have been described. Amongst existing methods of sample justification are indirect and direct ones, and are used to verify content (Fig. 5). The indirect methods result in identification of the degree of destruction of amino acid composition and the rate of amino acid *racemization* [POINAR *et al*. 1996]. The rate of conversion from the L to D form of aspartic acid (Asp) is one of the fastest among amino acids, and is characterized by constant activation energy similar to that of the depuration process. Standardization of the
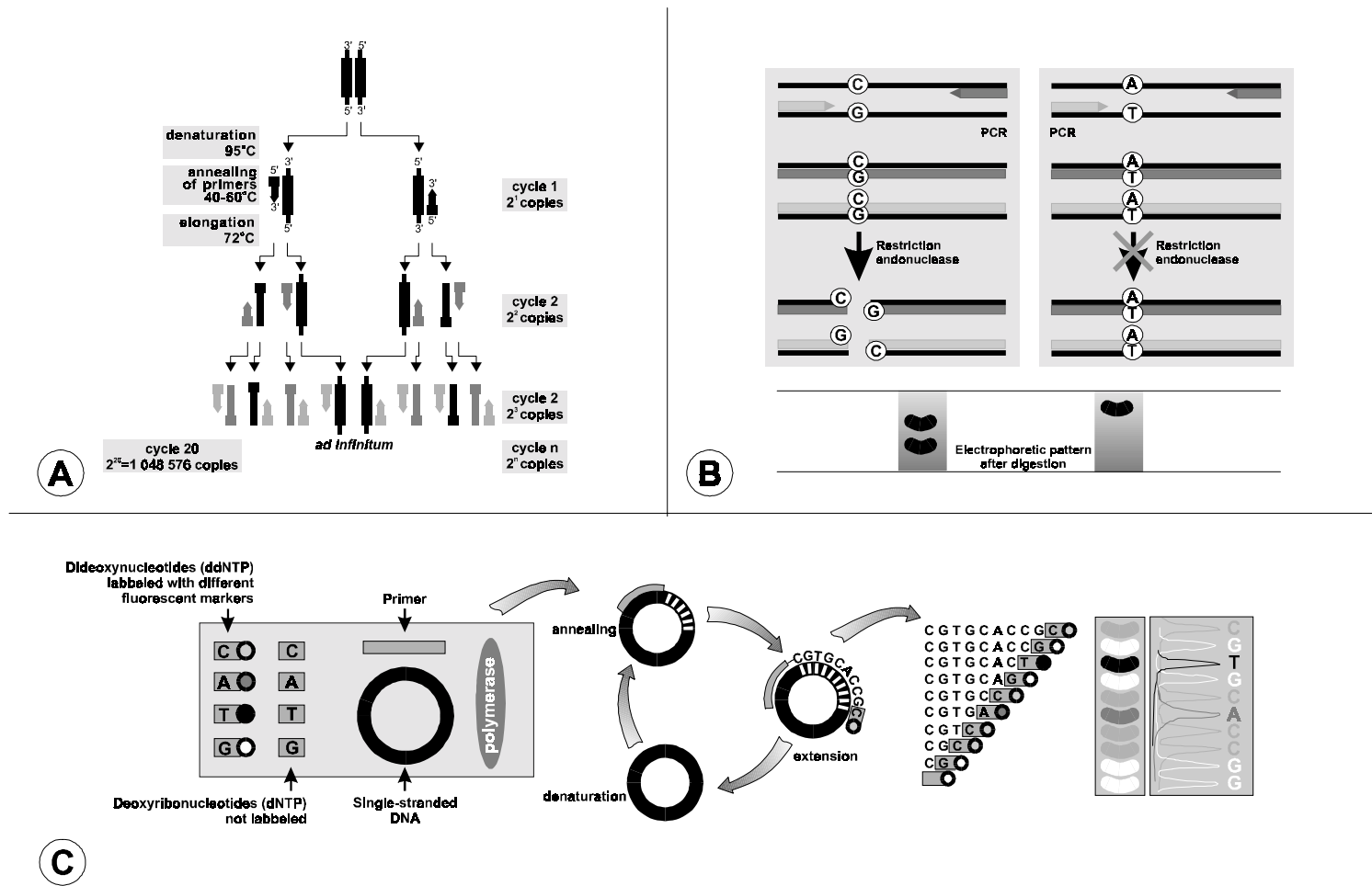
Fig. 4. Methods of basic significance for studies on ancient nucleic acids: (A) PCR, (B) restriction analysis, and (C) sequencing. For understanding the details see BROWN [1999, 2001].
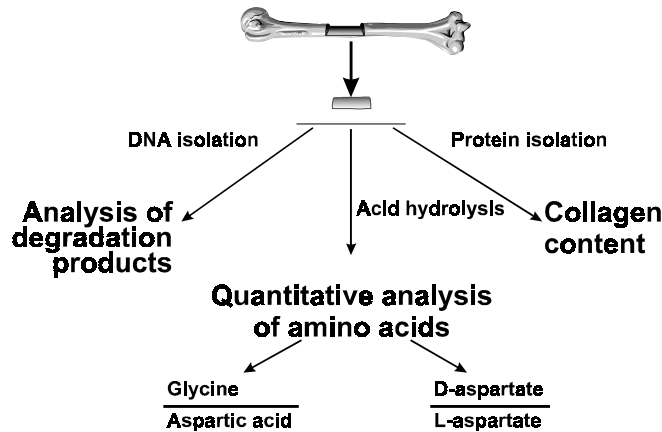
Fig. 5. Scheme of procedures to be undertaken when verifying macromolecular content of an ancient specimen before DNA extraction.

original method allows one to approximate the average length of aDNA degradation products. If the value of log(D/L) exceeds 0.08 there is no need to attempt DNA extraction since the length of degraded DNA fragments will be well below 100 bp.

