

Katarzyna Bąbol-Pokora, Adam Prośniak, Renata Jacewicz, Jarosław Berent

Pentapleks SNP – rozkład częstości alleli w populacji centralnej Polski¹

SNP pentaplex – the allele frequency database of central Poland population

Z Katedry i Zakładu Medycyny Sądowej Uniwersytetu Medycznego w Łodzi
Kierownik: prof. zw. dr hab. n. med. Stefan Szram

Analiza SNP jest jedną z najnowszych metod stosowanych do identyfikacji osobniczej w genetyce sądowej, zwłaszcza w analizach spraw nietypowych i w badaniach silnie zdegradowanego materiału dowodowego. W pracy przedstawiono wyniki badań populacyjnych, przeprowadzone metodą minisekwencjonowania, z użyciem pentapleksu biallelicznych markerów SNP (rs2294067, rs2282160, rs2070764, rs2277216, rs1063739), dla populacji centralnej Polski.

SNP analysis is among the most contemporary method for personal identification in forensic genetics, especially in analyses of untypical cases and in studies of degraded DNA samples from crime scenes. This paper shows the results of Central Poland population studies carried out using minisequencing and SNP-pentaplex, which contains 5 biallelic loci rs2294067, rs2282160, rs2070764, rs2277216 and rs1063739. DNA fragments were amplified in one multiplex PCR reaction and SNPs were identified by the minisequencing method. The combined PD was 0.9895253125, which makes the pentaplex useful for forensic applications.

Słowa kluczowe: SNP, minisekwencjonowanie, baza populacyjna
Key words: SNP, minisequencing, population database

WSTĘP

SNP (single nucleotide polymorphism) to utrwalone mutacje punktowe powstające w wyniku tran-

zycji lub transwersji. Występują w całym genomie – zarówno w regionach kodujących, jak i niekodujących – co 100-300 nukleotydów [1]. Ich zastosowanie w genetyce sądowej umożliwia analizę wysoce zdegradowanego materiału biologicznego dzięki temu, że amplifikacji poddawane są znacznie krótsze fragmenty niż w przypadku układów STR. SNP są doskonałym narzędziem do badania śladów z miejsc przestępstw, które są często silnie zdegradowane i analiza układów typu STR nie jest możliwa [2]. Szansa znalezienia w zdegradowanym materiale genetycznym niezdegradowanego fragmentu zawierającego SNP, o długości kilkudziesięciu nukleotydów, jest znacznie większa niż znalezienie fragmentu kilkusetnukleotydowego z układami typu STR, który uda się zamplifikować. Metoda ta ma większą czułość od metod opartych na analizie układów typu STR, dzięki znacznie większej ilości potencjalnych markerów identyfikacji osobniczej, jakimi są SNP.

Do analizy SNP wykorzystano reakcję wydłużania starterów, czyli minisekwencjonowanie. Jest to zmodyfikowana reakcja sekwencjonowania i polega na wydłużaniu specjalnie zaprojektowanych starterów [3, 5, 7] (tabela I) przy użyciu znakowanych fluorescencyjnie ddNTP [3, 4, 5]. Startery przyłączają się końcem 3' do każdego SNP, czyli tuż przed miejscem mutacji, a następnie są wydłużane o jeden komplementarny do SNP dideoksynukleotyd, znakowany odpowiednim, w zależności od zasady, fluorochromem (ryc. 1). Do startera może się przyłączyć tylko jeden ddNTP, ponieważ jest zabloko-

¹ Temat opracowany w ramach prac własnych Uniwersytetu Medycznego w Łodzi nr 502-11-373

wany w 3' końcu. Powstałe w ten sposób fragmenty (starter wydłużony o 1 dideoksynukleotydy) są rozdzielane na żelu i identyfikowane na podstawie barwy znacznika.

