ANCIENT DNA STUDY ON HUMAN FOSSILS FOUND IN COSTIŞA, ROMANIA, DATING FROM THE BRONZE AGE

GEORGETA CARDOŞ[1,2,3], ALEXANDER RODEWALD[2], ANDREI D. SOFICARU[4]

[1] "Victor Babeş" National Institute, Bucharest, Romania
[2] Institute of Human Biology, University of Hamburg, Hamburg, Germany
[3] University of Bucharest, Bucharest, Romania
[4] "Francisc Rainer" Institute of Anthropology, Bucharest, Romania

Keywords: aDNA, HVRI, HVRII, HumVWA31A
Abstract: Our study has focused on the ancient DNA (aDNA) analysis from some skeletal remains (bones and teeth) of 6 old individuals found at Costişa, Romania, dating from the Bronze Age. We analysed mtDNA (HVRI and HVRII regions) and nuclear DNA (HumVWA31A, DYS393 and DYS392) in order to understand their genetic relationships and amplified a sequence of the Amelogenin gene for identifying their genetic sex. Our results showed that the teeth of 3 old individuals contained PCR amplifiable DNA, all of them being females at the genetic level, in agreement with the anthropological study. The nuclear and mtDNA marker analysis revealed their genetic kinship between each other and with individuals from other human old populations from Romania, dating from the Bronze and Iron Age.

Cuvinte-cheie: ADN vechi, HVRI, HVRII, HumVWA31A
Rezumat: Studiul nostru a fost focalizat pe analiza ADN-ului vechi din ramăşite scheletice (oase si dinţi) fosile ale 6 indivizi din epoca bronzului, descoperite la Costişa, Romania. Am analizat ADN mitocondrial (regiunile HVR I si HVR II) si ADN nuclear (markerii Hum VWA31A, DYS392 si DYS393) pentru a întelege relatiile lor de înrudire genetică şi am amplificat o secvenţă a genei Amelogeninei pentru a identifica sexul genetic al acestor indivizi. Rezultatele noastre au relevat prezenţa de ADN amplificabil prin PCR din dinţii a 3 dintre indivizii luaţi în studiu, toţi cei 3 fiind de sex feminin la nivel genetic, în acord cu studiul antropologic. Analiza markerilor de ADN nuclear si mitocondrial a relevat înrudirea genetică dintre indivizii studiaţi, dar şi cu indivizi din alte populaţii umane vechi de pe teritoriul Romaniei, din epoca bronzului şi a fierului.

INTRODUCTION

The extraction of ancient DNA (aDNA) from fossil remains has opened up a new research area with many implications, providing access to the basic molecules of organism evolution and awarding the possibility of studying evolution at the molecular level over an unlimited time scale. The study of ancient DNA extracted from human fossils has led to a better understanding of the human beings and the evolution of the human genome, as well as the spread of human populations throughout the Earth, confirming or infirming some anthropological data[1].

[1] S. Paabo, *Ancient DNA: Extraction, characterization, molecular cloning, and enzymatic amplification*, Proc. Natl. Acad. Sci. 86, 1989, p. 1939-1943; C. Cattaneo, S. Dimartino, S. Scali, O. Craig, M. Grandi, R.J. Sokol, *Determining the human origin of fragments of burnt bones: a comparative study of histological, immunological and DNA techniques*, Forensic Science International 102, 1999, p. 181-191; A. Gotherstrom, M. J. Collins, A. Angerbjorn and K. Liden, *Bone preservation and DNA amplification*, Archaeometry 44/3, 2002, p. 395-404; Susanne Hummel, *Ancient DNA Typing. Methods, Strategies and Applications.* Springer-Verlag Berlin Heidelberg New York, 2001, p. 159-217; M. Krings, A. Stone, W.R. Schmitz, H. Krainizki, M. Stoneking and S. Päbo, *Neanderthal DNA Sequences and the Origin of Modern Humans,* Cell 90, 1997, p. 21-30.

DNA – "the molecule of life"

The cell is the structural, functional and genetic basic unit of all forms of life. All eukaryotic cells have an external double layer-membrane, which encloses the cytoplasm with nucleus and several types of organelles (e.g., mitochondria, Golgi apparatus, ribosomes, etc.). Nucleus and mitochondria contain the DNA molecules, which encode all the structural and functional traits of the cells and implicitly of the organisms.

DNA is a macromolecule with a double helical structure, built up from four types of nucleotides, each of them containing a sugar, a phosphate group (making up the DNA backbone) and a nucleobase. There are only four nucleobases in the DNA structure: Adenine (A), Guanine (G), Cytosine (C) and Thymine (T) (Fig 1), their succession making up the universal genetic code. The hydrogen bonds between the complementary nucleobases on the two strands stabilize the double helix structure of the DNA molecule[2]. The length of a double stranded DNA molecule is usually expressed in complementary base pairs (bp).