Another indirect method is realized through evaluation of collagen-type debris content in bone which seems to be a good indicator of macromolecular preservation [BROWN *et al.* 1988]. Results of collagen extraction from the Mezmaiskaya Neandertal rib fragment amounting to 22% of the average obtained from a modern bone with carbon content close to 42% and nitrogen content close to 15% are a good recommendation for the condition of prehistoric remains [OVCHINNIKOV *et al.* 2000; DENIRO 1985]. Those results suggest a low level of diagenetic modification and a high probability of getting a relatively high extraction yield of ancient DNA.

Direct methods applied to verify the degree of DNA damage are based on the identification of its degradation products. HÖSS *et al.* [1996] found a correlation between the amount of products formed as a result of oxidation of a DNA molecule [DIZDAROGLU 1991] and the amount of fragments long enough to be amplified. When using gas chromatography and mass spectrometry at least eight oxidation products were identified quantitatively. Amongst the compounds assayed, hydantoin derivatives of pyrimidines were dominant and their amount correlated inversely with the ability to amplify ancient sequences. 5-hydroxy-5-methylhydantoin (5-OH-5-MeHyd) and 5-hydroxyhydantoin (5-OH-Hyd) which are the products of thymine oxidation (e.g., as an effect of gamma irradiation) [BREIMER and LINDAHL 1985] are quantitatively identified in all ancient DNA samples studied. Consequently, if the concentration of a derivative, e.g., 5-OH-5-MeHyd, exceeds 5 nmol, the isolation of aDNA fails. 5-OH-5-MeHyd usually predominates among derivatives present in an

examined specimen. However, no correlation exists between the main product of purines oxidation, i.e., 8-hydroxy-guanine (8-OH-Gua), which is in fact formed due to hydroxyl attack [DIZDA-ROGLU 1992], and the availability of sequences for analysis.

Controls and criteria for aDNA authenticity

Previous studies have shown that DNA as a molecular and informative entity can persist in the remains of ancient organisms, and that it is much better preserved in the bones and teeth than in soft tissues [HIGUCHI et al. 1984; PÄÄBO 1986, 1989]. A number of methods were proposed [e.g., LAWLOR et al. 1991] for successful isolation and extraction of DNA from the remains of living organisms, including fossils of the humans who inhabited Earth in the past [HERRMANN and HUMMEL 1993].

Identification of sample contamination has revealed one of the most critical issues in ancient DNA research. It presents extreme technical difficulties since only minute amounts are available, molecules are degraded, and a high risk of contamination exists. Isolates of nucleic acid from ancient material yield 1-5% of molecules coming from the tissues of contemporary individuals. PCR sensitivity can be a potential problem or even a major limitation, since a false positive signal may arise from even a very low amount of contaminating molecules (e.g., amplicons from previous reactions). The need to authenticate results became obvious when the results of a number of high-profile studies were shown to be unrepeatable [AUSTIN et al. 1997a,b; STANKIEWICZ

et al. 1998]. Precautions should be taken during aDNA analysis mainly to avoid contamination by contemporary molecules, which are pervasive in the closer and farther environment. Researchers who try to obtain aDNA, especially from human tissues, should be aware of the fact that in the material they work with only a minute amount of nucleic acid has persisted since the death of an organism, whereas a large quantity of contaminating DNA, both from outside and inside the laboratory, enters the sample. Studied specimens may be enriched with modern DNA or very frequently with bacterial DNA [SIDOW et al. 1991]. The extraction of DNA from the biological material of various other species is much simpler to perform. *The first* precaution is to carry out the extraction and preparation in a lab in which it is completely impossible to have contact with contemporary DNA. The primary obstacle in the analysis of ancient nucleic acid is the difficulty in confirming its existence, especially mtDNA. Although it is commonly recommended and accepted that two separate laboratories should carry out the same preparation schedule and compare results, limitations of that method are described [KRINGS et al. 1997]. The *second* precaution against modern contaminating molecules there is the routine provision of a sterile environment, involving UV irradiation of lab facilities, washing of used equipment with oxidizing agents, as well as protecting the sample under examination from contamination with the researcher's DNA by the wearing of face shields and special clothing [HANDT et al. 1994].

