



## FREQUENCY ASSESSMENT OF 19 SNPs IN A POPULATION FROM SOUTH-WEST POLAND

Iwona CHROMIK, Małgorzata MAŁODOBRA, Tadeusz DOBOSZ

*Department of Forensic Medicine, Molecular Technique Unit, Medical University of Wrocław, Poland*

### Abstract

This study presents the allele frequency distribution of 19 autosomal single nucleotide polymorphisms (SNPs). Allele and genotype frequencies of previously selected markers were assessed in 100 unrelated individuals originating from South-West Poland. No deviations from Hardy-Weinberg equilibrium (*HWE*) were revealed at any of the investigated loci within the analyzed population. All forensically relevant statistical parameters (i.e.  $Het_o$ ,  $Het_{exp}$ ,  $PD$ ,  $PE$ ,  $PIC$ ,  $PD_{total}$ ,  $PE_{total}$ ) were estimated as well. Finally, the results confirmed that these 19 loci are useful markers for forensic identification analysis.

### Key words

Single nucleotide polymorphism; SNP; Human identification; Forensic genetics; SNP genotyping.

*Received 21 May 2012; accepted 23 November 2012*

### 1. Introduction

Despite the fact that genetic analysis of short tandem repeat (STR) loci is currently the most powerful and widely presented method used in human identification caseworks, the use of single nucleotide polymorphisms in such analyses, especially degraded DNA, has generated a robust discussion among forensic scientists. In 2003 a large project called *SNPforID*, involving 5 partnership countries, was initiated. The main goal of this project was to create a universal SNP panel for human identification in all populations [12, 13]. In 2006, the *SNPforID* consortium reported the results of an inter-laboratory exercise. A total of 11 European and one US forensic genetic laboratories tested a subset of a 52 SNP-multiplex PCR kit. The results showed that the total SNP locus dropout was at a level of 2.8%, and in poor quality samples it was at a level of more than 50%. The overall rate of discrepant SNP allele assignments was 2%. Two laboratories reported 60% of all the discrepancies [17]. The outcomes of the collaborative exercise were assessed by the authors as

surprisingly good, but they also showed how hard it is to create an objective universal SNP panel for large scale population analysis. Additionally, SNPs very often show dissimilar frequencies in different populations, causing problems with creating an optimal SNP multiplex designed for global use. Hence the aim of the presented work was: selection of new SNPs from the NCBI database, examination of the allele frequency distribution in a population from South-West Poland and also assessment of statistical parameters.

### 2. Materials and methods

#### 2.1. Population and SNPs selection

One hundred unrelated individuals originating from South-West Poland were tested (DNA Bank, Molecular Technique Unit of the Medical University of Wrocław). According to the main criteria of SNP selection [15, 18] (i.e. the size of the amplicon generated from an optimally designed primer of length

less than 50 bp, reported heterozygosity in the range 0.3–0.6 units in the European population, localization in intron parts of the chromosome, flanking DNA sequence free from interfering polymorphisms, such as nucleotide substitution or polynucleotide sequences in potential primer binding sites), 18 biallelic and one triallelic SNPs were selected from the NCBI database (<http://www.ncbi.nlm.nih.gov/snp>) and the exact chromosomal locations were determined using BLAST (<http://genome.ucsc.edu/cgi-bin/hgBlat>).