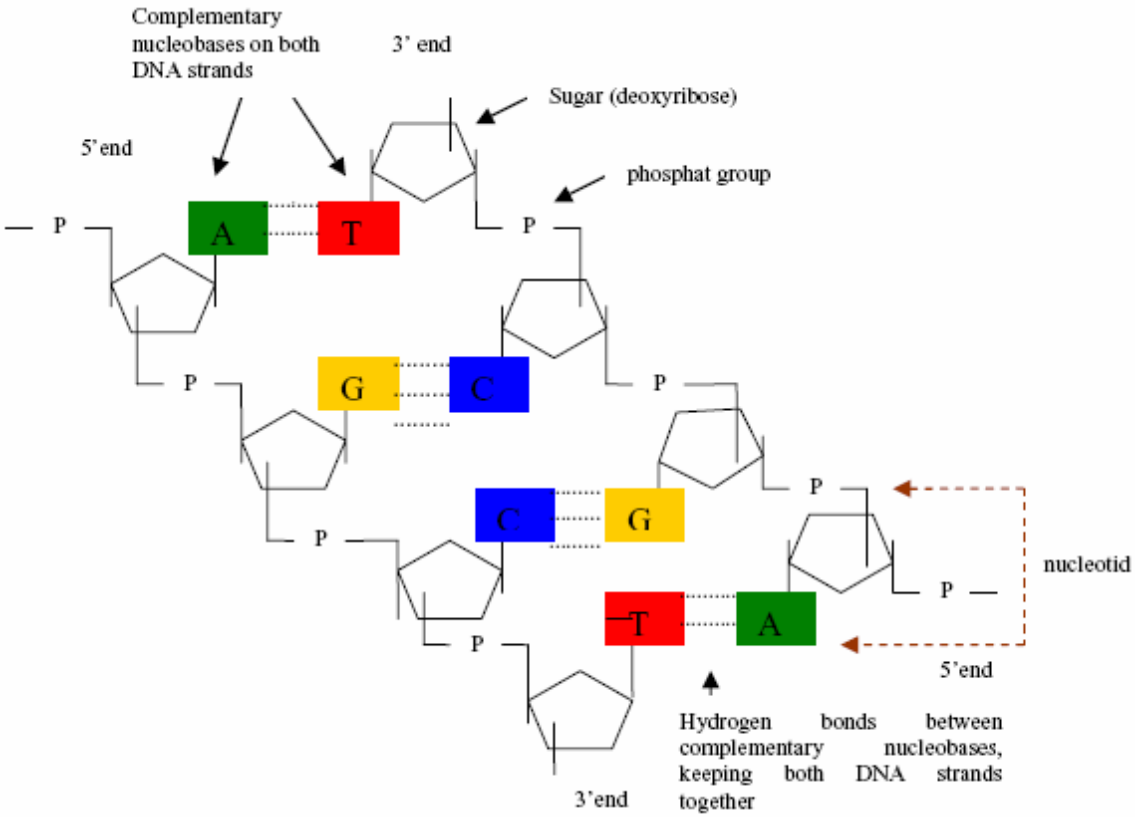

Fig. 1. Schematic representation of the DNA molecule (here only a 4 complementary base pair DNA sequence). A= Adenine; G= Guanine; C= Cytosine; T= Thymine; P= phosphate group. To make up the double helix structure, the two DNA strands are antiparallelly oriented, one of them from the 3'-end to the 5'-end, and the other one from the 5'-end to the 3'-end.

The genetic information encoded in the DNA is copied in the RNA molecules during the transcription process. The mRNA[3] molecules transport the genetic message to the cytoplasm, where it is translated into protein molecules during the translation process[4]. Each individual has its own genetic

[2] The complementary nucleobases are A with T and C with G, and in this manner, they are normally pairing.

[3] Messenger RNA.

[4] Translation is the protein synthesis process performed by ribosomes ("the protein synthesis machinery" of the cell), in which the genetic message carried by messenger RNA is "read" and "translated" into an aminoacid sequence of a protein.

information that determines the structure and functionality of its body and which is further transmitted to offsprings.

The cellular genome, which is localized in the nucleus (99%) and mitochondria (1%), contains all DNA molecules of an organism. The complete human genome has 3.25×10^9 bp (3250 Mbp[5]) in length[6].

The nuclear genome is organized into chromosomes, built from DNA and protecting proteins. Each human sexual cell possesses a **haploid** nuclear genome containing a complete set of genes of 23 chromosomes. Each human body cell has a **diploid** nuclear genome, consisting of 2 sets of chromosomes inherited from the two parents, $2\times23 = 46$ chromosomes in total (that is 23 pairs of homologous chromosomes[7]: 22 pairs of autosomal chromosomes and 1 pair of sex determining chromosomes). The sex chromosomes are XX in females and XY in males, the Y chromosome being essential for masculinity development[8].

During meiosis[9], a genetic recombination process with reciprocal genetic material changing occurs between homologous chromosomes, resulting recombined haploid gametes. These haploid sexual cells (each of them containing a random haploid set combination of chromosomes) will fuse at the time of conception, resulting in the diploid zygote (egg cell), which will develop a new organism. Thus, each individual contains a particular, unique diploid set combination of chromosomes, except the monozygotic twins[10], who are genetically identical[11].

The DNA material shows only 3% coding regions with genes and 97% uncoding regions. Each individual possesses two variants call **alleles** of each gene (one inherited from the mother and one from the father) at a particular genetic locus on homologous chromosomes. The two alleles can be identical in the case of **homozygote** individuals or different in the case of **heterozygote** individuals. This characterization of the alleles present at a genetic locus is called **genotype**[12].

Genes in the nucleus consist of **exons**[13] and **introns**[14]. Different factors (internal or external factors) can produce a modification in the DNA sequence, named **mutation**. There are several types of mutation, point mutations[15], for instance, occur most frequently. The evolution forces keep only the **favorable mutations** in gene exons, which increase fitness of organisms, the **deleterious mutations** being excluded. Mutations occurring in the uncoding part of DNA are accumulated during evolution, as long as they do not affect the survival capacity of the individuals. Thus, a great diversity of DNA sequences named **DNA polymorphisms** exist in the uncoding DNA and intron regions in a population with many individuals.

There are several types of DNA polymorphisms in the uncoding regions; in our study, **STR**[16] **markers** are relevant. These DNA markers are also named **microsatellites** and consist of basic 2-5 bp[17] sequence repeated several times. Many of those DNA polymorphisms have several alleles[18] in a population and a high heterozygosity rate, being valuable tools in forensic and population studies.

---

[5] Mega base pairs; 1 Mbp = 1,000,000 bp.

[6] Hummel, *op. cit*., p. 19-20.

[7] In a pair of homologous chromosomes, one chromosome is inherited from the mother and the other one from the father. As concerns the homologous sex chromosomes, the Y chromosome is always inherited from the father, while the mother can only transmit the X chromosome to her offsprings.

[8] T. Strachan and A. P. Read, *Human Molecular genetics* [3], London and New York, 2003, p. 86-93

[9] Meiosis is the process of cell division forming gametes.

[10] Monozygotic twins develop from the same egg cell.

[11] Strachan and Read, *op. cit.,* p. 86-89.

[12] J. M. Butler, *Forensic DNA Typing- Biology & Technology behind STR Markers*, New York, 2001, p. 13-24.

[13] Exons are protein-codifying parts of a gene.

[14] Introns are DNA uncoding sequences between the exons of a gene.

[15] Point mutation = a nucleobase from a DNA sequence is replaced by other nucleobase, changing the genetic information of that DNA sequence.

[16] STR= short tandem repeats.

[17] bp= base pair.

[18] Each allele of a STR marker has a different number of repeats of the basic sequence, for example the DYS392 STR marker has a four-base sequence repeat, [AGAT], and 11 alleles identified in the human population so far; the allele number 11 contains 11 repeats of the base sequence, the allele 12 contains 12 repeats of the base sequence and so on.