In respect of nuclear DNA, HERR-MANN and HUMMEL [1997] presented

an elegant experimental design based on the analysis of a short tandem repeat (STR). Multiple extractions and PCR controls should be performed to enable detection of sporadic or a low copy number of contaminating molecules [HANDT et al. 1996; COOPER 1994]. However, results should be repeatable from the same as well as different DNA extracts from a specimen. Two different labs should perform the same set of samples. In addition, different primer sets should be used to increase the chance of detection, e.g., foreign mitochondrial gene insertion to nucleus [GREENWOOD et al. 1999], or contamination by a PCR by-product. PCR sequences obtained must be verified by the cloning of amplicons. This procedure allows determining the ratio of endogenous to exogenous sequences, to identify damage-induced errors or any inserts (Fig. 6).
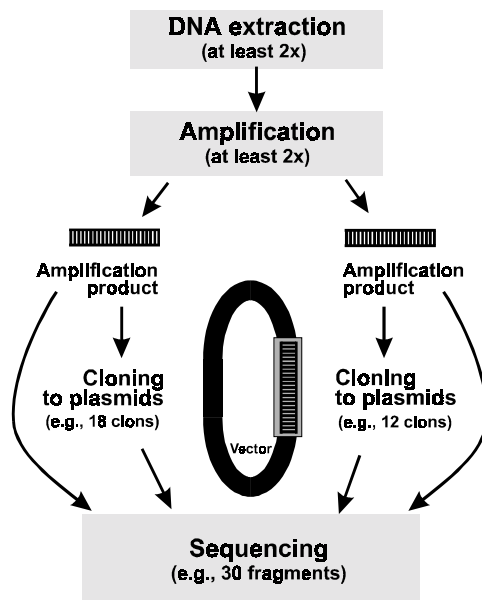

Fig. 6. Methodological precautions taken during ancient DNA retrieval.

It is extremely difficult to exclude contamination if the genuine template is present in a small number of copies and if their amplification is possible only by some of the reactions. The number of template molecules present in the sample determines the choice of the proper procedure of amplification. The fact that 1000 or more molecules are present in a specimen excludes the need to verify misincorporations in the amplified sequence. The possibility to measure the number of copies is offered by the use of competitive PCR or real-time quantitative PCR [HANDT et al. 1996].

Some authors [KOLMAN and TUROSS 2000], however, argue that the above-listed standard precautionary measures are insufficient to identify contamination, and if identical sequences are found in ancient specimens as well as in control DNA it precludes convincing proof that the extracted DNA is of ancient origin.

Comparison of mtDNA suggests no Neandertal contribution to the ancestry of modern humans

Although only a few published papers describe the results of the analysis of mtDNA isolated from the remains of three Neandertals, the sequences compared have clearly shown that Neandertals did not contribute mitochondrial DNA to modern humans. The very first paper on the Neandertal mtDNA appeared in 1997 [KRINGS et al. 1997] as one of the results obtained under an interdisciplinary project carried out by the Rhineland Museum since 1991. The research focused on the Neandertal type specimen found in the Neander Valley near Düsseldorf in 1856, Germany (con-

sidered to be 100-30 kyr). As part of this project, a 3.5 g section of the right *humerus* was removed and used for analysis of the first hypervariable region (HVR1) of mtDNA. Almost two years later the next paper [KRINGS *et al*. 1999] offered the results of sequence analysis of the second hypervariable region (HVR2). Besides the comparison of both mtDNA sequences belonging to humans, Neandertals, and chimpanzees, all positions where insertions/deletions between the studied species occurred were excluded. Altogether 600 positions were identified encompassing that from 16,024 to 16,365 and from 73 to 340, excluding 10 of them.

The other study described the analysis of mtDNA isolated from the remains of a Neandertal infant found near Mezmayskaya, Caucasus (29 kyr) [OVCHINNIKOV *et al*. 2000]. A 345-bp fragment of HVR1 was sequenced and compared to the previously obtained one [KRINGS *et al*. 1997, 1999]. The third Neandertal remains (those found in Vindija, Croatia and dated at 42 kyr) were also analyzed with the purpose of finding the primary structure of HVR1 and HVR2. Although the sample size for the Neandertals is far from desired, it is noteworthy that the probability for obtaining the deepest genetic divergence within this sample was 50%, since in a random mating population it equals $(n − 1) / (n + 1)$; were n is a number of representatives of the population under study [KRINGS *et al*. 2000]. The number of sequences of modern humans contained in the database exceeded 5,000. All of the above made it possible to estimate genetic diversity not only between Neandertals and modern humans, but also amongst Neandertals themselves [KRINGS *et al*.

2000]. A distance-base tree which exhibits the relation between the analyzed sequences as well as the percentage of the positions identified as changed are usually applied to study distances between studied groups. The results obtained enriched by comparison with apes (chimpanzees and gorillas) suggest that Neandertals were more similar to modern humans than the apes, since the diversity within the mtDNA molecule was more like that of ours and far different from that of apes.

Low species-wide diversity of the Neandertal mtDNA similar to contemporary humans and high degree of sequence diversity compared to modern humans may suggest [KRINGS *et al*. 1997, 1999, 2000; OVCHINNIKOV *et al*. 2000] that:
− It is unlikely to be a representation of an ancestral lineage of modern human mtDNA;
− It is likely to represent the same type of expansion as contemporary human population [CANN *et al*. 1987];
− It is likely that the lineage leading to Neandertal man diverged approximately 500 kyr;
− Interbreeding is not excluded [NORDBORG 1998], but Neandertals went extinct without contributing their mtDNA to modern human gene pool.