## 2.2. DNA extraction and SNP genotyping

DNA from blood samples was extracted using the QIAamp DNA Mini Kit (Qiagen), according to the manufacturer's instructions. 19 SNPs were amplified by PCR in 10  $\mu$ l reactions containing 5  $\mu$ l Multiplex PCR Master Mix (Qiagen), 1  $\mu$ l mixture of all primers with a final concentration of 0.2  $\mu$ M each, 1 ng template DNA and 3  $\mu$ l H<sub>2</sub>O. The PCR protocol was as follows: initial activation at 95°C for 15 min, followed by 30 cycles at 94°C for 30 s, 55°C for 60 s, 72°C for 60 s, and the final extension at 72°C for 10 min. The thermal cycling program was carried out on a GeneAmp 9700 (Applied Biosystems). Excess of PCR primers and dNTP were removed by the addition of 1  $\mu$ l Shrimp Alkaline Phosphatase (SAP, 1 U/ $\mu$ l, Fermentas) and 0.4  $\mu$ l Exonuclease I (Exo I, 20 U/ $\mu$ l, Fermentas) to 4  $\mu$ l PCR product and incubation at 37°C for 30 min and then at 80°C for 15 min. Single base extension (SBE) reactions were performed in 5  $\mu$ l reaction volumes containing 0.8  $\mu$ l purified PCR product, 1.5  $\mu$ l SNaPhot reaction mix (Applied Biosystems), 0.25  $\mu$ l SBE primer mix (0.1–1  $\mu$ M of each primer) and 2.45  $\mu$ l H<sub>2</sub>O. Extension reactions were incubated as follows: 25 cycles of 96°C for 10 s, 57°C for 5 s and 60°C for 30 s. Excess nucleotides were removed by addition of 1  $\mu$ l Shrimp Alkaline Phosphatase to the SBE mix and incubated at 37°C for 60 min, then 80°C for 15 min. 1  $\mu$ l SBE products was mixed with 10  $\mu$ l Hi-Di formamide (Applied Biosystems) and 0.35  $\mu$ l GeneScan-120 LIZ internal standard (Applied Biosystems). The SBE products were analyzed by a 4-capillary ABI PRISM 3130 Genetic Analyzer (Applied Biosystems) using filter set E5. After electrophoresis, the data was analyzed by GeneMapper ID v3.2 software (Applied Biosystems). The sets of primers used in the PCR and SBE reaction are presented in Table I.

## 2.3. Statistical analysis

Allele frequencies and statistical parameters of the selected SNPs were calculated with POWER STATS

v.12 (Promega). Divergence from Hardy-Weinberg expectations (*HWE*) was determined by the  $\chi^2$  test. The following statistical parameters: expected heterozygosity ( $H_{exp}$ ), observed heterozygosity ( $H_o$ ) [14], power of discrimination (*PD*) [8], power of exclusion (*PE*) [4, 20] and polymorphism information content (*PIC*) [2] were estimated.

## 3. Results and discussion

The European SNPforID consortium has developed a multiplex of 52 SNPs for forensic purposes, with the amplification of all 52 loci in a single reaction followed by two single base extension (SBE) reactions, which are detected with capillary electrophoresis. Such SNPs as suggested by the SNPforID scientists are promising. The number of scientists involved in this large project is still increasing [3, 9, 16]. However, the dissimilar frequencies of the same SNP loci analysed in different populations are inclining scientists towards developing new sets of SNPs [10, 11] as well as towards initiating other large projects aimed at developing universal SNPs for worldwide individual identification [22].

In this study, allele frequencies of a new 19-loci SNP were estimated. Calculations as well as statistical parameters are summarized in Table II. No deviations from HWE were observed in any locus ( $P > 0.05$ ). The total power of discrimination is  $PD_{total} = 0.999999921$ , which means the performance of the present system is adequate for application to forensic identification. Other parameters, such as total power of exclusion and polymorphism information content mean that the present set of SNPs qualifies for paternity testing, but according to Gill [6], the minimum number of SNPs in a multiplex panel which gives reliable results is 50–60 SNP with allele frequencies 0.3–0.7. Moreover, the number of SNPs used in human identification cases giving reliable results with similar match probability to commonly used STR panels (such as AmpF/STR®Identifiler™ or PowerPlex®) is 40–50, with an optimal allele frequency of approximately 0.5 [1, 5]. This proves that reasonable numbers of loci are able to separate individuals from one another. SNPs may be amplified from only 100 pg of genomic DNA and the length of the amplicons ranges from as short as 40 bp (primer size plus one polymorphic nucleotide) upwards. The high sensitivity and the short amplicon sizes make the assay very suitable for typing degraded DNA samples, and the low mutation rate of SNPs makes them very useful for the identification of disaster victims or relationship testing [7, 21].

