

HVI MITOCHONDRIAL DNA DATA IN A POPULATION SAMPLE FROM SOUTH-EAST POLAND

Dorota MONIES, Marzanna CIESIELKA, Roman MADRO

Chair and Department of Forensic Medicine, Medical Academy, Lublin, Poland

ABSTRACT: A population database was generated from 106 unrelated individuals living in South-East Poland. Mitochondrial DNA from hypervariable region I was determined by PCR amplification and direct sequencing. Sequence comparison led to the identification of 85 mitochondrial lineages, as defined by 67 variable positions. The most frequently occurring lineage comprised 11 individuals, whilst 57 types were observed in only one individual. A statistical analysis of the results for this population showed a genetic diversity of 0.94. The probability of two random individuals showing identical mtDNA haplotypes is 0.06. The average number of nucleotide differences between individuals is 3.74 bp.

KEY WORDS: Mitochondrial DNA; HVR-I; Polymerase chain reaction; Population study; Automatic sequencing.

Problems of Forensic Sciences, vol. LVII, 2004, 88–99

Received 7 April 2004; accepted 19 May 2004

INTRODUCTION

Human mitochondrial DNA is a double-stranded closed circular molecule, approximately 16 569 bp in length, including a 1.1 Kb-long noncoding region called the “control region” or “D-loop containing region” [1]. This represents a highly variable sequence. The variable sites in the control region consist of two hypervariable segments, hypervariable region I and hypervariable region II, of approximately 400 bp each [8]. Human mitochondrial DNA has become a useful tool for forensic studies due to its high copy number, maternal inheritance and high levels of sequence polymorphisms [4, 5]. Sequence polymorphisms are useful markers for personal identification in forensic practice, particularly when genomic DNA can not be analysed because of low quantity or degradation. Furthermore, automated sequence analysis of PCR samples has allowed the use of mtDNA in routine casework.

Knowledge of the frequencies with which certain mitochondrial DNA sequences occur in a given population is of crucial importance for the application of mitochondrial markers in forensic expert investigations [6]. The aim of this study was to determine the sequence of one hypervariable segment –

HVI – in a South-East Polish population of 106 unrelated individuals using direct fluorescent-based sequencing.

MATERIALS AND METHODS

Blood samples from 106 unrelated individuals from South-East Poland were extracted by the phenol–chloroform method [7] and quantified by the QuantiBlot kit (Applied Biosystems).

Primer L15996: 5'-CTCCACCATTAGCACCCAAAGC-3' and primer H 16401: 5'-TGATTTACGGAGGATGGTG-3' were used to amplify the mitochondrial HVI region. PCR amplification was performed in 100 µl reaction containing: 2 ng of DNA template, 0.5 µM of each primer, 200 µM of dNTPs, 1.5 mM of MgCl₂, 10 µl of 10 × PCR buffer and 2 U of AmpliTaq Gold polymerase (Applied Biosystems). PCR was carried out in a thermal cycler (TriBlock, Biometra) using the initial heat soak of 9 min denatured at 95°C, followed by 36 cycles of 94°C for 40 s, 55°C for 45 s, 72°C for 60 s; the cycling program was concluded with 72°C incubation for 5 min, followed by storage at 4°C until the samples were collected. The quality of PCR products was investigated using gel electrophoresis in 2% Nusieve agarose after 45 min of migration at 100 V in 1 × TBE and visualised by ethidium bromide staining.

PCR products were purified using Microcon columns (YM-100, Millipore Co., Bedford, MA, USA) and then both strands were sequenced using Cy5 labelled primers (AmliCycle Sequencing Kit, Applied Biosystems). The ALFexpress II DNA sequencer (Amersham Pharmacia Biotech) was used for separation and detection of chain termination products. Sequences were aligned and compared with the Cambridge reference sequence (GeneBank accession: M63933) by BLAST search.

Genetic diversity was calculated according to Tajima's method [10] and random match probability was calculated according to that of Stoneking [9].

RESULTS AND DISCUSSION

Table I shows variable nucleotide positions of the mtDNA HVI region of 106 unrelated people from South-East Poland compared to the reference sequence. In this region, 65 different haplotypes were observed in 106 individuals of which 57 were unique, and 8 were found in at least 2 persons. In this region, 20% of the haplotypes present were identical with the Anderson sequence. In the total length of 370 bp, we obtained 67 variable positions.

TABLE I (CONT.)