Według Brennera [6] zastosowanie 4 SNP w badaniach ojcostwa daje podobną siłę dyskryminacji co użycie 1 układu STR, natomiast w badaniach porównawczych śladów biologicznych 1 układ STR można zastąpić 2,6 loci SNP. Tym samym, pomimo niskiej jednostkowej polimorficzności, analiza statystyczna z użyciem zestawu kilkudziesięciu SNP pozwala na uzyskanie podobnych wyników co badanie kilkunastu ukła-

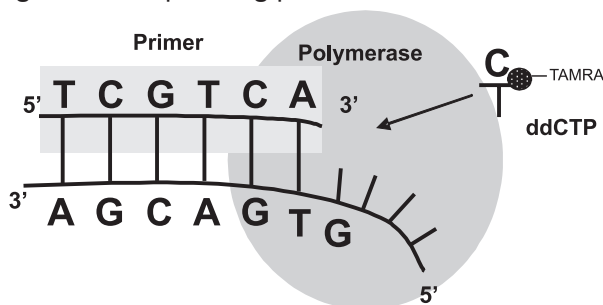
dów typu STR [3, 4, 6]. Przykładowo w analizie śladów biologicznych zestaw SGM Plus może zostać zastąpiony zestawem 26 loci SNP. Metoda z użyciem SNP może zatem zminimalizować problemy związane z identyfikacją silnie zdegradowanego materiału biologicznego, gdzie zastosowanie metod opartych na analizie STR nie daje zadowalających wyników.

Celem badań było praktyczne opracowanie i wdrożenie innowacyjnej metody analizy SNP do analiz identyfikacyjnych oraz opracowanie bazy częstości alleli wybranych pięciu SNP dla populacji centralnej Polski.

Tabela I. Wielkość produktów amplifikacji i minisekwencjonowania badanego pentapleksu SNP.
Table I. PCR and minisequencing product sizes of SNP pentaplex.

SNP	Allel Allele	Locus	Produkt amplifikacji PCR product size	Wielkość produktu minisekwencjonowania Minisequencing product size	Sekwencje starterów do minisekwencjonowania Sequences of primers used for minisequencing
rs2070764	A/T	8q13,3	123 bp	35 nt	5'ttttgattgagctaaaaattatgaataactgat3'
rs2294067	G/C	8p23,3	99 bp	28 nt	5'aggctggctgggaaggggcctctgtgg3'
rs2282160	A/G	9p22	93 bp	50 nt	5'tttttttttttttttttttgggctcagagcaatacactctcttctctcc3'
rs1063739	A/C	8q24,3	85 bp	42 nt	5'ttttttttttagcacaggtgaccggagccaggcggtgagg3'
rs2277216	T/C	10p13	71 bp	57 nt	5'tttttttttttttttttttttttgtctagcctcacctcgacctggaggcac3'

Ryc. 1. Schemat reakcji minisekwencjonowania.
Fig. 1. Minisequencing process scheme.



MATERIAŁ I METODY

Materiał do badań stanowiły próbki krwi pobrane od 80 niespokrewnionych osób biorących udział w standardowych badaniach identyfikacyjnych w Katedrze i Zakładzie Medycyny Sądowej Uniwersytetu Medycznego w Łodzi.

DNA wyizolowano metodą kolumnkową z zastosowaniem zestawu Sherlock AX (A&A Biotechnology). Multipleksowa reakcja amplifikacji została przeprowadzona z użyciem pięciu par starterów (opracowanych wg 3,5,7), obejmujących następujące loci SNP: rs2294067, rs2282160, rs2070764, rs2277216, rs1063739. Amplifikowane fragmenty miały wielkość od 71 bp do 123 bp (tabela I). Produkty PCR zostały oczyszczone za pomocą zestawu kolumniek ultrafiltracyjnych MiniElute® (Qiagen). Do minisekwencjonowania użyto zestawu ABI Prism® SnaPshot™ (Applied Biosystems). Zastosowane w multipleksowej reakcji startery miały różną długość (27 nt – 56 nt) co umożliwiło identyfikację badanych SNP poprzez rozdział elektroforetyczny. Produkty minisekwencjonowania, po oczyszczeniu zestawem DyeEx® (Qiagen), zostały rozdzielone

w poliakrylamidowym żelu denaturującym w sekwencjonatorze ABI Prism 377 z użyciem LIZ™ 120 jako wewnętrznego standardu wielkości. Analiza końcowa – genotypowanie przeprowadzono za pomocą programu Gene Scan™.

