WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Mitochondrial DNA Profiling in Combination with Autosomal Loci as a Valuable Tool for Obtaining Statistical Significance in Forensic Cases

Renata Jankova [a]*, Zlatko Jakovski [a], Silvana Filipova [b]

[a] *Institute of Forensic medicine, criminalistic and medical deontology, Faculty of Medicine, Majka Tereza 19, Ss."Cyril and Methodius", University in Skopje, North Macedonia*

[b] *Center for rehabilitation of verbal communication pathology, Belgradska 15, 1000 Skopje, Republic of North Macedonia*

A B S T R A C T

The majority of forensic casework utilizes DNA markers of the nuclear genome such as short tandem repeats (STRs) scattered throughout the non-coding regions of autosomal or sex chromosomes. However, there are some scenarios in which the analysis of nuclear DNA markers is impossible or impractical, or not sufficient for statistical calculation of the results. In number of such cases, mitochondrial DNA (mtDNA) analysis has proven to be a valuable tool. Several characteristics such as high level of polymorphisms, multiple copies of the mtDNA genome in each cell and maternally inheritance make it useful in forensic genetic analysis. Mitochondrial DNA profiling is mainly used in human identification and in the analysis of crime scene samples when conventional DNA analysis is difficult. In identifying human remains sometimes it may be possible to obtain large amount of nuclear DNA suitable for profiling, but there may be no suitable relative with whom to compare, so statistical calculation for the results of analyzed nuclear markers (STRs) may be not sufficient for statistical significance.

This article presents some forensic cases where statistical significance of the results, under requirements of the legal system, had been assessed after analysis of additional genetic markers such as mtDNA.

Keywords: mtDNA analysis, statistical calculation, forensic cases.

1. Introduction

Three main characteristics of the mtDNA genome make it a valuable forensic marker (Conell et al., 2024). First, it contains relatively high levels of polymorphisms, which allow individuals to be differentiated. Second, there are big number of copies in each cell increasing the possibility of recovering mtDNA from forensic samples such as bone, teeth, hair or degraded remains which contain low amount of DNA making conventional DNA analysis difficult (Robin et al., 1988). Third, it is maternally inherited without recombination and must be treated as a single locus haplotype. Mother passes along her mtDNA type to her children unchanged and therefore siblings and maternal relatives have an identical mtDNA sequence. This can be helpful in solving missing persons cases or mass disaster identification. Since distantly related maternal relatives should possess the same mtDNA type, this extends the number of useful reference samples that may be used to confirm the identity of a missing person, but will likely reduce the significance of a match in forensic cases (Gill et al., 1994; Stone et al., 2001). Except for these major features of mtDNA genome that make it an important forensic DNA marker, it is also a valuable additional system in forensic analysis when statistical significance of the results after STR analysis is not attained, such as in some cases of identification or kinship analysis when there are not suitable relatives for comparison. It is important to know that mtDNA can never have the power of discrimination that an autosomal STR marker can, since its inheritance is uniparental.

1.1 Mitochondrial DNA

Mitochondrial DNA is a separate genome found in mitochondria, organelles that exist in the cytoplasm of eukaryote cells. Human mtDNA was first sequenced in 1981 in the laboratory of Frederick Sanger in Cambridge, England (Anderson et al., 1981), first named Anderson sequence or Cambridge Reference sequence (CRS) and was after being re-sequenced and known as revised Cambridge reference sequence (rCRS) (Andrews et al., 1999). The human mtDNA genome is a circular genome only 16569 base pairs (bp) long. Analysis of the human mtDNA has revealed very economic use of the DNA with only few non-coding bases in the genome, except in a region called the D-loop region or control region. Since the control region is non-coding (it plays role in replication and transcription of mtDNA), the restrictions are fewer for nucleotide variability and polymorphisms between individuals are more abundant than in similarly sized portions of the coding region. Accumulation of polymorphisms in this region is because the sequence does not code for any substances necessary for the cell's function. As the most polymorphic, the control region of mtDNA particularly two hypervariable regions commonly referred to as hypervariable region I (HV1) and hypervariable region II (HV2), are mostly the focus in forensic DNA studies. Occasionally a

third portion of the control region, known as HV3, is examined to provide more information regarding the tested sample. Control region spans form nucleotide position 16024 to 16365, and from 1 to 576.