TABLE I. SETS OF PRIMERS USED IN PCR AND SNAPSHOT REACTION

SNP (rs)	Sequence of primers	PCR product size (bp)	SNaPshot product size (nt)	Final concentration in SNaPshot reaction [ $\mu$ M]
41397044 (C/T)	F tctccaattacacaacggatc	47	–	–
	R ggccactgtaaatctagatg		23	1.0
45504994 C/G	F tgcaaactgattgtctgagcaga	48	–	–
	R attgggagtaattgagcctgactg		24	0.2
7617204 A/G	F ttctccagactgatacgcgttc	49	27	1.0
	R cgatgcctgataaaagcacaattac		–	–
8176044 A/G	F gatgactacagttctctaccta	48	25	1.0
	R ttgagttgtgggggtctgaa		–	–
7026 C/T	F agctattttacctcccgcctccc	49	25	1.0
	R aagacgtcaggtggggggaccagca		–	–
41424948 C/T	F agccaagttaggcaagcttagtc	48	25	0.4
	R ggaccaattcatctgaaggcttg		–	–
2741098 C/T	F gctgtaggaggtgcagcatca	43	–	–
	R catccaggcattgctccttg		22	0.4
11082466 A/G	F gagctcaagtgccttctcc	41	21	1.0
	R agtttcttgccaagaagcta		–	–
3757853 C/G	F aagcatccccagccttaccat	46	23	0.5
	R ggcacatgagatgaggagatgaa		–	–
3760578 A/G	F tgacttcttatgcaaatgg	41	21	1.0
	R atttctatcaggacagtgg		–	–
2738077 C/G	F catagccctgctcctgggta	41	21	1.0
	R agactccaggctgctgagga		–	–
9274703 A/C/G	F tcagaattggaggatgatggag	46	23	0.3
	R ttcccatactcctactccttca		–	–
12075 A/G	F tccttccagatggagacgactatg	44	23	0.5
	R gcagctgcttccaggttgca		–	–
2553728 C/T	F ttcaggttctgcccctttg	43	22	0.8
	R tcccaaattatcttctctac		–	–
1135062 A/G	F gccatgttctcacttcggc	42	22	0.4
	R accacctcagtcactcaccgg		–	–
7630674 A/G	F gaaactgtgagaacagttagtgt	49	–	–
	R cagcttttaatcagcagcttct		25	1.0
9955503 A/T	F taatctctattgaacatagga	44	23	1.0
	R ggtgattaatctatccatgc		–	–
8090908 A/G	F aacatagtggtccagatcttct	44	23	1.0
	R cagcttgggcctggttacag		–	–
2741083 C/T	F gtgaggagtctggattggc	44	24	1.0
	R actgctgaagaaattcatalcgt		–	–

TABLE II. ALLELE FREQUENCIES AND STATISTICAL PARAMETERS OBTAINED FOR 19-LOCI SNPS IN 100 UNRELATED INDIVIDUALS FROM SOUTH-WEST POLAND

SNP (rs)	Allele 1	Allele 2	Allele 3	$Het_o$	$Het_{exp}$	$PD$	$PE$	$PIC$	$HWE$ (p-value)
41397044 C/T	0.615	0.385	–	0.590	0.474	0.541	0.279	0.36	> 0.97
45504994 C/G	0.495	0.505	–	0.550	0.500	0.596	0.235	0.37	> 0.99
7617204 A/G	0.545	0.455	–	0.650	0.500	0.512	0.355	0.37	> 0.95
8176044 A/G	0.600	0.400	–	0.500	0.480	0.605	0.188	0.36	> 0.99
7026 C/T	0.650	0.350	–	0.580	0.455	0.626	0.133	0.36	> 0.96
41424948 C/T	0.725	0.275	–	0.450	0.399	0.545	0.147	0.32	> 0.99
2741098 C/T	0.385	0.615	–	0.430	0.474	0.626	0.133	0.36	> 0.99
11082466 A/G	0.430	0.570	–	0.640	0.490	0.516	0.342	0.37	> 0.95
3757853 C/G	0.455	0.545	–	0.650	0.496	0.512	0.355	0.37	> 0.95
3760578 A/G	0.445	0.555	–	0.650	0.494	0.502	0.369	0.37	> 0.95
2738077 C/G	0.450	0.550	–	0.460	0.495	0.638	0.155	0.37	> 0.99
9274703 A/C/G	0.315	0.375	0.310	0.730	0.663	0.766	0.488	0.59	> 0.99
12075 A/G	0.530	0.470	–	0.640	0.498	0.524	0.342	0.37	> 0.96
2553728 C/T	0.380	0.620	–	0.620	0.471	0.492	0.342	0.36	> 0.95
1135062 A/G	0.705	0.295	–	0.530	0.416	0.525	0.215	0.33	> 0.96
7630674 A/G	0.295	0.705	–	0.430	0.416	0.569	0.133	0.33	> 0.99
9955503 A/T	0.360	0.640	–	0.440	0.461	0.610	0.140	0.35	> 0.99
8090908 A/G	0.325	0.675	–	0.450	0.439	0.585	0.188	0.34	> 0.99
2741083 C/T	0.625	0.375	–	0.550	0.469	0.565	0.235	0.36	> 0.98
$PD_{total}$	0.999999921								
$PE_{total}$	0.996622								