The role of particular DNA sequences as markers of the origin of particular populations

Mitochondrial DNA

Until recently most of the research on ancient nucleic acids was performed on mtDNA isolated at the very beginning

of aDNA studies from soft tissues of ancient remains [HIGUCHI *et al*. 1984; PÄÄBO 1985*a,b*; LAWLOR *et al*. 1991]. Then, several authors confirmed it was also possible to extract the macromolecule from bones [HAGELBERG *et al*. 1989; HORAI *et al*. 1989].

Since the beginning of aDNA studies the mtDNA molecule was considered to be a rich source of molecular information for evolutionary studies due to its exceptional characteristics. In almost all cells of *eukaryotic* organisms the molecule is represented by high number of copies (100-1000 mitochondria per cell) in contrast to nuclear DNA. This favors survival and availability of mtDNA, which in turn makes it well suited for research. The very well characterized human sequence of 16, 569 base pairs apparently lacks recombination [OLIVO *et al*. 1983], but exhibits a high rate of substitution [BROWN *et al*. 1979] and represents maternal mode of inheritance [GILES *et al*. 1980]. Such features are advantageous for evolutionary and phylogenetic studies. It is noteworthy that almost all studies on evolution of *H. sapiens* and based on variation of mtDNA primary structure have been confined to a small region of the whole molecule, amounting to 7%, and located within highly variable regions of the mitochondrial genome [INGMAN *et al*. 2000]. Although the above features are advantageous some complications due to extreme variation in substitution rate are frequently encountered during studies. Parallel mutations found within informative sequences [TAMURA and NEI 1993] result in difficulties while estimating genetic distance, thus making phylogenetic inferences doubtful [MADDISON *et al*. 1992].

Y chromosome

Although mtDNA is a very rich source of information, it only reflects the genetic history of the female though, fortunately, it is not the only locus for retrieval in evolutionary studies involving humans. Another location within our nuclear genome suitable for evolutionary purpose, i.e., Y chromosome, was proposed in 1995 [PÄÄBO 1995; DORIT *et al*. 1995]. Almost 60 Mb nonrecombining part of Y chromosome, with the exception of the two pseudoautosomal regions, is transmitted strictly from father to son [JOBLING and TYLER-SMITH 1995]. This characteristic renders the chromosome to be one of the most versatile locations of haplotypic genotyping system within human genome. These sequences were used to follow migration routes of our male ancestors from the recent past [FOSTER *et al*. 1998] through centuries [HAMMER *et al*. 2000], or the history of evolution of species including humans [HAMMER *et al*. 1998]. Moreover, Y chromosome sequences were used to estimate the time which passed from the most recent common ancestor [BERTRANPETIT 2000; SHEN *et al*. 2000]. Among the sequences/polymorphisms studied on Y chromosome, are biallelic markers characterized by low mutation rate, which represent unique mutation events, moderately-fast evolving sequences with an average mutation frequency of 0.2% per generation and fast evolving loci with mutation frequency of 6-11% per generation [DE KNIJFF 2000].

Microsatellites

Some DNA regions are extremely susceptible *in vivo* to modifications such as insertions and deletions. Among them

are microsatellites also known as short tandem repeats (STRs), i.e., simple 1-6 base pair in length repeats that are found in all *eukaryotic* organisms [HAMADA *et al*. 1982]. The most frequent motif is $(GT)_n$ were n equals 10-30 [BRINK-MANN *et al*. 1996]. These are widely dispersed within the genome. The most abundant microsatellite sequences in the human genome are copied 50,000 to 100,000 times [MIESFELD *et al*. 1981]. Allelic variation of microsatellites, mainly due to the number of tandem repeats, is responsible for a significant degree of polymorphism [WEBER and MAY 1989]. There are no methodological problems in evaluating of the length of an individual microsatellite. Cloning requires a set of primers directed to unique flanking sequences, which usually amplify 100-300 bp fragments. All the above characteristics, together with a high degree of heterozygosity (up to 90%), contribute to the generation of allele diversity which is very useful in searching for evolutionary events [WEBER and WONG 1993]. The sequences may also be used as genetic markers, e.g., for genetic mapping/linkage analysis [WEBER and MAY 1989], for detecting genomic instability in tumors [LOEB 1994] or in forensics [WIEGAND *et al*. 1993].

Final remarks

One of the first studies whose purpose it was to explain phylogenetic relationships was reported by HIGUCHI *et al*. [1984] who confirmed that a representative of an extinct species (a museum specimen from which mtDNA was isolated) could be classified under genus *Equus*. PÄÄBO [1985*a,b*] was the first to

attempt to recover and analyze ancient human DNA and was soon followed by others [DORAN *et al*. 1986; HAGELBERG *et al*. 1989].

