

1989, 7, 736. -2. Budowle B., Moretti TR., Baumstark AL., Defenbaugh DA., Keys KM.. Population data on the thirteen CODIS core short tandem repeat loci in African Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. *J. Forensic Sci.* 1999 6, 1277-86 -3. Chang-En Pu, Ching-Mei Hsieh, Meng-Yi Chen, Fang-Chin Wu, Chien-Feng Sun. Genetic variation at nine STR loci in populations from the Philippines and Thailand living in Taiwan. *Forensic Science International* 1999, 106, 1-6. -4. Entrala C, Lorente J., Lorente M., Alvarez J., Budowle B., and Villanueva E. Population Studies and Casework Application with the New GenePrint™ Silver STR III Multiplex (D16S539, D7S820, D13S317). *J. Forensic Sci.* 1999 5, 1032-4. -5. Garofano L., Pizzamiglio M., Vecchio C, Lago G., Floris T., D'Errico G., Brembilla G., Romano A., Budowle B. Italian population data on thirteen short tandem repeat loci: HUMTH01, D21S11, D18S51, HUMVWFA31, HUMFIBRA, D8S1179, HUMTPOX, HUMCSF1PO, D16S539, D7S820, D13S317, D5S818, D3S1358, *Forensic Sci. Int.* 1998, 97, 53-60. -6. Lahiri D.K., Bye S., Nurnberger J.L Hodes M.E., Crisp M., A non-enzymatic and non-enzymatic extraction method gives higher yields of genomic DNA from whole-blood samples than do nine other methods tested. *Journal of Biochemical and Biophysical Methods* 1992, 25, 193-205. -7. Walsh P.S., Metzger D.A. and Higuchi R. Chelex®100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 1991, 10, 506-513. -8. Wiegand P., Schurenkamp M., Schuttler U. DNA extraction from mixtures of body fluid using mild preferential lysis. *Int. J. Leg. Medicine* 1992, 104, 359-360.

Adres pierwszego autora:

Katedra i Zakład Medycyny Sądowej
80-210 Gdańsk
ul. Curie Skłodowskiej 3A

**Ryszard Pawłowski, Anita Dettlaff-Kąkol, Grzegorz Jezierski,
Agnieszka Maciejewska, Regina Paszkowska, Monika Reichert**

Genetyka populacyjna dziewięciu loci typu STR z zestawu Profiler Plus w próbce populacyjnej z obszaru Polski

Population genetics of nine STR loci from the Profiler Plus kit in a population sample from Poland

Z Katedry i Zakładu Medycyny Sądowej AM w Gdańsku,
p.o. Kierownik: dr hab. Z. Szczerkowska - profesor AM

Praca przedstawia wyniki badań populacyjnych 9 loci typu STR (D3S1358, VWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317 i D7S820) zawartych w komercyjnym zestawie Profiler Plus. Próbkę DNA pochodzącą od 706 niespokrewnionych osobników poddano amplifikacji metodą multiplex PCR a następnie produkty zidentyfikowano z zastosowaniem elektroforezy kapilarnej na automatycznym sekwenatorze DNA (ABI Prism 310). W badanej próbce populacyjnej wszystkie badane loci znajdowały się w równowadze opisanej równaniem Hardy-Weinberga. Obliczona wartość prawdopodobieństwa przypadkowej zgodności wynosi 1.25×10^{-11} dając tym samym średnie prawdopodobieństwo - przypadkowej zgodności profili DNA w populacji polskiej wynoszące 1 do 83 miliardów. Zestaw Profiler Plus dzięki bardzo wysokiej sile dyskryminacyjnej jest szczególnie przydatny do badań identyfikacyjnych z zakresu genetyki sądowej.

This paper presents the allele frequency distributions for the nine (D3S1358, VWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317 and D7S820) loci present in the commercially available Profiler Plus kit. DNA samples of 706 individuals from Northern Poland were amplified in a multiplex reaction with subsequent automatic detection using capillary electrophoresis (ABI Prism 310 DNA sequencer). In the analysed population sample all loci met the Hardy-Weinberg equilibrium conditions. The calculated probability of identity is 1.25×10^{-11} giving on average 1 in 83 billion probability of identity in our population sample. Owing to high discrimination power the Profiler Plus kit is highly useful for forensic identification purposes.