1.2 Producing a mtDNA profile

The first step in mtDNA profiling is to extract DNA from a sample. The processes used to extract mtDNA are the same as those used for nuclear DNA extraction. After the extraction of DNA, amplification of mtDNA using polymerase chain reaction (PCR) follows, using specific primers for region of interest. Since the most mtDNA variations between individuals are found within the control region, particularly in hypervariable region I (HV1) and hypervariable region II (HV2), these regions are examined by PCR amplification followed by sequencing analysis. Sequencing analysis is commonly performed using Sanger sequencing chemistry (Sanger et al. 1997; Wilson et al., 1995) and obtained results are reported as variations compared to the rCRS.

1.3 Population databases

There are population databases with thousands of mtDNA profiles in them which are used in estimating the expected frequency of mtDNA haplotypes that are observed in casework when there is a match between compared samples. Information in the database is also important for reliable estimation of the frequency for a random match. A European DNA Profiling Group mitochondrial DNA population database project (EMPOP) has gathered thousands of mtDNA sequences and constructed a high-quality mtDNA database that can be accessed online (Parson et al., 2004; Parson, Dur, 2007). While searching through this database, exact mtDNA profile is entered into a database and its frequency within particular population is examined. The larger the number of unrelated individuals in the database, the better the statistics will be for a random-match frequency estimate. In EMPOP database, mtDNA profiles from Macedonian population are also included (Zimmermann et al., 2007; Jankova-Ajanovska et al.,2014). One of the biggest weaknesses of mtDNA analysis is that some haplotypes are rather common in various population groups.

2. Methodology

2.1 Forensic cases with application of mitochondrial DNA profiling in reporting statistics

Case 1: Upon the request of the prosecution, the process of identification of the death body found in a river, in stage of advanced decomposition, was performed. On the surface of the body the adipocere was formed and remained internal organs were in advanced stage of putrefaction. This combination of postmortem changes is commonly seen in corpses that have been submerged in water for a long period of time. With anthropological studies it was determined that the body is a female in her eighth decade of life. The estimated post mortem interval was at least 6 months. These three pieces of information are important for the investigation authorities to make a connection with families that reported missing person within the given postmortem interval and with similar descriptions as ours. DNA was extracted from the femur of the corpse and autosomal STR DNA profile was obtained. It was compared with DNA profile from the supposed daughter of the deceased and probability of match was calculated. For statistical calculation a presumption of prior probability must be made. Prior probability is defined according to the event or the story about the case. In this case the prior probability was determined to be 50%, because in this case there are two possibilities, one being that the deceased body is the mother of the daughter and the other that it is not, so one or the other is true. Hypothesis testing is the formal procedure for using statistical concepts and measures in performing decision-making.The statistical calculation of DNA result was performed using DNA VIEW statistical program for STRs, under hypothesis that the deceased body is the mother of the daughter or it is not (Charles et al., 1992-2020).

D:M+Man

/

Under this hypothesis the obtained likelihood ratio (LR) was LR=188 and posterior probability 99,4%. A likelihood ratio (LR) involves a comparison of the probabilities of the evidence under two alternative propositions. Probability is the number of times an event happens divided by the number of opportunities for it to happen. In the context of forensic DNA evidence, under requirements of the legal system, posterior probability must be at least 99.9% (Hoppe et al., 2002) for decision making in this hypothesis testing. Since the result of posterior probability was 99,4%, additional markers such as mtDNA were analyzed. The choice was on mtDNA analysis because if the deceased woman is the mother of the supposed daughter, they will have the same mtDNA profile. Table 1 and 2 shows the results from the analysis of hypervariable region HV1 and HV2 of the control region of mtDNA.

Table 1. Results of mtDNA sequence analysis in hypervariable region 1

HV1 (position 16024-16365)		
Anderson sequence	16224	16311
	T	T
NN	C	C
Supposed daughter	C	C

Table 1 shows the polymorphism (two base substitutions) on analyzed samples on positions 16224 and 16311 in comparison to the Anderson sequence. Polymorphism on both positions is the same in NN body and supposed daughter.

Table 2. Results of mtDNA sequence analysis in hypervariable region 2

HV2 (position73-340)				
Anderson	73	94	263	315.1
sequence	A	G	A	-
NN	G	A	G	C
Supposed daughter	G	A	G	C

Table 2 shows the polymorphism (three base substitutions and one insertion on position 315) on analyzed samples on positions 73, 94, 263 and 315 in comparison to the Anderson sequence. Polymorphism on all positions is the same in NN body and supposed daughter.