One should emphasize here the advantages of analysis of the primary structure of DNA sequences properly isolated from ancient remains. Until recently, the frequencies of different genetic variants were examined through the mobility in the gel of proteins, expression products of particular alleles. Nowadays, direct determination of the sequences of nucleotides forming a gene and its surroundings are available. Now we can not only reveal a difference between two alleles, but also determine how they differ from each other. The number of differences reflects the time that has passed between the most recent common ancestor and the groups under study. It also reflects the history of a particular group in the sense of changing size – reduction in size means reduction in nucleotide diversity, e.g., in speciation events. Comparative studies also allow us to reveal phylogenetic relationships of extinct animals. One can also reveal from coprolites the genetic characteristics of both the consumer and his diet, and thus his behavior as well as flora and fauna present in that time.

In conclusion, it should be noted that the data obtained are of no value if genetic and anthropological descriptions are contradictory.

References

AHERN T., A.M. KLIBANOV, 1985, *The mechanism of irreversible enzyme inactivation at 100C*, Science, **228**, 1280-84

AUSTIN J.J., A.J. ROSS, A.B. SMITH, R.A. FORTEY, R.H. THOMAS, 1997*a*, *Problems with reproducibility – does geologically ancient*

DNA survive in amber-preserved insects? Proc. R. Soc. Lond. B., **264**, 467-474

Austin J.J., A.B. Smith, R.H. Thomas, 1997*b*, *Paleontology in a molecular world: The search for the authentic ancient DNA*, Trends Ecol. Evol., **12**, 303-306

Bertranpetit J., 2000, *Genome, diversity and origins: The Y chromosome as a storyteller*, Proc. Natl. Acad. Sci. USA, **97**, 6927-29

Breimer L.H., T. Lindahl, 1985, *Thymine lesions produced by ionizing radiation in double-stranded DNA*, Biochemistry, **24**, 4018-22

Brinkmann B., A. Sajantila, H.W. Goedde, H. Matsumoto, K. Nishi, P. Wiegand, 1996, *Population genetic comparisons among eight populations using allele frequency and sequence data from three microsatellite loci*, Eur. J. Hum. Genet., **4**, 175-182

Brown T.A., 1999, *Genomes*, BIOS Scientific Publ., Oxford

Brown T.A., 2001, *Genomy*, (Polish edition), Wyd. Naukowe PWN, Warszawa

Brown T.A., D. Nelson, J.S. Vogel, J.R. Southon, 1988, *Improved collagen extraction by modified Longin method*, Radiocarbon, **30**, 171-177

Brown W.M., M. George Jr., A.C. Wilson, 1979, *Rapid evolution of animal mitochondrial DNA*, Proc. Natl. Acad. Sci. USA, **76**, 1967-71

Cann R.L., M. Stoneking, A.C. Wilson, 1987, *Mitochondrial DNA and human evolution*, Nature, **325**, 31-36

Cooper A., 1994, *DNA from museum specimens*, [in:] *Ancient DNA*, B. Herrmann & S. Hummel (eds.), Springer Verlag, New York, pp. 149-165

de Knijff P., 2000, *Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on human Y chromosome*, Am. J. Hum. Genet., **67**, 1055-61

DeNiro M.J., 1985, *Postmortem preservation and alteration of in vivo bone collagen isotope ratios in relation to palaeodietary reconstruction*, Nature, **317**, 806-809

Dizdaroglu M., 1991, *Chemical determination of free radical-induced damage to DNA*, Free Radical Biol. Med., **10**, 225-242

Dizdaroglu M., 1992, *Oxidative damage to DNA in mammalian chromatin*, Mutat. Res., **275**, 331-342

Doran G.H., D.N. Dickel, W.A. Ballinger Jr., F.O. Agee, P.J. Laipis, W.W. Hauswirth, 1986, *8000 year old human brain tissue: Anatomical, cellular and molecular analysis*, Nature, **323**, 803-806

Dorit R.L., H. Akashi, W. Gilbert, 1995, *Absence of polymorphism at the ZFY locus on the human Y chromosome*, Science, **268**, 1183-85

Foster E.A., M.A. Jobling, P.G. Taylor, P. Donnelly, *et al.*, 1998, *Jefferson fathered slaves last child*, Nature, **396**, 27-28

Frederico L.A., T.A. Kunkel, B.R. Shaw, 1990, *A sensitive genetic assay for the detection of cytosine deamination: Determination of rate constants and the activation energy*, Biochemistry **29**, 2532-37

Giles R.E., H. Blanc, H.M. Cann, D.C. Wallace, 1980, *Maternal inheritance of human mitochondrial DNA*, Proc. Natl. Acad. Sci. USA, **77**, 6715-19

Greenwood A.D., C. Capelli, G. Possnert, S. Pääbo, 1999, *Nuclear DNA sequences from late Pleistocene megafauna*, Mol. Biol. Evol., **16**, 1466-73

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

Hamada H., M.G. Petrino, T. Kakunaga, 1982, *A novel repeated element with Z-DNA forming potential is widely found inevolutionary diverse eukaryotic genomes*, Proc. Natl. Acad. Sci. USA, **79**, 6465-69

Hammer M.F., T. Karafet, A. Rasanayagam, E.T. Wood, *et al.*, 1998, *Out of Africa and back again: Nested cladistic analysis of human Y chromosome variation*, Mol. Biol. Evol., **15**, 427-441