The mtDNA[19] **genome** is frequently used in population genetic research because of some particular traits, namely: 1) its presence in multiple copies per cell (1,000–10,000); 2) it is only maternally inherited; 3) it does not undergo recombination and 4) its mutation rate is about 10 times faster than the average nuclear genes[20].

All members of one maternal lineage show the same mtDNA – representing a **haplotype** and demonstrate the same features codified by the mt genome – this is a **haplogroup**. If the paternal mtDNA enters into the egg cell, it is specifically excluded in the early few hours after conception[21].

All mt genes have no introns and the uncoding DNA spacers are few and short.

The longest uncoding mtDNA sequence is the D-loop or the control region (CR) which shows the highest sequence polymorphism of the whole mt genome by far. This region is a 900 bp uncoding mt region located nearby the nomenclatorial origin of the mt genome and contains the two hypervariable regions: HVRI (341 bp) and HVRII (267 bp)[22] (Fig. 2).

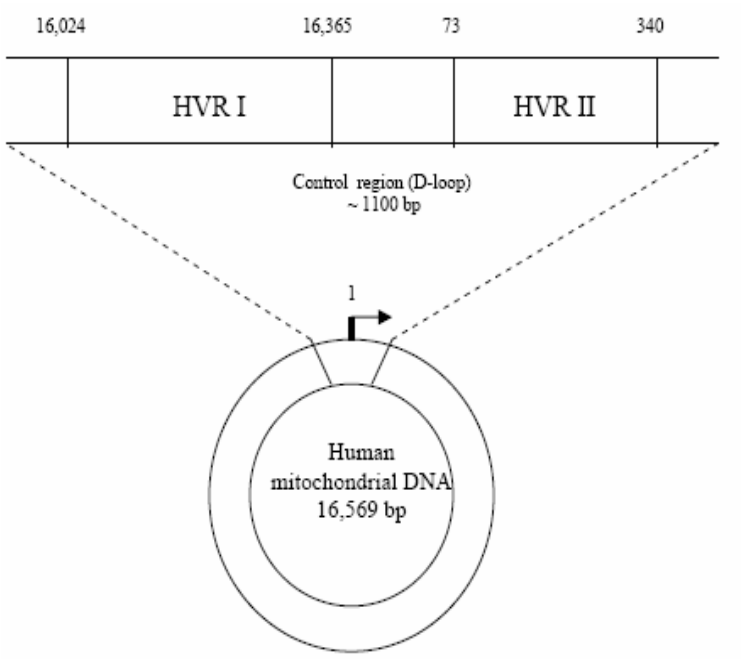

Fig. 2. Human mitochondrial genome – schematic representation (circular, covalent closed, double stranded DNA molecule). The control region (or D-loop) contains the HVR HVR II and I (341 bp, between 16,024- 16,365 bp) (267 bp, between 73- 340 bp) regions. The base numbering system begins arbitrarily inside the control regions, at base pair no. 1, as Anderson *et al.* proposed in 1981 for the European standard reference sequence.

The two HVR[23] regions reveal about 3% variability between two unrelated individuals and their high polymorphic sites are clustered into some hotspots[24].

[19] Mitochondrial DNA.

[20] A. Rokas, E. Laudoukakis, E. Zouros, *Animal mitochondrial DNA recombination revisited.* Trends in Ecology and Evolution 18/8**,** 2003, p.411-417; K.Tamura, M. Nei, *Estimation of the number of nucleotide substitution in the control region of mitochondrial DNA in humans and chimpanzees,* Mol Biol Evol 10**,** 1993, p. 512-526; M. Richards *et al*., *Tracing European Founder Lineages in the Near Eastern mtDNA Pool*, Am.J.Hum.Genet. 67, 2000, p. 1251-1276.

[21] Rokas, Laudoukakis, Zouros, *op. cit.,* p. 512-524.

[22] Hummel, *op. cit*., p. 23-24, S. Anderson, A. Bankier, B. Barell, M. De Bruijn, A. Coulson, J. Drouin, I. Eperon, D. Nierlich, B. Roe, F. Sanger, P. Schreier, A. Smith, R. Staden, I Young, *Sequence and organization of the human mitochondrialgenome*, Nature 290, 1981, p. 457-465.

[23] HVR= hipervariable region.

[24] M. Thomas, P. Gilbert, E. Willerslev, A. J. Hansen, I. Barnes, L. Rudbeck, N. Lynnerup and A. Cooper, *Distribution Patterns of Postmortem Damage in Human Mitochondrial DNA*, Am.J.Hum.Genet. 72, 2003, p. 32-47; Hummel, *op. cit.*, p. 23-24.

Due to the high levels of polymorphism, the HVRs are in general suitable for aDNA analysis without additional use of other mitochondrial or chromosomal markers; the analysis of those mtDNA regions enables investigators to identify individuals to a family lineage level[25] down to the maternal lineage. In this way several studies traced back the maternal inheritance of mtDNA, concluding that the first female ancestor of all modern humans lived in Africa about 150,000 years ago (The "African Eve" hypothesis)[26].

To study genetic information (DNA) inheritance down to the paternal lineage the Y chromosome is used as it is strictly paternally inherited[27].

Peculiarities in paleogentics

In any aDNA analysis, the most crucial step is DNA extraction, because of the status of DNA preservation in fossils. Immediately after death, the degradation of DNA molecules is induced by some endogenous nucleases and continued by environmental conditions[28]. Moreover, some environmental factors such as high temperature and acidic pH act as diagenetic factors and speed up the DNA degradation process. Several processes (such as oxidation, hydrolization, and irradiation) occur and destroy the DNA molecules over time. For instance, oxidative processes modify a number of cytosine and thymine rests to hydantoins, which inhibit DNA polymerases and PCR amplification of the DNA molecules. After a while, the accumulative effects of these degradative processes can lead to a total degradation of all the DNA molecules from fossils. It has been assumed that in physiological concentrations of salt, neutral pH and 15 ^{0}C, all DNA molecules from fossils are degraded in approximately 100,000 years[29].

Our study is part of an extended paleogenetic research focused on mtDNA and ncDNA[30] polymorphisms in old human populations dating from the Bronze and Iron Age, from Romania. Here, we present a molecular genetic analysis on some skeletal remains of 6 old individuals found at Costişa, dating from the Bronze Age.

MATERIAL AND METHODS

The biological material was represented by human fossil bones and teeth of 6 individuals from the Bronze Age, found at Costişa.

The human remains were preserved in graves at about 60 cm depth and archaeologists between 2001 and 2004 dug them up. The fossils were preserved at room temperature after exhumation.