Analiza statystyczna, którą przeprowadzono przy użyciu programów komputerowych GDA [9] i DLP [10], objęła zgodność z prawem Hardy’ego-Weinberga (exact p) oraz parametry przydatności pentapleksu SNP w standardowych analizach genetyki sądowej: tj. heterozygotyczność obserwowaną i oczekiwaną (Het obs.; Het exp) oraz jej błąd [11], współczynnik informacji o polimorfizmie (PIC) [12], minimalną i maksymalną [13] oraz średnią szansę ojcostwa (PI) [14], siłę dyskryminacji (PD) [15] oraz siłę wykluczenia dla pełnych (PE_{trio}) [16] i niepełnych trójek (PE_{duo}) [17].

WYNIKI I DYSKUSJA

Utworzona baza populacyjna 5 loci SNP składa się ze 160 alleli. Ocena testem exact wykazała zgodność z prawem Hardy’ego-Weinberga we wszystkich loci. Tabela II przedstawia uzyskany rozkład często-

ści występowania poszczególnych wariantów allelicznych badanych SNP oraz parametry statystyczne obliczone dla każdego SNP w obrębie populacji centralnej Polski. Łączna siła dyskryminacji $PD_{całk}$ pentapleksu SNP wyniosła 0,9895253125, natomiast łączna siła wykluczenia PE_{trio} dla trójek osiągnęła wartość 0,6200206292, a w sprawach bez matki PE_{duo} wyniosła 0,4353405089. Stwierdzono wystąpienie nieznacznych różnic w częstościach poszczególnych alleli w badanej populacji w porównaniu z populacją japońską [5] i bardziej znaczących, zwłaszcza w układzie rs2277216, w porównaniu z inną opublikowaną populacją japońską [3] (tabela III).

Zestaw pięciu SNP może mieć zastosowanie w analizie porównawczej dowodów rzeczowych, zwłaszcza gdy mamy do czynienia ze zdegradowanym materiałem biologicznym i badanie z użyciem markerów STR, czy nawet bardziej skutecznych w tego typu analizach markerów miniSTR, nie daje jednoznacznych wyników. Do badania pokrewieństwa zestaw pięciu markerów jest niewystarczający i należałoby go poszerzyć, może być jednak dobrym narzędziem pomocniczym w analizach spraw nietypowych, takich jak przypadki mutacyjne bądź niepełne.

Tabela II. Rozkład częstości alleli oraz parametry statystyczne badanych 5 loci SNP.

Table II. The allele frequencies and statistical parameters of five SNP loci.

SNP	Allel Allele	Częstość Frequency	p	PD	Het			PIC	PE		PI		
					exp	obs	błąd		trio	duo	min	średnia	max
rs2294067	C	0,488	0,498	0,625	0,503	0,55	0,04	0,375	0,187	0,125	0,976	1,006	2,051
	G	0,512											
rs2070764	T	0,625	0,343	0,608	0,472	0,525	0,04	0,359	0,179	0,110	0,800	0,946	2,667
	A	0,375											
rs1063739	A	0,488	0,653	0,625	0,503	0,475	0,04	0,375	0,187	0,125	0,976	1,006	2,051
	C	0,512											
rs2282160	G	0,513	0,178	0,625	0,503	0,425	0,04	0,375	0,187	0,125	0,976	1,006	2,051
	A	0,487											
rs2277216	C	0,794	0,084	0,494	0,329	0,263	0,037	0,274	0,137	0,054	0,630	0,746	4,848
	T	0,206											

Tabela III. Porównanie częstości alleli pięciu loci SNP w populacji centralnej Polski i dwóch populacji japońskich [3, 5].