Hammer M.F., A.J. Redd, E.T. Wood, M.R. Bonner, *et al.*, 2000, *Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes*, Proc. Natl. Acad. Sci. USA, **97**, 6769-74

Handt O., M. Höss, M. Krings, S. Pääbo, 1994, *Ancient DNA: Methodological challenges*, Experientia, **50**, 524-529

Handt O., M. Krings, R.H. Ward, S. Pääbo, 1996, *The retrieval of ancient human DNA sequences*, Am. J. Hum. Genet., **59**, 368-376

Herrmann B., S. Hummel, 1993, *Recovery and analysis of genetic material from paleontological, archeological museum, and forensic specimens*, [in:] *Ancient DNA*, B. Herrmann & S. Hummel (eds.), Springer Verlag, New York, pp. 1-12

Herrmann B., S. Hummel, 1997, *Genetic analysis of past populations by DNA studies*, [in:]

Advances in research on DNA polymorphisms, ISFH Hakone Symposium Program Committee (eds.), Toyoshoten, Tokyo, pp. 33-47

HIGUCHI R., B. BOWMAN, M. FREIBERGER, O.A. RYDER, A.C. WILSON, 1984, *DNA sequences from a quagga an extinct member of the horse family*, Nature, **312**, 282-284

HOFREITER M., H. N. POINAR, W.G. SPAULDING, K. BAUER, P.S. MARTIN, G. POSSNERT, S. PÄÄBO, 2000, *A molecular analysis of ground sloth diet through the last glaciation*, Mol. Ecol., **9**, 1975-84

HORAI S., K. HAYASAKA, K. MURAYAMA, N. WATE, H. KOIKE, N. NAKAI, 1989, *DNA amplification from ancient human skeletal remains and their sequence analysis*, Proc. Jap. Acad., **65**, 229-233

HÖSS M., P. JARUGA, T.H. ZASTAWNY, M. DIZDAROGLU, S. PÄÄBO, 1996, *DNA damage and DNA sequence retrieval from ancient tissue*, Nucleic Acid Res., **24**, 1304-07

INGMAN M., H. KAESSMANN, S. PÄÄBO, U. GYLLENSTEN, 2000, *Mitochondrial genome variation and the origin of modern humans*, Nature, **408**, 708-713

JOBLING M.A., C. TYLER-SMITH, 1995, *Fathers and sons: The Y-chromosome and human evolution*, Trends Genet., **11**, 449-456

KASAI H., S. NISHIMURA, 1984, *Hydroxylation of deoxyguanosine at the C-8 position by ascorbic acid and other reducing agents*, Nucleic Acid Res., **12**, 2137-45

KOLMAN C.J., N. TUROSS, 2000, *Ancient DNA analysis of human populations*, Am. J. Phys. Anthropol., **111**, 5-23

KRINGS M., C. CAPELLI, F. TSCHENTSCHER, H. GEISERT, *et al.*, 2000, *A view of Neandertal genetic diversity*, Nature Genet., **26**, 144-146

KRINGS M., H. GEISERT, R.W. SCHMITZ, H. KRAINITZKI, S. PÄÄBO, 1999, *DNA sequence of the mitochondrial hypervariable region II from the Neandertal type specimen*, Proc. Natl. Acad. Sci. USA, **96**, 5581-85

KRINGS M., A. STONE, R.W. SCHMITZ, H. KRAINITZKI, M. STONEKING, S. PÄÄBO, 1997, *Neandertal DNA sequences and the origin of modern humans*, Cell, **90**, 19-30

LAWLOR D.A., C.D. DICKEL, W.W. HAUSWIRTH, P. PARHAM, 1991, *Ancient HLA genes from 7,500-year old archaeological remains*, Nature, **349**, 785-788

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

LOEB L.A., 1994, *Microsatellite instability: Marker of a mutator phenotype in cancer*, Cancer Res., **54**, 5059-63

MADDISON D.R., M. RUVOLO, D.L. SWOFFORD, 1992, *Geographic origins of human mitochondrial DNA: Phylogenetic evidence from control region sequences*, Syst. Biol., **41**, 111-124

MIESFELD R., M. KRISTAL, N. ARNHEIM, 1981, *A member of a new repeated sequence family which is conserved throughout eucaryotic evolution is founded between the human delta globin genes*, Nucleic Acid Res., **9**, 5931-47

MULLIS K.B., F.A. FALOONA, 1987, *Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction*, Methods Enzymol., **155**, 335-350

NORDBORG M., 1998, *On the probability of Neanderthal ancestry*, Am. J. Hum. Genet., **63**, 1237-40

OLIVO P.D., M.J. VAN DE WALLE, P.J. LAIPIS, W.W. HOUSEWIRTH, 1983, *Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial DNA D-loop*, Nature, **306**, 400-402

OVCHINNIKOV I.V., A. GÖTHERSTRÖM, G.P. ROMANOWA, V.M. KHARITONOV, K. LIDEN, W. GOODWIN, 2000, *Molecular analysis of Neanderthal DNA from the northern Caucasus*, Nature, **404**, 490-493