To prevent any contamination, sterile protective equipment, instruments and reagents were used. Solutions were commercially acquired when possible; otherwise, they were autoclaved and UV-treated.

All extractions were performed in a dedicated laboratory of the Institute of Human Biology-University of Hamburg, with three different rooms for ancient DNA (aDNA) extraction (constantly irradiated with UV lamps), PCR amplification and cleaning-up the PCR products.

Prior to extraction, the surface of the teeth and bones were thoroughly washed with 5M HCl, rinsed with sterile distilled water and dried under a UV-lamp for 20-30 min on each side to remove any previous contamination. Then, each human fossil remain was crushed and ground to a fine powder in a metal box with a heavy metal piece, both sterilized at 200^{0}C for at least 2 hours. The bones and teeth powder was preserved at -20^{0}C.

[25] Hummel, *op. cit.*, p.19-25.

[26] A. von Haesler, A. Sajantila & S. Paabo, *The genetical archaeology of the human genome*, Nat. Genet. 14, 1996, p. 135-140.

[27] M.W. Gray, G. Burger, B. F. Lang, *The origin and early evolution of mitochondria,* Genome Biology 2/6, 2001, p. 1018.1- 1081.5

[28] Thomas *et al.*, *op. cit.* (n. 24), p. 32-47.

[29] Paabo, *op. cit.,* p. 1939-1943; Hummel, *op. cit.,* p. 66-78.

[30] Nuclear DNA.

The DNA extraction

For aDNA extraction we used the three following methods:
▪ the phenol-based DNA extraction method as described by Hummel[31] and modified[32];
▪ the guanidine tiocianat and silica-based method described by Hoss and Pääbo[33] and modified[34];
▪ the Invisorb Forensic Kit (Invitek)-based DNA extraction method according with the manufacturer recommendations and modified[35].

To monitor any contamination during each extraction, an extraction blank, without biological sample, was processed together with each fossil piece.

The PCR amplification

PCR[36] is an *in vitro* amplification reaction, performed by a thermostable enzyme DNA-polymerase, the Taq-DNA-Polymerase[37] being usually used. The invention of the PCR in the 1980s opened real perspectives in aDNA study, allowing detecting and amplifying even one DNA molecule persisted in a fossilised biological material.

A typical PCR consists in 30-40 cycles based on 3 stages, performed on 3 temperature steps, namely the DNA template denaturation at 94-95 ^{0}C, the primers annealing at variable temperature, usually between 55 ^{0}C- 70 ^{0}C, and the elongation step at 72^{0}C. Theoretically, at the end of each cycle, the DNA amount is doubled. Thus the PCR allows an exponential amplification of the target DNA sequence flanked by the two specific oligonucleotid[38] primers used, in accordance with the equation: $Y= A (1+F)^n$, where Y= product yield [ng DNA], A= DNA target concentration [ng DNA], F= reaction efficiency and n= number of amplification cycles[39] (Fig. 3).

Practically, the reaction efficiency F is below 1 (or below 100%), even with pure and intact DNA template. In the aDNA field the PCR efficiency is lower (F between 0.4-0.7) because of organic and inorganic contaminants from aDNA extracts and DNA degradation, the PCR product being often undetectable. To compensate the low efficiency of the PCR reaction, in aDNA studies the number of PCR amplification cycles is usually raised over 40 cycles, depending on the number of DNA template molecules and the purity of the DNA extract[40].

In our study, the mitochondrial HVR I and HVR II DNA regions were amplified by PCR reactions, each of them being amplified in two fragments, with different sets of specific primers[41].

Four nuclear DNA markers were amplified by PCR, namely a short DNA sequence from the Amelogenin gene on the sex chromosomes, HumVWA31 marker[42] on the chromosome 12, DYS393 and DYS392 markers on the chromosome Y.

---

[31] Hummel, *op. cit.*, p. 227-229.

[32] G. Cardoş, V. Stoian, N. Mirițoiu, A. Comşa, A. Kroll, A. Rodewald, *Paleo-mtDNA analysis and population genetic aspects of old Thracian populations from South-East of Romania*, Rom. Journal of Legal Medicine,. XII/ 4, 2004, p. 239-246.

[33] M. Hoss and S. Paabo, *DNA extraction from Pleistocene bones by silica-based method*, Nucleic Acids Res. 2, 1993, p. 3913-3914.

[34] Cardoş *et al.*, *op. cit.*, p. 242.

[35] *Ibidem*.

[36] Polymerase Chain Reaction.

[37] The Taq-DNA-Polymerase (or simply Taq-Pol or Taq) is a thermostable polymerase, extracted from *Thermus aquaticus*, a bacterium that lives in hydrothermal vents and hot springs.

[38] Oligonucletid primer= a short sequence of single strand DNA, usually of 18-22 nucleotides, synthetically constructed, complementary with one end of the DNA target fragment amplified by PCR. In the annealing step, each oligonucleotid primer (or simply primer) will recognize and couple with its complementary sequence from the DNA template molecule.

[39] Hummel, *op. cit*., p. 85-86.

[40] *Ibidem*, p. 81-85.

[41] Cardoş *et al.*, *op. cit.*, p. 241; primer sequences available on request.

[42] von Willebrand Factor.