$Het_o$  – observed heterozygosity,  $Het_{exp}$  – expected heterozygosity,  $PD$  – power of discrimination,  $PE$  – power of exclusion,  $PIC$  – polymorphic information content,  $PE_{total}$  – total power of exclusion total,  $PD_{total}$  – total power of discrimination.

This paper is the second part of a description of a large project whose main goal is to create an SNP panel with the recommended number of SNPs. The first part, describing a heptaplex SNP, was published in a previous paper [19]. Newly selected 7-loci SNPs and 19-loci SNPs were analyzed in the Polish population for the first time. The results presented for a South-West Polish population are comparable with data for the European population accessible in the NCBI database (<http://www.ncbi.nlm.nih.gov/pubmed/SNP>). Conclusions suggest potential universal usability of the selected SNP markers for forensic analysis; however, in order to perform a full verification, it is necessary to extend the research to a larger sample, including other European countries.

## References

1. Amorim A., Pereira L., Pros and cons in the use of SNPs in forensic kinship investigation: a comparative analysis with STRs, *Forensic Science International* 2005, 150, 17–21.
2. Botstein D., White R. L., Skolnicki M. [et al.], Construction of a genetic linkage map in man using restriction fragment length polymorphisms, *The American Journal of Human Genetics* 1980, 32, 314–331.
3. Drobic K., Børsting C., Rockenbauer E. [et al.], Typing of 49 autosomal SNPs by SNaPshot in the Slovenian population, *Forensic Science International: Genetics* 2010, 5, e125–127.
4. Fung W. K., Chung Y. K., Wong D. M., Power of exclusion revisited: probability of excluding relatives of the true father from paternity, *Internal Journal of Legal Medicine* 2002, 116, 64–67.

5. Gill P., Werrett D. J., Budowle B. [et al.], An assessment of whether SNPs will replace STRs in national DNA databases – Joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDM), *Science & Justice*, 2004, 44, 51–53.
6. Gill P., An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes, *Internal Journal of Legal Medicine* 2001, 114, 204–210.
7. Hughes-Stamm S. R., Ashton K. J., van Daal A., Assessment of DNA degradation and the genotyping success of highly degraded samples, *Internal Journal of Legal Medicine* 2011, 125, 341–348.
8. Jones D. A., Blood samples: Probability of discrimination, *Journal of the Forensic Science Society* 1972, 12, 355–359.
9. Khodjet-el-Khil H., Fadhlaoui-Zid K., Cherni L. [et al.], Genetic analysis of the SNPforID 34-plex ancestry informative SNP panel in Tunisian and Libyan populations, *Forensic Science International: Genetics* 2011, 5, e45–47.
10. Kim J. J., Han B. G., Lee H. I. [et al.], Development of SNP-based human identification system, *Internal Journal of Legal Medicine* 2010, 124, 125–131.
11. Lou C., Cong B., Li S. [et al.], A SNaPshot assay for genotyping 44 individual identification single nucleotide polymorphisms, *Electrophoresis* 2011, 32, 368–378.
12. Musgrave-Brown E., Ballard D., Balogh K. [et al.], Forensic validation of the SNPforID 52-plex assay, *Forensic Science International: Genetics* 2007, 1, 186–190.
13. Musgrave-Brown E., Ballard D., Fondevila M. [et al.], Forensic validation of the Genplex SNP typing system – Result of an inter-laboratory study, *Forensic Science International: Genetics, Supplement Series* 2008, 1, 389–393.
14. Nei M., Roychoudhury A. K., Sampling variances of heterozygosity and genetic distance, *Genetics* 1974, 76, 379–390.
15. Phillips C., Lareu V., Sanchez M. [et al.], Selecting single nucleotide for forensic applications, *International Congress Series* 2004, 1261, 18–20.
16. Ruiz Y., Chiurillo M. A., Borjas L. [et al.], Analysis of the SNPforID 52-plex markers in four Native American populations from Venezuela, *Forensic Science International: Genetics* 2012, 6, e142–e145.
17. Sanchez J. J., Børsting C., Balogh K. [et al.], Forensic typing of autosomal SNPs with a 29 SNP-multiplex – results of a collaborative EDNAP exercise, *Forensic Science International: Genetics* 2008, 2, 176–183.
18. Sanchez J. J., Phillips C., Børsting C. [et al.], A multiplex assay with 52 single nucleotide polymorphisms for human identification, *Electrophoresis* 2006, 27, 1713–1724.
19. Sękowska M., Musiał D., Lebioda D., Chromik I. [et al.], Heptaplex SNP-an allele frequencies database of the south-west Poland population. Announcement of population data, *Problems of Forensic Science* 2010, 82, 184–189.
20. Weir B. S., Genetic data analysis II. Methods for discrete population genetic data, Sinauer Associates, Sunderland 1996.
21. Westen A., Matai A., Laros J. [et al.], Tri-allelic SNP markers enable analysis of mixed and degraded DNA samples, *Forensic Science International: Genetics* 2009, 3, 233–241.
22. Zeng Z., Wang L., Feng Q. [et al.], Evaluation of 96 SNPs in 14 populations for worldwide individual identification, *Journal of Forensic Science* 2012, 57, 1031–1035.