PÄÄBO S., 1985a, *Preservation of DNA in ancient Egyptian mummies*, J. Arch. Sci., **12**, 411-417

PÄÄBO S., 1985b, *Molecular cloning of ancient Egyptian mummy DNA*, Nature, **314**, 644-645

PÄÄBO S., 1986, *Molecular genetic investigations of ancient human remains*, Cold Spring Harbour Symp. Quant. Biol., **51**, 441-446

PÄÄBO S., 1989, *Ancient DNA: Extraction, characterization, molecular cloning, and enzymatic amplification*, Proc. Natl. Acad. Sci. USA, **86**, 1939-1943

PÄÄBO S., 1995, *The Y chromosome and the origin of all of us*, Science, **268**, 1141-42

POINAR H.N., M. HÖSS, J.L. BADA, S. PÄÄBO, 1996, *Amino acid racemization and the preservation of ancient DNA*, Science, **272**, 864-866

SHEN P., F. WANG, P.A. UNDERHILL, C. FRANCO, *et al.*, 2000, *Population genetic implications from sequence variation in four Y chromosomal genes*, Proc. Natl. Acad. Sci. USA, **97**, 7354-59

SIDOW A., A.C. WILSON, S. PÄÄBO, 1991, *Bacterial DNA in Clarkia fossils*, Philos. Trans. R. Soc. Lond. B., **333**, 429-433

STANKIEWICZ B., H. POINAR, D. BRIGGS, R. EVERSHED, G. POINAR, 1998, *Chemical presentation of plants and insects in natural resins*, Proc. R. Soc. Lond. B., **265**, 641-647

TAMURA K., M. NEI, 1993, *Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees*, Mol. Biol. Evol., **10**, 512-526

WEBER J.L., P.E. MAY, 1989, *Abundant class of human DNA polymorphism which can be typed using polymerase chain reaction*, Am. J. Hum. Genet., **44**, 388-396

WEBER J.L., C. WONG, 1993, *Mutation of human short tandem repeats*, Hum. Mol. Genet., **2**, 1123-28

WIEGAND P., T. BAJANOWSKI, B. BRINKMANN, 1993, *DNA typing of debris from fingernails*, Int. J. Legal Med., **106**, 81-83

Streszczenie

Ocena morfologiczna szczątków kopalnych pozostałych po pokoleniach wcześniej żyjących organizmów stanowiła do niedawna główne źródło informacji o przodkach istot współcześnie żyjących. W następstwie rozwoju wysoko zaawansowanych metod badawczych, takich jak łańcuchowa reakcja polimerazy DNA (PCR) i sekwencjonowanie DNA, pojawiła się możliwość korzystania z nowego, zupełnie różnego jakościowo, obfitego źródła informacji o osobniku i gatunku, do którego należy. PCR jest metodą, dzięki której dowolny niezdegradowany fragment DNA, choćby jednej znalezionej cząsteczki, może być powielony w milionach kopii. Uzyskane kopie umożliwiają sekwencjonowanie, które pozwala odpowiedzieć na najważniejsze pytanie – jaka jest kolejność par zasad w badanym odcinku łańcucha. Powszechnie uważa się, że struktura pierwszorzędowa „natywnego" DNA jest zmienną zależną od czasu, charakterystyczną zarówno dla osobnika jak i gatunku.

Terminem „archaiczny DNA" (aDNA) określane są jego cząsteczki, a raczej ich mniejsze lub większe fragmenty pochodzące z komórek i tworzące się po śmierci organizmu. W pierwszej fazie są one konsekwencją destrukcji chemicznej spowodowanej hydrolitycznym działaniem enzymów nukleolitycznych, uwolnionych z lizosomów. W dalszym etapie, niszczenie struktury chemicznej pierwotnej makrocząsteczki następuje w wyniku działania wszechobecnego tlenu i wody. Najczęstszym przejawem destrukcji jest hydrolityczne odłączenie puryn, rzadziej pirymidyn, oraz grup aminowych. Pod wpływem tlenu dezintegracji ulega pierścień pirymidynowy i purynowy, dezoksyryboza oraz grupa metylowa tyminy. Powstające zmiany wpływają na ciągłość struktury aDNA, co z kolei ogranicza możliwości powielania jego fragmentów, jak to ma miejsce w przypadku odłączania puryn. Produkt utlenienia cytozyny – uracyl jest niewłaściwie rozpoznawany w trakcie syntezy nowego łańcucha, a hydantoiny, produkty utlenienia tyminy są inhibitorami PCR.