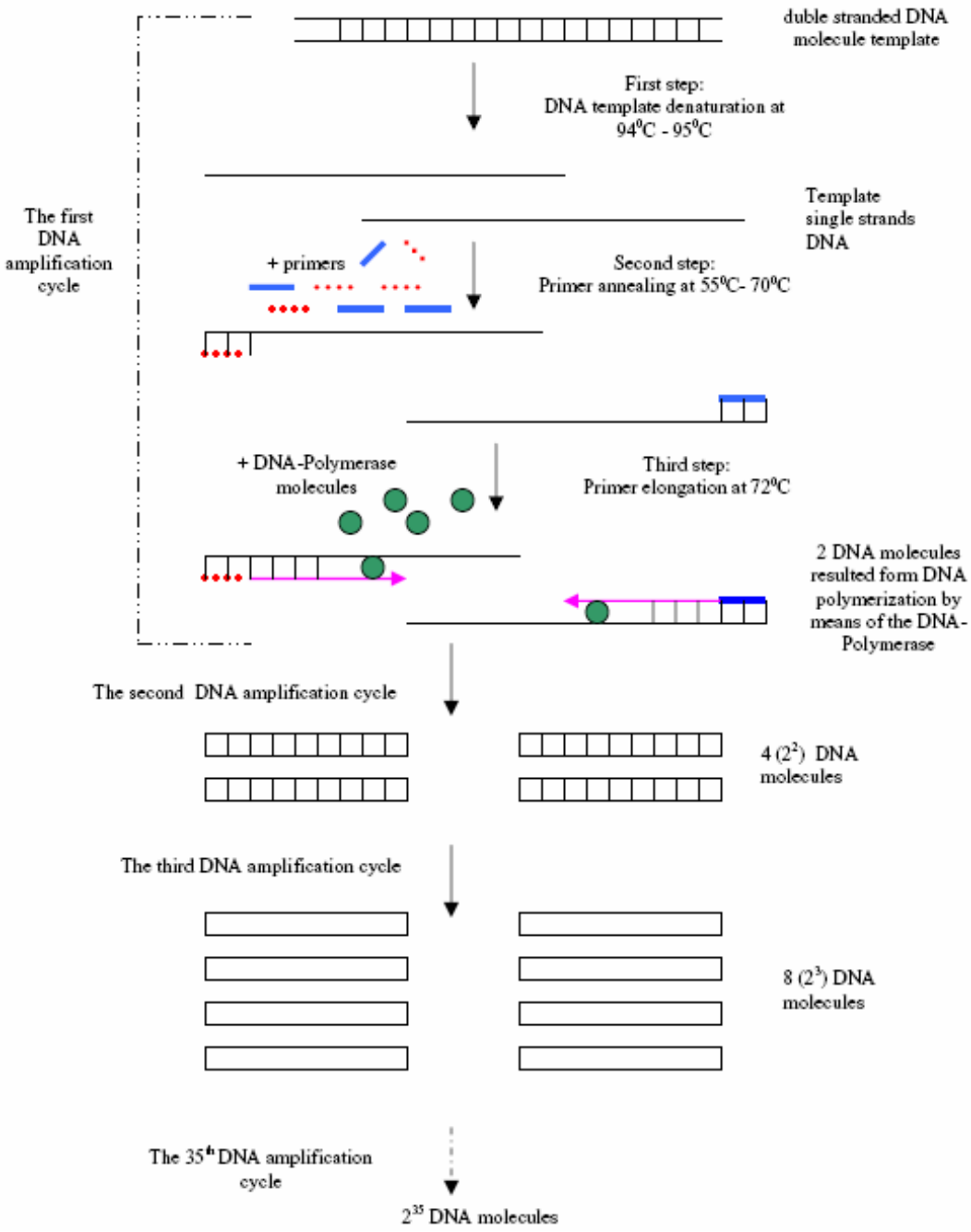

Fig. 3. The Polymerase chain reaction (PCR) – schematic representation (with a theoretically 100% reaction efficiency here).

The Amelogenin gene is localised on the sex chromosomes X (in the Xp22.31-p22.1 region) and Y (in the Yp11.2), and it codes a protein of the dental enamel (Fig. 4).

This gene is used in forensic genetics for sex-typing assay, because the first intron of the gene shows a 6 bp deletion[43] on the X chromosome. We amplified this region with a specific primer pair[44], and we obtained a 106 bp DNA sequence from the X chromosome and a 112 bp DNA sequence from the Y chromosome. In the case of male individuals, two PCR products were obtained (corresponding to the both sequences mentioned above) and in the case of females only the one from the X chromosome.

[43] 6 bp missing
[44] Hummel, *op. cit.*, p. 235.

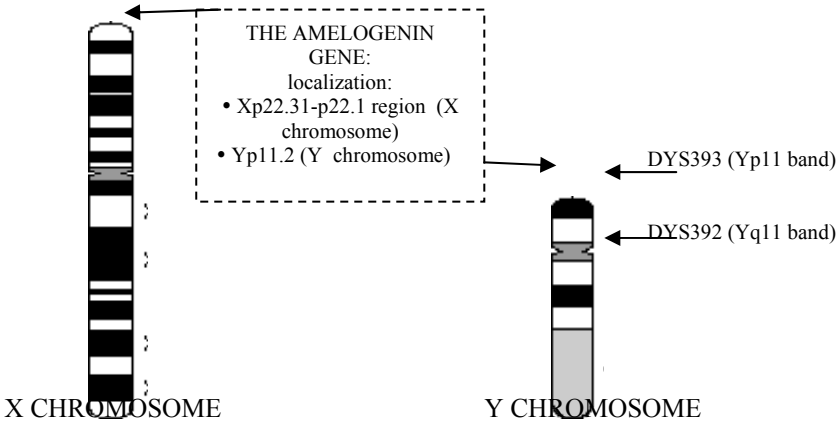

Fig. 4. Physical map of sex chromosomes X and Y, particularly the patterns of dark and light bands obtained after Giemsa staining. Generally, the genes are localized in the light bands. Localization of the Amelogenin gene on the X and Y chromosomes and of the DYS392 and DSY393 STR markers on the Y chromosome.

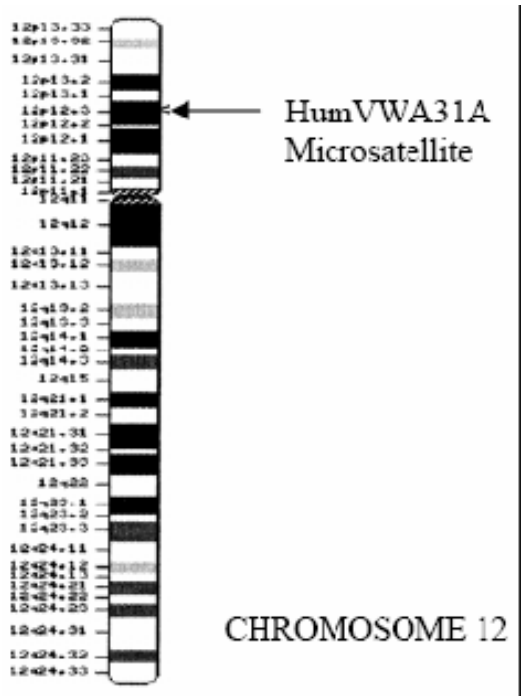

Fig. 5. Physical map of chromosome 12, after Giemsa staining; the arrow shows the HumVWA31A Microsatellite localization.

The DYS393 STR marker is a trinucleotid microsatellite[45], localized in Yp11 band of the Y chromosome (Fig. 4).