Figures 1 and 2 shows electropherogram analysis of mtDNA from the femur of the remains, as raw data (figure 2) and as analyzed sequence (figure3).

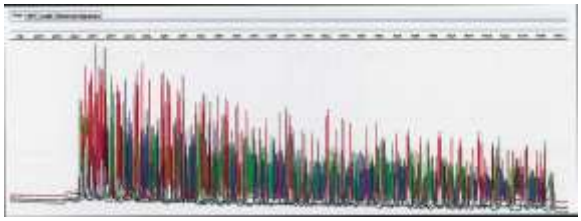

Figure 1. Raw data of mtDNA sequence from the sample of the bone (femur) from NN body

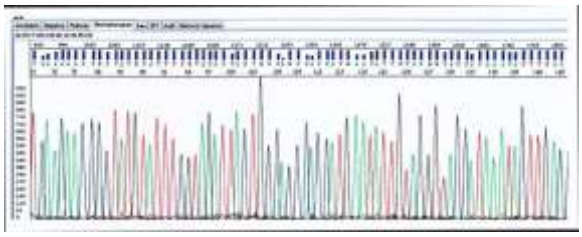

Figure 2. Electropherogram of the mtDNA sequence, extracted from the bone

The results from mtDNA analysis were included in statistical calculation, so posterior probability of analysis of STRs was used as prior probability in cumulative posterior calculation. LR for mtDNA sequence was calculated using EMPOP-European population database for mtDNA.

Cumulative LR was calculated as following:

LRc=LRstr x LRmtdna=188 x 201 = 37788

Cumulative posterior probability:

PPc=LRc x prior/LRc x prior + (1-prior) x 100%

=37788 x 0,994/ 37788 x 0,994 + 0,006 x 100%

=37561,272/37561,278 x 100% = 99,99998%.

Posterior probability of 99,99998% is strong evidence for the tested hypothesis that the deceased body is the mother of the supposed daughter.

Case 2: The second case represents a case of kinship analysis where a woman requested to be tested in order to find out if she was an adopted child in the family, or she was a sibling with a man. For disclosure of this forensic case, an analysis of mtDNA was required, since the analysis of their autosomal STRs after statistical calculation of the results, didn't show a conclusive result. Under hypothesis that the woman and man are siblings or not related (S=sister, B=brother),

S: Ma+Pa

B: Ma+Pa

/

the obtained likelihood ratio was 27,9 and posterior probability was 93%. For reaching the necessary criteria of 99,9%, analysis of additional genetic markers such as mtDNA was included. The results of the analysis of their mitochondrial DNA as shown in table 3 and 4 showed that they have the same polymorphisms on positions that differ from the reference Anderson sequence.

Table 3. Results of mtDNA sequence analysis in hypervariable region 1

HV1 (position 16024-16365)			
Anderson	16069	16126	16256
sequence	C	T	C
sister	T	C	T
brother	T	C	T

Table 3 shows the polymorphism (three base substitutions) of analyzed samples, on positions 16069, 16126 and 16256 in comparison to the Anderson sequence. Polymorphism on all positions is the same in the tested man and woman (supposed siblings).

Table 4. Results of mtDNA sequence analysis in hypervariable region 2

HV2 (position73-340)							
Anderson	73	185	188	228	263	295	315.1
sequence	A	G	A	G	A	C	-
sister	G	A	G	A	G	T	C
brother	G	A	G	A	G	T	C

Table 4 shows the polymorphism (six base substitutions and one insertion on position 315) of analyzed samples on positions 73, 185, 188, 228, 263, 295 and 315 in comparison to the Anderson sequence. Polymorphism on all positions is the same in the tested man and woman (supposed siblings).

Upon statistical calculation of the results from analysis of autosomal STRs and mtDNA, the obtained likelihood ratio and posterior probability were as follows:

LRc=LRstr x LRmtdna= 27,9 x 201 =5608

PPc= LRc x prior/LRc x prior + (1-prior) x 100%

=5608 x 0,5/5608 x 0,5 + 0,5 x 100%

=99,97%.

So, the cumulative posterior probability was 99,97% in favor of the hypothesis that they are a brother and a sister.