Proces otrzymywania i identyfikacji aDNA, tylko z pozoru łatwy, można podzielić na trzy etapy. Pierwszy etap, najtrudniejszy ze względu na zanieczyszczający, wszechobecny tzw. obcy DNA, sprowadza się do izolowania właściwych cząsteczek aDNA. Drugi etap, polega na powieleniu, w wyniku łańcuchowej reakcji polimerazy, nielicznych, wyizolowanych jego cząsteczek. Ostatni etap to identyfikacja struktury pierwszorzędowej aDNA. Przed przystąpieniem do pierwszego etapu, pośrednio lub bezpośrednio ocenia się stopień degradacji chemicznej struktury pozyskanych szczątków. Ocena bezpośrednia obejmuje analizę składu produktów degradacji zasad purynowych i pirymidynowych, wśród których dominują hydantoiny (główna pochodna tyminy – 5-hydroksy-5-metylohydantoina). Stężenie wyższe niż 5 nmoli dyskwalifikuje próbę jako źródło aDNA. Częściej stosuje się analizę pośrednią stopnia uszkodzenia aDNA, która polega na ocenie stopnia racemizacji aminokwasów, np. przemiana L-Asp / D-Asp, ponieważ wartości energii aktywacji tego procesu i reakcji odłączania puryn są porównywalne. Warto podkreślić, że nie wykazano bezpośredniej zależności

pomiędzy wiekiem próby i stopniem racemizacji, który najpewniej zależy od warunków w jakich pozostają szczątki. Jeszcze inną pośrednią metodą oceny stopnia degradacji DNA jest oznaczenie ilościowe kolagenu.

Po przystąpieniu do próby wyizolowania aDNA pojawia się kolejna, najtrudniejsza do pokonania przeszkoda – identyfikacja zanieczyszczenia, którym jest głównie DNA współczesny. Ocenia się, że ze szczątków izolowane jest nie więcej niż 1-5% wyjściowej zawartości DNA. Głównym problemem jest czułość metody PCR. Zazwyczaj ilość zanieczyszczającego DNA, porównywalna z ilością aDNA, jest współpowielana podczas PCR. Należy podejmować wiele działań zapobiegających zanieczyszczaniu badanej próby. *Po pierwsze*, laboratorium gdzie prowadzone są badania musi być odizolowane od możliwości kontaktu z DNA współczesnym. Najlepiej tę samą próbę analizować w dwóch niezależnych ośrodkach, w kilku powtórzeniach, z kilku różnych ekstrakcji. Jeżeli liczba kopii jest mniejsza niż 1000, istnieje potrzeba wielokrotnej weryfikacji struktury pierwszorzędowej powielanego fragmentu. Możliwość powielania długich fragmentów (ponad 500 pz) jest dowodem na obecność niezdegradowanego, tj. współczesnego DNA. *Po drugie*, należy stosować sprzęt jednorazowy, sterylizację szeroko rozumianego otoczenia, gdzie prowadzone są badania (UV, przemywanie w środkach utleniających) oraz zabezpieczenia przed zanieczyszczaniem przez pracownika wykonującego badania (odzież ochronna, maski, rękawiczki).

Z metodyką izolowania aDNA i jego daleko zaawansowaną, poprawną analizą można zapoznać się przeglądając wyniki badań dotyczących informacji genetycznej uzyskanej z próbek zachowanych kości trzech przedstawicieli człowieka neandertalskiego. Analizowano mitochondrialny DNA (mtDNA), a w szczególności jego wysoce zmienne odcinki HVR1 i HVR2. Na podstawie ilości różnic w strukturze pierwszorzędowej badanych odcinków cząsteczek mtDNA neandertalczyków i współczesnego człowieka czas życia wspólnego przodka oszacowano na 500 tys. lat temu, z sugestią, że neandertalczycy najpewniej wymarli bez pozostawienia śladów swego mtDNA w strukturze cząsteczki mtDNA anatomicznie nowoczesnego człowieka. Nie wyklucza to jednak możliwości krzyżowania się przedstawicieli obydwu form. Na podstawie porównania mtDNA trzech przedstawicieli człowieka neandertalskiego można także przypuszczać, że jego populacja reprezentowała podobny typ ekspansji jak populacja człowieka współczesnego.

Pierwsze badania archaicznego DNA przeprowadzano na cząsteczkach znajdujących się w mitochondriach, znacznie później wprowadzono badanie DNA jądrowego jako źródła informacji. mtDNA występuje w komórce w znacznej liczbie kopii, nie ulega rekombinacji, charakteryzuje się wysokim stopniem zmienności i dziedziczony jest wyłącznie po matce. Te cechy zadecydowały, że stał się dobrym materiałem do badań. Podczas gdy mtDNA może być odzwierciedleniem ewolucji kobiet, nierekombinujący fragment chromosomu Y o długości 60 Mpz jest jego odpowiednikiem dla linii męskiej. Wśród fragmentów genomu szczególnie podatnych na występowanie modyfikacji o charakterze delecji i insercji są wielokrotnie powtórzone tzw. odcinki mikrosatelitarne (1-6 pz). Liczba powtórzeń sama w sobie stanowi cechę osobnika lub grupy.

Najważniejsze korzyści płynące z możliwości identyfikacji struktury pierwszorzędowej aDNA są następujące. Kumulacja różnic określa odległość czasową pomiędzy przedstawicielami grup i ich wspólnym przodkiem. Zmniejszenie różnorodności w strukturze pierwszorzędowej, tj. ilości różnic pomiędzy przedstawicielami tej samej grupy wskazuje na czasowe zmniejszenie liczebności, co ma miejsce np. w czasie specjacji. Możliwe jest także ustalenie filogenetycznych zależności form wymarłych.