Table III. Comparison of allele frequencies of five SNP loci between central Poland population and two Japanese populations [3, 5].

SNP	Allele Allele	Częstość Frequency	Częstość wg [3] Frequency [3]	Częstość wg [5] Frequency [5]
rs2294067	G	0,512	0,637	0,556
	C	0,488	0,363	0,444
rs2070764	A	0,375	0,446	0,453
	T	0,625	0,554	0,547
rs1063739	A	0,488	0,392	-
	C	0,512	0,608	-
rs2282160	A	0,487	0,544	0,493
	G	0,513	0,456	0,507
rs2277216	T	0,206	0,520	0,483
	C	0,794	0,480	0,517

PIŚMIENNICTWO

1. Weiner M. P., Hudson T. J.: Introduction to SNPs: Discovery of Markers for Disease. *Biotechniques* 2002, 32, S4-S13.

2. Lee H. Y., Park M. J., Yoo J. E., Chung U., Han G. R., Shin K. J.: Selection of twenty-four highly informative SNP markers for human identification and paternity analysis in Koreans. *Forensic Sci. Int.* 2005, 10, 148, 107-12.

3. Inagaki S., Yamamoto Y., Doi Y., Takata T., Ishikawa T., Imabayashi K., Yoshitome K., Miyaishi S., Ishizu H.: A new 39-plex analysis method for SNPs including 15 blood group loci. *Forensic Sci. Int.* 2004, 144, 45-57.

4. Doi Y., Yamamoto Y., Inagaki S., Shigeta Y., Miyaishi S., Ishizu H.: A new method for ABO genotyping using a multiplex single-base primer extension reaction and its application to forensic case-work samples. *Int. Legal Med* 2004, 6, 213-223.

5. <http://snp.ims.u-tokyo.ac.jp/>

6. Brenner C.: The power of SNPs – even without population data. Poster presentation at the 10th Promega Symposium on Human Identification, 1999, <http://dna-view.com>

7. Sohocki M. M., Sullivan L. S., Harrison W. R., Sodergren E. J., Elder F. F., Weinstock G., Tanase S., Daiger S.: Human glutamate pyruvate transaminase (GPT): localization 8q24.3, cDNA and genomic sequences, and polymorphic sites, *Genomics* 1997, 40, 247-252.

8. <http://www.ncbi.nlm.nih.gov/SNP/>

9. Lewis P. O., Zaykin D.: Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c), 2001, <http://lewis.eeb.uconn.edu/lewishome>

10. Berent J., Szram S.: DLP – a computer program for calculation of discrete locus parameters. *Forensic Sci.* 2003, 1, 46-47.

11. Nei M., Roychoudhury A. K.: Sampling variants of heterozygosity and genetic distance. *Genetics* 1974, 76, 379-390.

12. Botstein D., White R. L., Skolnick M., Davis R. W.: Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.* 1980, 32, 314-331.

13. Berent J. A., Miścicka-Śliwka D., Czarny J.: Średnie wartości szansy ojcostwa – obliczenia dla populacji polskiej. *Arch. Med. Sąd. i Kryminol.* 1999, 49, 11-15.

14. Brenner C., Morris J. W.: Paternity index calculations in single locus hypervariable DNA probes: validation and other studies. *Proceedings from the International Symposium on Human Identification* 1989, Promega, Madison 1990, pp.21-53.

15. National Research Council Report II. The Evaluation of Forensic DNA Evidence. National Academy Press, Washington, D.C. 1996, pp.96-97.

16. Weir B. S.: Genetic Data Analysis II. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts 1996, pp.209-211.

17. Fung W. K., Chung Y. K., Wong D. M.: Power of exclusion revisited: probability of excluding relatives of the true father from paternity. *Int. J. Legal Med.* 2002, 116, 64-67.

Adres pierwszego autora:
Katedra i Zakład Medycyny Sądowej
Uniwersytetu Medycznego w Łodzi
ul. Sędziowska 18a
91-304 Łódź
katarzynabol@wp.pl