Słowa kluczowe: Multiplex PCR, loci typu STR, genetyka populacyjna, elektroforeza kapilarna, Polska.

Key words: Multiplex PCR, STR loci, population genetics, capillary electrophoresis, Poland.

WSTĘP

Sekwencje typu N/NTR-STR są powszechnie stosowane do badań z zakresu genetyki sądowej (2). Opracowanie i zastosowanie techniki multiplex PCR stworzyło możliwość jednoczesnej i szybkiej amplifikacji wielu loci z minimalnej ilości DNA izolowanego np. z plam biologicznych (3,5,6). Połączenie techniki multiplex PCR i fluorescencyjnej detekcji produktów PCR otworzyło nowe możliwości automatycznej analizy wielu polimorficznych loci jednocześnie. Jednym z komercyjnie dostępnych zestawów typu multiplex jest zestaw ProfilerPlus firmy Perkin-Elmer (PE), zawierający 9 loci typu STR (D3S1358, VWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317 i D7S820) i marker płci w postaci locus *amelogeniny*. Warunkiem wprowadzenia nowego zestawu (lub nowego locus) do praktyki genetyki sądowej jest nie tylko poznanie wpływu różnorodnych czynników na czułość, wiarygodność i powtarzalność analizy (12), ale również poznanie częstości alleli obecnych w analizowanych loci. Głównym celem niniejszej pracy było zbadanie rozkładu częstości alleli dziewięciu loci typu STR zawartych w zestawie Profiler Plus w próbkę populacyjnej liczącej 706 osobników pochodzących z obszaru Polski oraz porównanie ich rozkładu z zaobserwowanymi w dwóch innych populacjach europejskich oraz populacji Polski północnej.

MATERIAŁY I METODY

Materiał do badań populacyjnych stanowiły próbki krwi lub wymazy z jamy ustnej pobierane jako materiał porównawczy do badań identyfikacyjnych śladów biologicznych pochodzące od osób z różnych regionów Polski. Na materiał ten złożyły się próbki niespokrewnionych 166 kobiet i 540 mężczyzn. Ekstrakcję DNA przeprowadzano metodą opisaną wcześniej (13) lub metodą z Chelex-100 (17). Ilość DNA oznaczano metodą hybrydyzacji z sondą swoistą dla DNA naczelnych (QuantiBlot, PE, USA) z chemiluminescencyjną metodą detekcji lub pomiarem fluorymetrycznym.

Amplifikację loci zawartych w zestawie ProfilerPlus prowadzono w warunkach opisanych przez producenta stosując termocykler 2400 lub robot do reakcji PCR 877 firmy PE. Warunki przygotowania próbek do elektroforezy kapilarnej i sam ich rozdział prowadzono w warunkach jak opisano uprzednio (10,11). Jako standardu wewnętrznego wielkości DNA używano GS500, znakowanego ROX (PE) lub ILS600 znakowanego CXR (Promega).

Sekwencjonowanie produktów PCR (allel 16 FGA i allel 33.1 locus D21S11) przeprowadzono z zastosowaniem odpowiednich starterów (1,14). Matryce do reakcji sekwencjonowania przygotowywano wg. metody opisanej przez Walsha i wsp. (16). Reakcje sekwencjonowania prowadzono z zastosowaniem zestawu firmy PE (BigDye Terminator sequencing kit) w warunkach opisanych wcześniej (17).

Analiza statystyczna. Weryfikację zgodności z równaniem Hardy'ego-Wein-

berga prowadzono testami exact i LR (7,18). Obserwowaną heterozygotyczność (H), siłę dyskryminacji (PD), średnią szansę wykluczenia (MEC) i zawartość informacji polimorficznej (PIC) obliczano jak uprzednio (10).