---

**Corresponding author**

Dr n. med. Iwona Chromik  
Uniwersytet Medyczny we Wrocławiu  
Katedra Medycyny Sądowej  
M. Curie-Skłodowskiej 52  
PL 50-369 Wrocław  
e-mail: chromik\_iwona@yahoo.pl

---



## OSZACOWANIE CZĘSTOŚCI WYSTĘPOWANIA 19 POLIMORFIZMÓW TYPU SNP W POPULACJI POLSKI POŁUDNIOWO-ZACHODNIEJ

### 1. Wstęp

Niezależnie od faktu, że genetyczna analiza krótkich powtarzających się tandemowo elementów DNA (STR) jest obecnie najbardziej powszechną i zapewniającą ogromną siłę dyskryminacji metodą stosowaną do identyfikacji człowieka, to jednak zastosowanie polimorfizmów zmian pojedynczego nukleotydu w takich analizach, szczególnie w przypadkach zdegradowanego DNA, powoduje intensywną dyskusję wśród genetyków sądowych. W 2003 roku zainicjowano ogromny projekt nazwany SNPforID realizowany przez 5 współdziałających ze sobą krajów. Głównym celem tego projektu było zaprojektowanie uniwersalnego panelu SNP do identyfikacji człowieka, wiarygodnego we wszystkich populacjach [12, 13]. W 2006 roku konsorcjum SNPforID przedstawiło wyniki międzynarodowych testów międzylaboratoryjnych. Jedenaście europejskich laboratoriów (i jedno amerykańskie) sprawdzało zestaw 52 SNP-multiplex PCR. Otrzymane rezultaty wykazały, że całkowite tempo wypadania (ang. drop out) alleli SNP *loci* było na poziomie 2,8%, a w ponad 50% obserwowano to zjawisko w próbkach bardzo złej jakości. Całkowity wskaźnik rozbieżności pomiędzy allelami SNP wyniósł 2%. Dwa laboratoria wykazały aż 60% rozbieżności [17]. Wyniki wspólnych doświadczeń zostały przez autorów ocenione jako zdumiewająco dobre, ale jednak wykazały, jak bardzo trudno jest uzyskać obiektywny, uniwersalny panel *loci* SNP do analizy populacyjnej o szerokim zakresie. Co więcej, polimorfizmy typu SNP bardzo często wykazują odmienne częstości pomiędzy różnymi populacjami, powodując tym samym problemy z zaprojektowaniem optymalnego multipleksu SNP do powszechnego zastosowania. Stąd celem tej pracy była selekcja nowych polimorfizmów typu SNP z bazy NCBI, sprawdzenie rozkładu częstości w populacji Polski południowo-zachodniej, a także oszacowanie ich parametrów statystycznych.