The DYS392 STR marker is a tetranucleotid microsatellite[46], localized in Yq11 band of the Y chromosome (Fig. 4). The primer sequences used in our study for PCR amplifying these two DNA STR markers are available on the STRBase website[47].

The HumVWA31A microsatellite is a tetranucleotid STR marker, localised in the 40th intron of the von Willebrand Factor gene, on the 12th chromosome, in the 12p12 region (Fig. 5). The primer set used in

[45] The base sequence repeat of the DYS393 STR marker is [TAT] $_{9\text{-}17}$.
[46] The base sequence repeat of the DYS392 STR marker is [AGAT] $_{6\text{-}17}$.
[47] http://www.cstl.nist.gov/div831/strbase/

our study amplified a 134-170 bp-DNA sequence (depending on the amplified allele). The primer sequences were:
HumVWA31A P1 forward: 5'-CCC TAG TGG ATG ATA AGA ATA ATC-3';
HumVWA31A P2 reverse: 5'-GGACAGATGATAAATACATAGGATGGATGG-3'.

The test for inhibitors consisted of amplifying the control DNA (DNA K_{562}) in the presence of variable amounts of aDNA extracts.

All the PCR reactions were carried out at least in duplicate, in a *Perkin Elmer GeneAmp PCR System 9600 Thermal Cycler* and an *Eppendorf Mastercycler Gradient Cycler*, the reaction mixture and the PCR profile were presented elsewhere[48]. In order to monitor contamination events, one to three negative controls, with DNA-free water instead of DNA extract, were used for each PCR.

The presence of PCR products was demonstrated by UV visualisation by electrophoresis on agarose gels 2% and photographed with an *Olympus C-4000* digital camera.

The alleles of the Amelogenin gene fragment and of the DNA microsatellite markers were separated by electrophoresis on non-denaturating 8% polyacrylamide gels, and they were visualised after silver staining[49]. The DNA alleles identified against male and female DNA positive control (in the case of the Amelogenine gene) and against a HumVWA31A DNA allelic ladder[50].

The DNA sequencing, alignment and analysis

The mtDNA PCR products were sequenced at MWG (Germany) by the Sanger method[51], after cleaning them with the *MBS spin PCRapace Kit (50)* (Invitek). Every aDNA sequence was amplified and sequenced at least twice and was compared with the corresponding sequences of DNA handlers and the control DNA, in order to detect any contamination.

The mtDNA sequences were compared with the modern European Reference Sequence CRS[52] and similar sequences of some other old individuals dating from the Bronze and Iron Age from Romania (our unpublished data), using the Bioedit Program[53]. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 3.1[54].

RESULTS AND DISCUSSION

Here we studied at DNA level the skeletal remains of 6 individuals found in graves from the Costişa necropolis, dating from the Bronze Age. We extracted DNA from old bones and teeth by the three methods described in literature and modified by adapting them to the preservation status of fossils. Most of the fossil teeth were better preserved than human old bones, probably due to the enamel layer, which protects teeth against damage induced by environmental factors.

We obtained amplifiable mitochondrial (Fig. 6) and nuclear DNA from the teeth of 3 out of 6 individuals, namely:

---

[48] Cardoş *et al.*, *op. cit.*, p. 242.

[49] Butler, *op. cit.*, p. 135-163.

[50] The HumVWA31A DNA allelic leader[50] was made in our laboratory by amplifying DNA samples of some human individuals with known alleles, from a previous study made here by Susanne Mueller-Scholtz.

[51] F. Sanger, S. Nicklen, A.R. Coulson, *DNA sequencing with chain terminating inhibitors*, Proc.Natl.Acad.Sci. 74, 1977, p. 5463-5467.

[52] Cambridge Reference Sequence; Anderson *et al.*, *Sequence and organization of the human mitochondrial genome*, Nature 290, 1981, p. 457-465.

[53] (Version 5.0.9.; Tom Hall, Department of Microbiology, North Carolina State University). T.A. Hall, *BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT*, Nucl. Acids. Symp. Ser. 41, 1999, p. 95-98.

[54] S Kumar, K Tamura, and M Nei, *MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment*, Briefings in Bioinformatics 5, 2004, p. 150-163.

1. M2, S III, 2004, noted here the individual 14A;
2. M3, S V, 2004, noted here the individual 15A;
3. M1, square E IV, 2001, noted below the individual 16A.

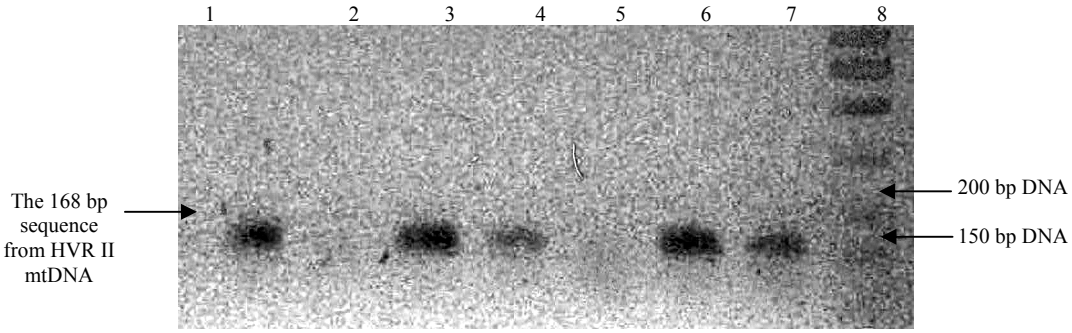

Fig. 6. The 168 bp fragment from HVRII mtDNA on 2% agarose gel. Lines 1 and 3: the individual 14A; Lines 2 and 5: the PCR negative controls; line 4: the individual 1 A; line 6: the positive control DNA; line 7: the individual 16A; line 8: the 100 bp DNA weight marker.

The skeletal remains of the other 3 individuals were in an advanced degradation status, containing no amplifiable DNA sequences by PCR. Moreover, in some DNA extracts, Taq-DNA Polymerase inhibitors were detected by the inhibitor test. Any trial to remove those inhibitors by some DNA cleaning methods failed.