WYNIKI I DYSKUSJA

Łącznie przy zastosowaniu zestawu Profiler Plus amplifikacji poddano 1048 próbek DNA pochodzących z materiału porównawczego. Z tej ilości, po odrzuceniu próbek pochodzących od osób spokrewnionych, analizie statystycznej poddano 706 profili DNA. Ze wszystkich przebadanych próbek, w zaledwie kilku przypadkach, konieczne było powtórne zamplifikowanie lub nastryknięcie produktów na kapilarę. Zaobserwowano, że dla trzech loci znakowanych NED (D5S818, D13S317 i D7S820) uzyskuje się najniższy sygnał amplifikacyjny w porównaniu do pozostałych znakowanych JOE i 5-FAM. Najwyższe sygnały amplifikacyjne obserwowano dla loci D3S1358, VWA i FGA znakowanych 5-FAM. Zaobserwowana sytuacja nie przynosi problemów interpretacyjnych w przypadku badania materiału porównawczego, jednak podczas analizy zdegradowanego materiału biologicznego, czy próbek będących mieszaninami DNA często dochodzi do utraty sygnałów amplifikacyjnych dla „żółtych loci” stwarzając tym samym trudności natury interpretacyjnej.

Tabela 1 przedstawia zaobserwowane częstości alleli dla 9 loci zawartych w zestawie Profiler Plus w populacji polskiej zaobserwowane w próbkę populacyjnej 706 niespokrewnionych osób obu płci z obszaru Polski. Dla poszczególnych loci stwierdzono od 8 (D13S317) do 17 (FGA) alleli. Jednocześnie zidentyfikowano trzy rzadkie allele FGA*16, D21S11*33.1 oraz D3S1158*20. Obecność dwóch pierwszych potwierdzono sekwencjonowaniem, w przypadku trzeciego (D3S1158) wykrytego niedawno, nie potwierdzono jeszcze sekwencjonowaniem jego pewnej obecności. Allel ten wykazuje jednak dokładnie o 4pz większą długość niż allel 19 obecny w drabinie alleli tego locus. W wyniku sekwencjonowania 16 allela FGA stwierdzono następującą sekwencję w rejonie repetytywnym: (TTTC)₂ TTTT TTCT (CTTT)₂ CTCC (TTCC)₂, co odpowiada dokładnie sekwencji 16 allela wg. nomenklatury i danych przytoczonych przez Barbera i wsp. (1). W przypadku allela 33.1 z locus D21S11 (nie opisanego dotychczas w publikacjach) stwierdzono następującą sekwencję obszaru repetytywnego: FR- (TCTA)₂ (TCTG)₂ - CR - (TCTA)₂ - A-TCTA-FR, gdzie FR oznacza obszar flankujący, a CR obszar stały o długości 43 nukleotydów i sekwencji (TCTA)₂TA(TCTA)₂TCA(TCTA)₂TCCATA. Allel ten zgodnie z nomenklaturą International Society of Forensic Genetics nazwano 33.1.

W wyniku zastosowania odpowiednich testów (7, 18), stwierdzono że wszystkie badane loci znajdują się w równowadze opisanej równaniem Hardy'ego-Weinberga (tabela 1). Najbardziej polimorficzne podobnie jak dla próbki 202 osób z populacji Polski północnej (12), okazały się loci: D18S51, HUMFGA i D21S11, a najmniej D5S818. Podobne wyniki uzyskano podczas badań populacji austriackiej i hiszpańskiej (4).

Tabela I. Częstości alleli i współczynniki statystyczne charakteryzujące loci ProfilerPlus w populacji polskiej (n=706).

Table I. Allele frequencies and some statistical indices for Profiler Plus loci in a polish population (n=706).

Allele	D3S1358	VWA	FGA	D8S1179	D21S11	D18S51	D5S818	D13S317	D7S820
6									0.0007
7							0.0028		0.0106
8				0.0099			0.0007	0.1431	0.1353
9				0.0113		0.0007	0.0496	0.0878	0.1544
10				0.0616		0.0099	0.0765	0.0510	0.3116
11		0.0007		0.0751		0.0099	0.3123	0.3584	0.1962
12	0.0021	0.0007		0.1622		0.0942	0.3916	0.2436	0.1629
13	0.0021	0.0028		0.3222		0.0977	0.1565	0.0758	0.0248
14	0.1367	0.1034		0.2408		0.1487	0.0092	0.0390	0.0028
15	0.2380	0.1105		0.0963		0.1593	0.0007	0.0014	0.0007
16	0.2542	0.1898	0.0014	0.0177		0.1728			
17	0.2210	0.2826	0.0007	0.0021		0.1289			
18	0.1374	0.2224	0.0142	0.0007		0.0836			
19	0.0078	0.0751	0.0871			0.0460			
20	0.0007	0.0092	0.1459			0.0290			
20.2			0.0007						
21		0.0028	0.1905			0.0050			
21.2			0.0021						
22			0.1941			0.0120			
22.2			0.0142						
23			0.1218			0.0021			
23.2			0.0071						
24			0.1261						
25			0.0673						
25.2			0.0014						
26			0.0241			0.0028			
27						0.0212			
28			0.0014			0.1742			
29						0.2047			
29.2						0.0007			
30						0.2280			
30.2						0.0510			
31						0.0751			
31.2						0.0899			
32						0.0170			