### 2. Materiał i metody

#### 2.1. Wybór populacji i polimorfizmów typu SNP

Przetestowano sto niespokrewnionych osób pochodzących z Polski południowo-zachodniej (Bank DNA, Pracownia Technik Molekularnych Uniwersytetu Medycznego we Wrocławiu). Zgodnie z głównymi kryteriami wyboru SNP [15, 18] (tj. długość amplikonu otrzymanego z optymalnie zaprojektowanego primeru o długości mniejszej niż 50 par zasad, znanej heterozygotyczności

w zakresie 0,3–0,6 w populacji europejskiej, lokalizacji w intronowej części chromosomu, sekwencji flankującej wolnej od dodatkowych polimorfizmów, takich jak substytucje nukleotydowe lub polinukleotydowe sekwencje w potencjalnym miejscu wiązania primerów), wyselekcjonowano z bazy danych NCBI (<http://www.ncbi.nlm.nih.gov/snp>) 18 dwuallelicznych i 1 trialleliczny polimorfizm typu SNP, a dokładne lokalizacje na chromosomach zostały ustalone przy użyciu BLAST (<http://genome.ucsc.edu/cgi-bin/hgBlat>).

#### 2.2. Izolacja DNA i genotypowanie polimorfizmów SNP

DNA z próbek krwi wyizolowano przy użyciu QIAmp DNA Mini Kit (Qiagen) zgodnie z instrukcją producenta. 19 polimorfizmów SNP amplifikowano w reakcji PCR w 10  $\mu$ l objętości reakcji zawierającej 5  $\mu$ l Multiplex PCR Master Mix (Qiagen), 1  $\mu$ l mieszaniny wszystkich primerów z końcowym stężeniem 0,2  $\mu$ M każdego primeru, 1 ng matrycowego DNA i 3  $\mu$ l H<sub>2</sub>O. Protokół PCR był następujący: aktywacja początkowa w 95°C przez 15 min z następującymi po niej 30 cyklami w temperaturze 94°C przez 30 s, 55°C przez 60 s, 72°C przez 60 s oraz przy końcowym wydłużaniu w 72°C przez 10 min. Program cyklicznego profilu temperaturowego był przeprowadzany za pomocą urządzenia GeneAmp9700 (Applied Biosystems). Nadmiar primerów reakcji PCR i dNTP usuwano poprzez dodanie 1  $\mu$ l Shrimp Alkaline Phosphatase (SAP, 1 U/ $\mu$ l, Fermentas) i 0,4  $\mu$ l Exonuclease I (Exo I, 20 U/ $\mu$ l, Fermentas) do 4  $\mu$ l produktu, a następnie poddano inkubacji w 37°C przez 30 min, po czym w 80°C przez 15 min. Reakcji wydłużania pojedynczych zasad (SBE) dokonano w objętości 5  $\mu$ l zawierającej 0,8 oczyszczonego produktu PCR, 1,5  $\mu$ l mieszaniny reakcyjnej SNaPshot (Applied Biosystems), 0,25  $\mu$ l mieszaniny primerów SBE (0,1–1  $\mu$ M każdego z primerów) i 2,45  $\mu$ l H<sub>2</sub>O. Reakcje wydłużania przeprowadzono, jak następuje: 25 cykli w temperaturach 96°C przez 10 s, 57°C przez 5 s i 60°C przez 30 s. Nadmiar nukleotydów był usuwany poprzez dodanie 1  $\mu$ l Shrimp Alkaline Phosphatase do mieszaniny SBE i inkubację w 37°C przez 60 min, a następnie w 80°C przez 15 min. 1  $\mu$ l produktów SBE zmieszano z 10  $\mu$ l formamidu Hi-Di (Applied Biosystems) i 0,35  $\mu$ l wewnętrznego standardu GeneScan-120 LIZ (Applied Biosystems). Produkty SBE analizowano przy użyciu 4-kapilarowego analizatora ABI PRISM 3130 (Applied Biosystems) za pomocą zestawu filtrów E5. Po zakończeniu elektroforezy analizowano dane przy użyciu oprogramowania GeneMapper ID v.3.2

(Applied Biosystems). Zestaw primerów zastosowanych do reakcji PCR i SBE zamieszczono w tabeli I.

### 2.3. Analiza statystyczna

Częstości alleli i parametry statystyczne wybranych zmiennych miejsc typu SNP obliczono przy użyciu programu POWER STATS v. 12 (Promega). Odchylenie od oczekiwanej równowagi Hardy-Weinberga (*HWE*) określono za pomocą testu  $\chi^2$ . Oszacowano następujące parametry statystyczne: oczekiwaną heterozygotyczność ( $H_{exp}$ ), obserwowaną heterozygotyczność ( $H_o$ ), [14], siłę dyskryminacji (*PD*) [8], siłę wykluczenia (*PE*) [4, 20] oraz zawartość polimorficznej informacji (*PIC*).