As far as mtDNA is concerned, we obtained a sequence from HVR I for one individual and from HVR II for 3 individuals. The HVR II polymorphisms are presented here, the DNA sequence alignment was done in Bioedit Program[55] (Fig. 7). The only one short HVRI sequence obtained has no scientific relevance for a sequence analysis here, so far. The mtDNA sequences of the three old individuals from Costişa analyzed here were compared with the mtDNA CRS and with some similar sequences of other old individuals from the Bronze and Iron Age from Romania (not shown here).

```
                 110        120        130        140        150
                 ...|....|....|....|....|....|....|....|....|....|.
CRS              gagcaccctatgtcgcagtatctgtctttgattcctgcctcatcctatta
14A-Costisa      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
15A-Costisa      ..............................................C.....C....
16A-Costisa      ..............................................T......

                 160        170        180        190        200
                 ...|....|....|....|....|....|....|....|....|....|.
CRS              tttatcgcacctacgttcaatattacaggcgaacatacttactaaagtgt
14A-Costisa      ~~~~~~~~~~~~~~~~~~.....................C..........
15A-Costisa      ..............................................C..........
16A-Costisa      ..............................................

                 210        220        230        240
                 ...|....|....|....|....|....|....|....|...
CRS              gttaattaattaatgcttgtaggacataataataacaattga
14A-Costisa      ..........................................
15A-Costisa      ................
16A-Costisa      ................
```

Fig. 7. MtDNA of 3 old individuals from Costişa, dating from the Bronze Age, together with the European reference sequence, CRS. Comparing the aDNA sequences with the CRS sequence, the individual 14A shows a T→C (bp 195) substitution, the individual 15A shows three T→C (bp 146, 152 and 195) substitutions, and individual 16A shows a C→T (bp 150) substitution. A= Adenine; G= Guanine; C= Cytosine; T= Thymine. Colored points mean base identity.

[55] *Ibidem*, p. 150-163.

The mtDNA sequences of the three old individuals have 1 to 3 point mutations against the CRS sequence (Fig. 7): a T→C substitution at the 195 bp in the sequence of the individual 14A, the DNA sequence of the individual 15A has three T→C substitutions at the 146, 152 and 195 bp, and the DNA sequence of the individual 16A has a C→T substitution at 150 bp. For the individual no. 14A we obtained and sequenced only the second part of the HVRII region, in this case we can not say for certain whether there are other point mutations in the missing DNA sequence of that individual or not.

The point mutations found in the mtDNA sequences presented above were also found in various patterns, in corresponding mtDNA sequences of other old individuals from different archaeological sites from Romania, dating from the Bronze Age (Sultana-Malu Roşu) and Iron Age (Satu Nou-Valea lui Voicu, Jurilovca, Babadag and Enisala-Palanca). That showed the closed genetic kinship along the maternal lineage between the three old individuals from Costişa and some individuals found in other archeological sites dated from the Bronze and Iron Age. We also should note that the point mutations analyzed above are also found in Romanian modern population (our unpublished data), suggesting that some old individuals from the human populations living on the Romanian land in the Bronze and Iron Age, could participate to a certain extent in the foundation of the Romanian genetic pool.

As concerns ncDNA, we obtained the Amelogenin sequence of the same 3 individuals. Thus, the genetic analysis revealed that the three old individuals investigated here were females, only the 106 bp-DNA sequence were present in each amplified DNA sample (Figs. 8 and 9).

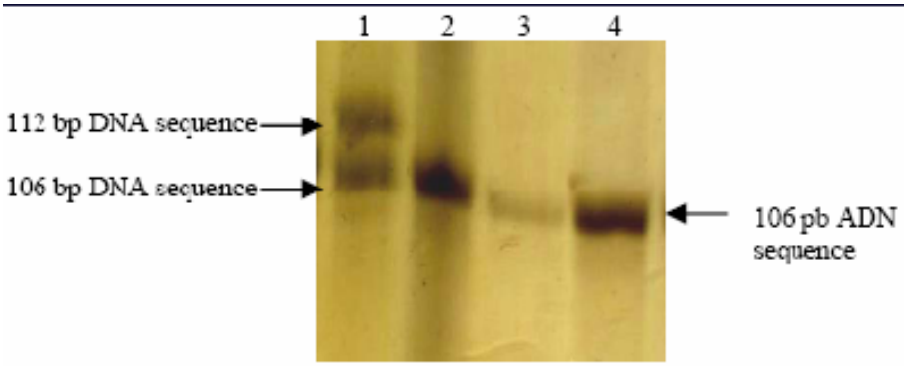

Fig. 8. The Amelogenin gene sequence on non-denaturating 8% polyacrylamide gel. Line 1: the male control DNA; line 2: the female control DNA; line 3: the individual 16A female; line 4: the individual 15A female.

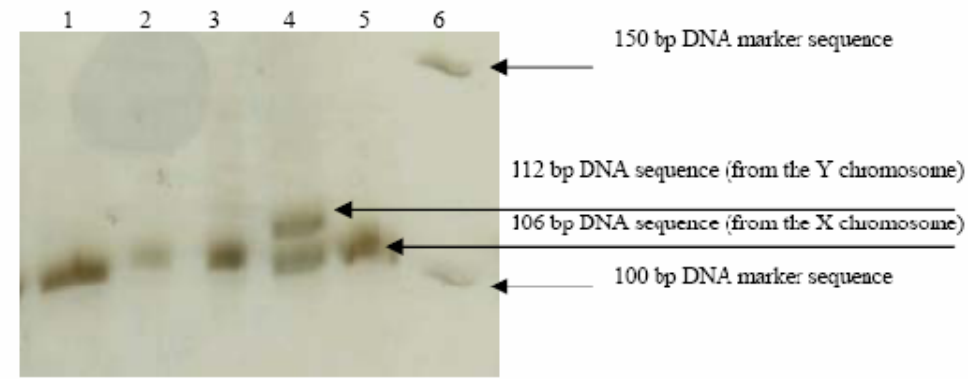

Fig. 9. The Amelogenin gene sequence on non-denaturating 8% polyacrylamide gel. Line 1: individual 14A female, line 2: the individual 15A female, line 3: the individual 16A female; line 4: the control male DNA; line 5: the control female DNA; line 6: the 100bp DNA marker.

The PCR amplification of the DNA markers localized on the Y chromosome led to negative results, proving once again the female genetic sex of those old individuals, already established by means of amplifying the Amelogenin gene.

The alleles for the vWA31 marker could be established for the same 3 individuals, as follows:
– the allele 15 for the individual 14A, the genotype being homozygote 15/15;
– the allele 18 for the individual 15A, the genotype being homozygote 18/18 (Fig. 10);
– the alleles 18 and 20 for the individual 16A, thus, the genotype being heterozygote 18/ 20.