Ciąg dalszy tabeli I

32.2						0.0963			
33						0.0021			
33.1.2						0.0007			
33.2						C.0099			
34.2						0.0035			
PD	0.923	0.936	0.964	0.926	0.958	0.970	0.873	0.918	0.927
PIC	0.760	0.779	0.845	0.765	0.831	0.865	0.669	0.744	0.766
MEC	0.586	0.618	0.720	0.602	0.699	0.752	0.478	0.576	0.599
test LR	0.532	0.503	0.964	0.788	0.980	0.656	0.596	0.974	0.283
Test Exact	0.425	0.447	0.886	0.591	0.977	0.207	0.453	0.898	0.139
Hobs.	0.785	0.785	0.853	0.798	0.860	0.875	0.711	0.758	0.785

Porównując wyniki analiz statystycznych dla obu próbek populacyjnych 202 osób z obszaru Polski północnej i 706 osobników z obszaru całej Polski można stwierdzić, że są one zbliżone. Największy wpływ powiększenia próbki populacyjnej zaobserwowano dla locus D13S317, gdzie stwierdzono największy wzrost wartości PD, PIC i MEC. Powiększenie próbki populacyjnej spowodowało też, zgodnie z oczekiwaniami wzrost ilości obserwowanych alleli. Łączna ilość zaobserwowana dla populacji 202 osób wyniosła 91, a dla 706 osób - 105. Uzyskane częstości alleli są zbliżone do tych, które zaobserwowano dla próbki populacyjnej z obszaru Polski północnej (12). Wyniki te świadczą o dużej homogenności rozkładu alleli dla badanych loci w populacji polskiej.

Obliczone prawdopodobieństwo przypadkowej zgodności profili DNA dla populacji 706 osób (1.26×10^{-11}) jest wyraźnie wyższe niż dla 202 osób (2.26×10^{-11}). Oznacza to średnie prawdopodobieństwo identyczności wynoszące odpowiednio 1 na 83 miliardy i 1 na 44 miliardy osób. Jest to wartość zbliżona do tych, które uzyskano dla próbek populacyjnych w Austrii i Hiszpanii wynoszące odpowiednio 1 na 96 i 109 miliardów osób (4,9).

Tabela II przedstawia porównanie homogenności rozkładu alleli loci Profiler Plus pomiędzy populacją polską a austriacką i hiszpańską. Porównanie homogenności rozkładu częstości alleli pomiędzy populacją polską a austriacką wykazało istotne statystycznie różnice dla loci: FGA, D18S51 i D13S317, natomiast porównanie z populacją z południowej Hiszpanii wykazało brak istotnych różnic tylko dla 2 loci (D7S820 i VWA).

Otrzymane współczynniki statystyczne charakteryzujące poszczególne loci zawarte w zestawie ProfilerPlus oraz ich łączna wartość świadczy o jego wysokiej przydatności do badań identyfikacyjnych w genetyce sądowej.

Tabela II. Porównanie homogenności rozkładu alleli loci ProfilerPlus pomiędzy populacjami polską austriacką i hiszpańską.
Table II. Comparison of homogeneity distribution of Profiler Plus alleles between Polish Austrian and Spanish populations.