Anders.	16221	16222	16223	16224	16234	16235	16243	16248	16256	16264	16265	16270	16271	16278	16288	16290	16291	16292	16293	16294	16295	16296
	C	C	C	T	C	A	T	C	C	C	A	C	T	C	T	C	C	C	A	C	C	C
1	.	.	.	C
2	T	.	.	T	T	T	.
3
4
5	C
6	T	.	.	T	.	.	T	.	T
7
8
9	T	.	.	T
10	.	.	.	C
11	C
12	.	.	.	C
13	T	.	.
14	T
15	T	.	T
16	T
17	.	.	.	C
18	T	.	.	T	T	.	.
19
20	T
21
22
23	T
24	T	.	.	T
25	T	.	.	T	T
26	T	.	.	T	T
27
28
29
30
31	.	.	T	C	T
32
33
34	G
35
36	.	.	T	.	T	.	T
37
38	.	.	T	.	T	T
39	T	T	.	T	.	.
40
41	T	.	.	T
42	.	.	T	.	T	C

Table II shows the distribution of nucleotide substitutions. The majority of mutations were transitions (94%), which is similar to other European Caucasians e.g. German [6] and Spanish [3]. We have a significantly lower frequency of transversions (5%), whereas an insertion was observed only once and deletions were not detected. Among the transitions, pyrimidine substitutions are predominantly found and C→T transitions occur with a particularly high frequency. Among the transversions there is no obvious preference for a particular type. However, we didn't observe single G→T and G→C changes. The majority of polymorphic sites are affected by only one mutation type. At only 3 (16256, 16293 and 16304) out of 67 positions (4.5%), were two different mutation events observed. In our investigations we found one insertion of a single C at np 16193.1. This type of polymorphism is probably due to instability within a stretch of C's between nps 16184 and 16193 [2].

TABLE II. OBSERVED NUCLEOTIDE SUBSTITUTIONS AND INSERTION FOUND IN THE HVS I ON THE L-STRAND OF THE CONTROL REGION OF MTDNA FOR 106 UNRELATED PERSONS FROM SOUTH-EAST POLAND

Mutation type	Numbers of positions	Total number of mutations
Transitions Py→Py		
C→T	29	87
T→C	16	76
Transitions Pu→Pu		
A→G	10	14
G→A	6	7
Total transitions	61	184
Transversions		
A→C	1	1
C→A	2	2
C→G	2	2
T→G	1	1
A→T	2	2
T→A	1	1
Total transversions	9	9
Insertion		
+ C	1	1

Pu – purine base; Py – pyrimidine base; A – adenine; C – cytidine; G – guanine; T – thymine.

A pairwise comparison of 106 sequences detected an average number of 3.74 bp differences between sequences. The genetic diversity was found to be 0.94 and probability of two randomly selected individuals from a population having identical mtDNA types – 0.06.

References:

1. Anderson S., Bankier A. T., Barrell B. G. [et al.], Sequence and organization of the human mitochondrial genome, *Nature* 1981, vol. 290, pp. 457–465.
2. Bendall K. E., Sykes B. C., Length heteroplasmy in the first hypervariable segment of the human mtDNA control region, *American Journal of Human Genetics* 1995, vol. 57, pp. 248–256.
3. Crespillo M., Luque J. A., Parades M. [et al.], Mitochondrial DNA sequences for 118 individuals from northeastern Spain, *International Journal of Legal Medicine* 2000, vol. 114, pp. 130–132.
4. Giles R. E., Blanc H., Cann H. M. [et al.], Maternal inheritance of human mitochondrial DNA, *Proceedings National Academy of Science* 1980, vol. 77, pp. 6715–6719.
5. Jorde L., Barnshad M., Questioning evidence for recombination in human mitochondrial DNA, *Science* 2000, vol. 288, p. 1931.
6. Lutz S., Weisser H. J., Heizmann J. [et al.], Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany, *International Journal of Legal Medicine* 1998, vol. 111, pp. 67–77.
7. Sambrook J., Fritsch E. F., Maniatis T., Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989.
8. Seo Y., Stradmann-Bellinghausen B., Rittner C. [et al.], Sequence polymorphism of mitochondrial DNA control region in Japanese, *Forensic Science International* 1998, vol. 97, pp. 155–164.
9. Stoneking M., Hedgecock D., Higouchi R. G. [et al.], Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes, *American Journal of Human Genetics* 1991, vol. 48, pp. 370–382.
10. Tajima F., Statistical method for testing the neutral mutation hypothesis by DNA polymorphism, *Genetics* 1989, vol. 123, pp. 585–595.