3. Discussion

When DNA profiles derived from evidence or human remains and reference sample fails to exclude an individual as a contributor of the evidence or as a relative to the missing person, statistical assessment is used to evaluate the significance of the association. Proper statistical conclusion requires careful formulation of the question to be answered, including, in this instance the requirements of the legal system. Hypothesis testing is the formal procedure for using statistical concepts and measures in performing decision-making. Usually, two hypothesis are formulated for testing. One hypothesis is put forward by the prosecution and the other by the defense. The outcome of the significance testing is very much dependent on how the question is framed as part of the hypothesis testing. In parentage testing and in kinship analysis different questions are usually being asked. While criminal casework involves looking for direct matches between evidence and suspect DNA profiles, relationship testing relies on additional assumption that involves genetic inheritance patterns and the possibility of mutation. In kinship analysis there is always a greater level of interpretation uncertainty with results due to inheritance pattern that vary depending on the relationship being considered and the possibility of an unrelated person sharing common alleles. According to the international criteria, statistical estimate of posterior probability of 99,9% is considered as significant. For statistical calculation in some forensic cases, as it is shown in this article, statistically significant results could be obtained after analysis of additional DNA markers such as mtDNA.

Statistical significance of the results must be reported under the legal system requirements and international standards. The frequency estimates for autosomal and mtDNA typing results obtained for a given sample may be combined. A combined likelihood ratio from the results of the analysis of autosomal and mtDNA for missing person searches or in kinship analysis, as shown in this article, can be generated and presented as valuable evidence to the court.

References

1. Connell R.J, Lea R.A, Haupt L.M, Griffiths L.R. (2024). Mitochondrial DNA analysis in population isolates: Challenges and Implications for Human Identification. Current Molecular Biology Reports.10:1-8. https://doi.org/10.1007/s40610-023-00155-4
2. Robin E.D, Wong R. (1988). Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. Journal of Cellular Physiology. 136:507-513 doi: 10.1002/jcp.1041360316.
3. Gill P, Ivanov P, Kimpton C, Piercy R, Benson N, et al. (1994). Identification of the remains of the Romanov family by DNA analysis. Nature Genetics. 6:130-135 doi: 10.1038/ng0294-130.
4. Stone A.C, Starrs J.E, Stoneking M. (2001). Mitochondrial DNA analysis of the presumptive remains of Jesse James. Journal of Forensic Sciences. 46:173-176 doi:10.1520/JFS14932J
5. Anderson S, Bankier AT, Barrel GB, DeBrujin MHL, Coulson AR, Drouin J et al. (1981). Sequence and organization of the human mitochondrial genome. Nature. 290:457-465doi: 10.1038/290457a0.
6. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nature Genetics. 23:147 doi: 10.1038/13779.
7. Sanger F, Nicklen S, Coulson R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. 74:5463-5467 doi: 10.1073/pnas.74.12.5463.
8. Wilson MR, DiZinno JA, Polanskey D, Replogle J, Budowle B. (1995). Validation of mitochondrial DNA sequencing for forensic casework analysis/ Int. J. Legal Med. 108:68-84. doi: 10.1007/BF01369907.
9. Parson W Brandstatter A, Alonso A, Brandt N, Brinkmann B, Caraccedo A, et al. (2004) The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercise: organization, results and perspective. Forensic Science Int. 139:215-226. doi: 10.1016/j.forsciint.2003.11.008.
10. Parson W, Dur A. (2007). EMPOP-a forensic mtDNA database. Forensic Science International: Genetics. 1:88-92. doi: 10.1016/j.fsigen.2007.01.018.
11. Zimmerman B, Brandstatter A, Duftner N, Niederwieser D, Spiroski M et al. (2007). Mitochondrial DNA control regio population data from Macedonia. Forensic Science International:Genetics. 1:e4-e9 doi:10.1016/j.fsigen.2007.03.002
12. Jankova-Ajanovska R, Zimmerman B, Huber G, Rock Alexander, Bodner M, Jakovski Z et al. (2014). Mitochondrial DNA control region analysis of three ethnic groups of the Republic of Macedonia. Forensic Science International: Genetics. 13:1-2doi: 10.1016/j.fsigen.2014.06.013.
13. Charles Brenner. 1992-2020. DNA-VIEW User's Manuel.
14. Hoppe JD, Kurth R, Sewing KF. (2002). Richtlinien fur die Erstattung von Abstammungsgutachten. Deutsches Arzteblantt, 99(10): A- 665