Polska x Hiszpania Południowa Poland x Southern Spain				
	χ^2	P	G-statystyka G-statistics	P
D3S1358	20,59	0,017	18,92	0,020
VWA	13,67	0,184	14,95	0,133
FGA	41,83	0,000	37,84	0,001
D8S1179	25,84	0,008	27,71	0,007
D21S11	34,22	0,007	31,18	0,017
D18S51	26,38	0,024	31,38	0,008
D5S818	20,93	0,010	19,41	0,014
D13S317	87,76	0,000	73,73	0,000
D7S820	7,74	0,556	8,12	0,551
Polska x Austria Poland x Austria				
	f	P	G-statystyka G-statistics	P
D3S1358	15,73	0,063	14,35	0,100
VWA	13,25	0,210	13,51	0,236
FGA	41,55	0,002	38,12	0,006
D8S1179	12,39	0,256	11,41	0,337
D21S11	8,19	0,882	9,06	0,010
D18S51	40,17	0,000	41,20	0,000
D5S818	1,22	0,994	1,64	0,995
D13S317	26,11	0,000	25,70	0,000
D7S820	10,37	0,311	10,48	0,326

1998, 98, 179-183. - 5. Furedi S., Budowle B., Wolier J., Padar Z.: Hungarian population data on six STR loci - HUMVWFA31, HUMTH01, HUMCSF1PO, HUMFES/FPS, HUMTPOX, and HUMHPRTB - derived using multiplex PCR amplification and manual typing. Int J Leg Med 1996, 109, 100-101. - 6. Greenhalgti M.J.: Casework experiences with a multiplex STR system. Adv. Forensic Haemogenetics 1996, 6, 249-251. - 7. Guo S.W, Thompson E.A.: Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 1992, 48, 361-372. - 8. Mills K.A., Even D., Murray J.C.: Tetranucleotide repeat polymorphism at the human alpha fibrinogene locus (FGA). Hum Mol Genet 1992, 1, 79-82. - 9. Nuehuber F., Radacher M., Meisner N., Tutsch-Bauer E.: Nine STR markers plus amelogenin (AmpFiSTR Profiler Plus): a forensic study in an Austrian population. Int J Leg Med 1999, 113, 60-62. - 10. Pawłowski R.: HUMFIBRA allele distribution in Northern Poland using capillary electrophoresis. Int J Leg Med 1999, 112, 139-141.

11. Pawłowski R., Branicki W., Kupiec T.: Y-chromosomal polymorphic loci DYS19, DYS390, DYS393 in a population sample from northern Poland. Electrophoresis 1999, 20, 1702-1706. - 12. Pawłowski R., Maciejewska A.: The forensic validation of multiplex containing nine STRs - population genetics in Northern Poland. Int J Leg Med 2000 (w druku) - 13. Pawłowski R., Maciejewska A., Paszkowska R., Welz A.: Frequencies of the five short tandem repeat systems (STR) in a population from North Poland. Int J Leg Med 1997, 110, 10-13. - 14. Sharma V., Litt M.: Tetranucleotide repeat polymorphism at the D21S11 locus. Hum Mol Genet 1992, 1(1), 67. - 15. Wallin J.M., Buoncristiani M.R., Lazaruk K.D., Fildes N., Holt C.L., Walsh P.S.: TWGDAM validation of the AmpFiSTR blue PCR amplification kit for forensic casework analysis. J Forensic Sci. 1998, 43, 854-870. - 16. Walsh P.S., Fildes N.J., Reynolds R.: Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acids Res 1996, 24, 2807-2812. - 17. Walsh P.S., Metzger D.A., Higuchi R.: Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques. 1991, 10(4), 506-13. - 18. Weir B.S.: Independence of VNTR alleles defined by fixed bins. Genetics 1992, 130, 873-878.

PIŚMIENNICTWO

1. Barber M.D., McKeown B.J., Parkin B.H.: Structural variation in the alleles of a short tandem repeat system at the human alpha fibrinogen locus. Int J Leg Med. 1996, 108, 180-185 - 2. Brinkmann B.: The STR approach. Adv Forensic Haemogenetics, 1996, 6, 41-51. - 3. Budowle B., Smerick J.B., Keys K.M., Moretti T.R., United States population data on the multiplex short tandem repeat loci - HUMTH01, TPOX, and CSF1 PO - and the variable number tandem repeat locus D1S80. J Forensic Sci 1997, 42, 846-849. - 4. Entrala C, Lorente M., Lorente J.A., Alvares J.C., Moretti T., Budowle B. Villanueva E.: Fluorescent multiplex analysis of nine STR loci: Spanish population data. Forensic Sci Int

Adres pierwszego autora:
Katedra i Zakład Medycyny Sądowej
80-210 Gdańsk
ul. Curie Skłodowskiej 3A