DANE POPULACYJNE REGIONU HVI MITOCHONDRIALNEGO DNA DLA REGIONU POLSKI POŁUDNIOWO-WSCHODNIEJ

Dorota MONIES, Marzanna CIESIELKA, Roman MADRO

WSTĘP

Mitochondrialny DNA (mtDNA) człowieka jest dwuniciową, kolistą molekułą o wielkości ok. 16 569 pz, zawierającą niekodujący („region kontrolny” lub „pętla D”) region o długości 1100 pz, który składa się z dwóch wysoce zmiennych segmentów: HVI i HVII (ang. hypervariable I, II) liczących po około 400 pz [1, 8]. Występuje on w komórce w dużej ilości kopii, dziedziczony jest w linii matczynej i wykazuje wysoki polimorfizm [4, 5]. Stanowi więc skuteczne narzędzie do osobniczej identyfikacji, zwłaszcza, gdy analiza DNA jądrowego jest niemożliwa (z powodu jego wysokiej degradacji lub zbyt małej ilości).

Wprowadzenie każdego nowego markera do rutynowych ekspertyz genetycznych musi być jednak poprzedzone badaniami populacyjnymi [6]. Celem pracy była zatem analiza zmian sekwencji w obrębie regionu HVI u 106 niespokrewnionych osób.

MATERIAŁY I METODY

Materiał do badań stanowiło 106 próbek, które uzyskano od niespokrewnionych osób z Polski południowo-wschodniej. DNA izolowano metodą organiczną [7], a jego stężenie określono przy użyciu zestawu QuantiBlot (Applied Biosystems). Jako starterów użyto:

1. L15996: 5'-CTCCACCATTAGCACCCAAAGC-3';
2. H16401: 5'-TGATTTACGGAGGATGGTG-3'.

100 µl mieszaniny reakcyjnej PCR zawierało: 2 ng DNA matrycowego, 0,5 µM każdego startera, 200 µM dNTP, 1,5 mM MgCl₂, 10 µl dziesięciokrotnie stężonego PCR bufor i 2 U AmpliTaq Gold polimerazy (Applied Biosystems). Reakcję przeprowadzono w termobloku (TriBlock, Biometra), stosując denaturację wstępną przez 9 min w 95°C, a następnie: 94°C – 40 s, 55°C – 45 s i 72°C – 60 s w 36 cyklach oraz wydłużanie końcowe przez 5 min w temperaturze 72°C. Uzyskany materiał przechowywano w temperaturze 4°C. Obecność amplikonów potwierdzano, rozdzielając produkty PCR, na 2% żelach agarozowych (przez 45 min w buforze 1 × TBE, przy 100 V), które wybarwiano bromkiem etydyny.

Tak zweryfikowane produkty PCR oczyszczano na kolumnach typu Microcon (YM-100, Millipore Co., Bedford, Stany Zjednoczone), po czym sekwencjonowano w obu kierunkach przy pomocy ALFexpress II (Amersham Pharmacia Biotech) z użyciem zestawów AmpliCycle Sequencing Kit (Applied Biosystems) oraz starterów wyznakowanych Cy5, a uzyskane sekwencje porównywano z referencyjną sekwencją Cambridge (GeneBank accession: M63933) przy pomocy programu BLAST.

Zmienność genetyczną (ang. genetic diversity) obliczono zgodnie z metodą Tajima [10], a prawdopodobieństwo przypadkowej zgodności (ang. random match probability) w sposób podany przez Stonekinga [9].

WYNIKI I ICH DYSKUSJA

Analizowano sekwencję o długości 370 pz, w obrębie której stwierdzono zmiany w 67 różnych pozycjach nukleotydowych (tabela I). Zidentyfikowano 65 różnych haplotypów, z których 57 pojawiło się tylko raz, a 8 wystąpiło co najmniej u dwóch osób. Całkowitą zgodność z sekwencją Andersona wykazało 20% haplotypów.

W tabeli II przedstawiono rodzaje mutacji. Podobnie jak w innych populacjach europejskich, np. w niemieckiej [6] ewentualnie hiszpańskiej [3], w większości (94%) przypadków były to tranzycje. Częstość transwersji wynosiła 5%. Stwierdzono tylko jedną insercję, nie wykryto natomiast żadnej delecji. Większość tranzycji polegała na substytucji C → T. W zakresie transwersji nie zidentyfikowano zmian: G → T i G → C, a w obrębie pozostałych możliwości nie zaobserwowano żadnych preferencji. Insercja polegała zaś na wstawieniu C między pozycjami 16193 i 16194, co prawdopodobnie wiąże się z niestabilnością odcinka 16184–16193 utworzonego przez łańcuch C [2]. Polimorficzność 64 pozycji była przy tym efektem jednego rodzaju mutacji. Jedynie w trzech miejscach (16256, 16293, 16304) wystąpiły dwa typy zmian.

W 106 sekwencjach pojedyncza zmiana nukleotydowa występowała średnio co 3,74 pz. Zmienność genetyczna wynosi zatem 0,94, a prawdopodobieństwo przypadkowego wystąpienia w tej populacji dwóch identycznych wariantów mtDNA – 0,06.