## 3. Wyniki i dyskusja

Europejskie konsorcjum SNPforID opracowało reakcję multipleksową 52 SNP na potrzeby analizy sądowej. Umożliwia ona amplifikację 52 *loci* w jednej reakcji z następującymi po niej dwoma reakcjami wydłużania pojedynczych nukleotydów (SBE), których produkty podlegają detekcji przy użyciu elektroforezy kapilarnej. Tego rodzaju zmienne miejsca typu SNP sugerowane przez naukowców z SNPforID wydają się obiecujące. Liczba osób zaangażowanych w ten projekt wciąż rośnie [3, 9, 16]. Jednakże odmienne częstości tych samych zmiennych miejsc SNP analizowanych w różnych populacjach skłaniają zaangażowanych w projekt naukowców do opracowania nowych zestawów SNP [10, 11], a także do tworzenia nowych projektów celem opracowania uniwersalnych polimorfizmów typu SNP do identyfikacji osób [22]. W tej pracy oszacowano częstości alleli nowych 19 *loci* SNP. Obliczenia, jak i statystyczne parametry, zawarto w tabeli II. W żadnym *locus* nie zaobserwowano odchylenia od równowagi Hardy-Weinberga ( $P > 0,05$ ). Całkowita siła dyskryminacji wynosiła  $PD_{total} = 0,999999921$ , co oznacza, że badany zestaw jest odpowiedni do identyfikacji człowieka na potrzeby sądu. Inne parametry, takie jak całkowita siła wykluczenia czy zawartość informacji polimorficznej, kwalifikują przedstawiony zestaw do testowania spornego ojcostwa, jednakże zgodnie z Gillem [6], minimalna liczba polimorfizmów typu SNP w panelu multipleksowym wymagana do otrzymania rzetelnych wyników wynosi 50–60 SNP z częstościami alleli od 0,3 do 0,7. Co więcej, liczba polimorfizmów typu SNP zastosowanych w celu identyfikacji człowieka, a dających wiarygodne wyniki z porównywalnymi prawdopodobieństwami przypadkowej zgodności takimi, jakie są osiągnąć przy użyciu powszechnie stosowanych zestawów zawierających polimorfizmy typu STR (AmpFSTR Identifier czy PowerPlex), wynosi 40–50 z optymalną częstością zbliżoną do 0,5 [1,5]. Zapewnia to uzasadnioną liczbę

*loci* niezbędną do odróżnienia jednej osoby od drugiej. Polimorfizmy typu SNP mogą być amplifikowane z zaledwie 100 pg genomowego DNA, a długość amplikonów jest szczególnie niska i sięga od 40 par zasad (długość primera plus polimorficzny nukleotyd). Wysoka czułość i krótkie rozmiary amplikonów powodują, że opracowany zestaw jest szczególnie użyteczny do genotypowania próbek bardzo zdegradowanych, a niskie tempo mutacji zmiennych miejsc typu SNP sprawia, że są one ogromnie użyteczne do identyfikacji ofiar katastrof czy do badania pokrewieństwa [7, 21].

Artykuł ten stanowi drugą część opisu wyników pracy w ramach dużego projektu, którego głównym celem było utworzenie panelu rekomendowanych polimorfizmów typu SNP. Pierwsza część opisująca SNP heptaplex została opublikowana we wcześniejszej pracy [19]. Współcześnie wyselekcjonowane 7-locusowe polimorfizmy typu SNP i panel 19 *loci* typu SNP zostały po raz pierwszy zanalizowane w polskiej populacji. Wyniki prezentowane dla populacji Polski południowo-zachodniej są porównywalne z danymi populacji europejskich, a dostępnymi poprzez bazę danych NCBI (<http://www.ncbi.nlm.nih.gov/pubmed/SNP>). Uzyskane wnioski potwierdzają uniwersalną użyteczność wybranych markerów typu SNP do analizy sądowej, jednakże, aby przeprowadzić pełną ich weryfikację, konieczne jest rozszerzenie badań do większych prób, w tym z innych krajów europejskich.