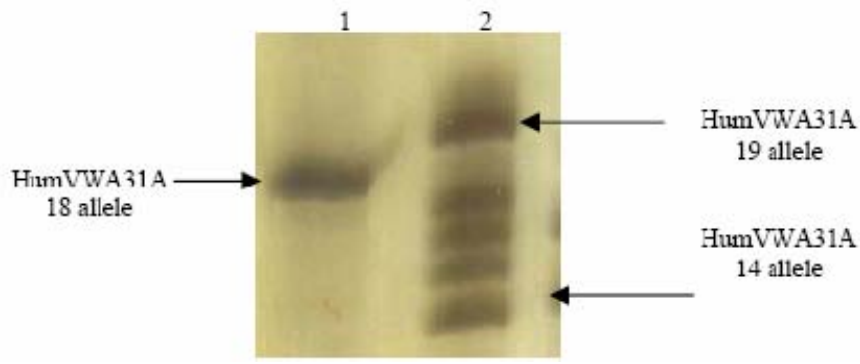

Fig. 10. The HumVWA31A Microsatellite. Line 1: the individual 15A- the 18 allele; line 2: the HumVWA31A DNA allelic ladder.

The alleles of the vWA31A marker identified in the three individuals from Costişa, dating from the Bronze Age, were also found in some old individuals from Sultana Malu Roşu (dating from the Bronze Age), Jurilovca and Niculiţel-Cornet (dating from the Iron Age) (our unpublished data). The allele 18 is also found very frequently in Romanian modern population (in European modern population as well) together with the allele 17. In the old individuals from Costişa the allele 17 was not found, either because of the small individual sample, or because of a genetic drift or a founder effect.

The genetic kinship of the 3 old individuals from Costişa with other old individuals was analyzed starting from of HVRII mtDNA sequences analyzed above. For this reason, the phylogenetic trees were constructed by means of the Neighbor-Joining Tree Method with bootstrap test (1,000 replicates), based on Kimura 2 parameter and Tamura si Nei formulas (Figs. 11 and 12).

The close genetic kinship on the maternal lineage can be noticed between the individuals 14A and 15A, they are clustered together in both phylogenetic trees, away from the individual 16A, which had a different haplotype, that means a different maternal lineage. The sequence of individual 16A clustered in the same cluster with the CRS sequence and the individual 23F[56] from the Iron Age found in Jurilovca (Figs. 11 and 12).

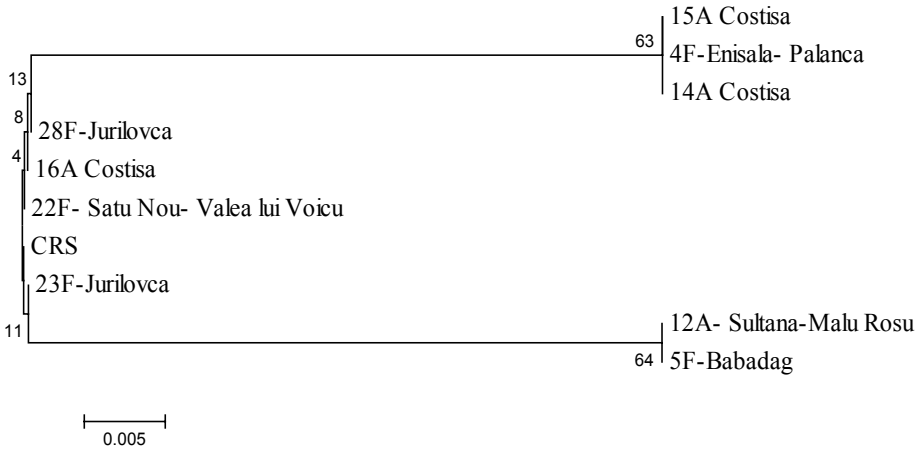

Fig. 11. The phylogenetic tree constructed by means of Neighbor-Joining tree method, based on Kimura 2 parameter formula (bootstrap 1000 replicates).

[56] 23F, 28f, 4F, 5F are indicatives of individuals from the Iron Age, found in archaeological sites named there. 12A is indicative of an old individual from Sultana Malu-Roşu, dating from the Bronze Age.

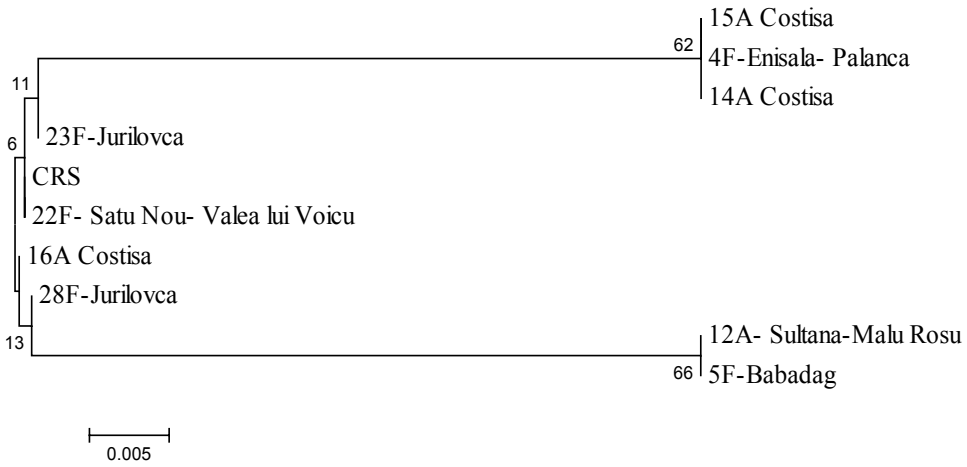

Fig. 12. The phylogenetic tree constructed by means of Neighbor-Joining tree method, based on Tamura and Nei formula (bootstrap 1000 replicates).

As concerns the vWA31A nuclear marker, the individuals 15A and 16A share the allele 18, which can be most likely inherited on the paternal lineage.

CONCLUSIONS

In the present study, 3 out of 6 old individuals dating from the Bronze Age found at Costişa were analyzed at the genetic level, their skeletal remains containing some amplifiable DNA.

The molecular analysis of the nuclear Amelogenin gene showed that all of those 3 individuals are females at genetic level, in agreement with the anthropological study.

Based on the mtDNA and ncDNA data obtained in this study, we can conclude that the individuals 14A and 15A are closely related to each other on the maternal lineage, the individual 16A representing a different maternal lineage, thus, a different mitochondrial haplotype, and the individuals 15A and 16A, sharing the allele 18 of vWA31A, are likely related on the paternal lineage.