

## POLYMORPHISM OF STR SYSTEMS (D9S304, D1S533, D7S820, FGA) IN POPULATION SAMPLES OF SOUTH-EAST POLAND

Piotr KOZIOŁ, Sylwia CHOCHOLSKA, Roman MAĐRO

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

**ABSTRACT:** Four STR systems: D9S304, D1S533, D7S820, FGA were analysed on 301 DNA samples. Amplification was carried out using MultiPlex II kits (Lifecodes). PCR products were separated using polyacrylamide gel electrophoresis and the silver staining method. In the D9S304 locus, 10 alleles were identified (the most frequent of which was allele 8 – 0.377), in the D1S533 locus 10 alleles (the most commonly occurring one was allele 13 – 0.307), 8 alleles in the D7S820 locus (the most frequent was allele 10 – 0.319), and 8 alleles in the FGA locus (with the most frequent allele 22 – 0.219) were observed. The calculated heterozygosity of analysed STR systems was 0.791, 0.804, 0.791 and 0.860 respectively. The distribution of genotype frequencies shows no deviation from the Hardy-Weinberg equilibrium. The high usefulness of the MultiPlex II kit in putative paternity testing was confirmed by the value of the Power of Discrimination coefficient ( $PD = 99.99$ ), Paternity Index ( $PI = 51$ ) and analysis of genotyping results in 87 paternity cases.

**KEY WORDS:** STRs; Population database; Linkage equilibrium; Forensic genetics; Paternity testing.

*Z Zagadnień Nauk Sądowych, z. XLIX, 2002, 84–98*

*Received 5 November 2001; accepted 10 January 2002*

### INTRODUCTION

The STR type of DNA polymorphism has been applied widely in forensic genetics since the development of the technique of amplifying many STR loci in one PCR reaction and genotyping and detection in one electrophoretic process [3, 4, 6]. However, simultaneous genotyping of a dozen or more STR systems requires the use of expensive equipment for automatic analysis of allelic fragments labelled with fluorochromes [11].

A cheap alternative technique of PCR detection is silver staining, application of which allows differentiation of from 3 to 4 loci in one electrophoretic separation. Carrying out a few multiplex reactions enables an amount of information to be gained comparable with that obtained by the fluorescent technique. An example, presented in the literature, may be the use (for amplification of a few loci and silver detection) of commercial kits produced by

Promega enabling amplification of 9 STR systems (CSF1P0, TPOX, TH01, F13A01, FESFPS, vWA, D16S539, D7S820, D13S317) in three separate PCR reactions [12, 17]. Similarly, the MultiPlex I kit (D18S849, D3S1744 and D12S1090) and MultiPlex II kit (D9S304, D1S533, D7S820, FGA) manufactured by Lifecodes allow genotyping of 7 STR systems. However, up to the present day these kits have not been tested on the Polish population.

The aim of the present study was to determine the genotypes of 4 DNA loci (D9S304, D1S533, D7S820 i FGA) in a population sample comprising 301 unrelated people from south-east Poland, determine frequencies of alleles, perform statistical calculation concerning allele segregation, compare results with other published population data and evaluate the usefulness of the multiplex in paternity cases.

#### MATERIAL AND METHODS

The study was performed on 301 DNA samples extracted from blood taken at the Department of Forensic Medicine, the Medical Academy in Lublin from unrelated donors of both genders living in south-east Poland and involved in disputed paternity cases. Analysis was also performed on DNA samples taken from children involved in 87 (disputed paternity) cases, in which routine analysis of 4 VNTR systems (D4S139, D5S110, D7S21, and D11S12) and 9 STR systems (PowerPlex 2.1) confirmed paternity with a probability close to certainty in 46 cases, and excluded the alleged father in 41 cases.

DNA was extracted with the standard phenol-chloroform extraction method. The concentration of the extracted DNA was measured using the slot blot method with a D17Z1 probe (Gibco-BRL, Gaththersburg, MD, USA). The MultiPlex II kit (Lifecodes Corporation) was used for simultaneous amplification of 4 STR systems (D9S304, D1S533, D7S820 i FGA) [18]. Each PCR reaction was conducted in 20  $\mu$ l in a Trio-block thermocycler (Biometra) according to the manufacturer's instructions. PCR products were separated on denaturing 6% polyacrylamide gels (4% T, 6% C, 8 M urea) in 0.5  $\times$  TBE buffer, pH = 8.3 at a constant power of 60 W for about 1.5 hours. DNA fragments were next silver stained [1], and genotypes were determined by comparing to an allelic ladder. The correctness of genotyping of FGA and D7S820 systems by this method was checked by comparison to results obtained for 170 persons analysed with the fluorescent method using the PowerPlex 2.1 kit delivered by Promega corporation and the FMBIO II Hitachi [10]. The size of allele 16 of the D1S533 locus was calculated using a local correlation method with video camera and BioImage image analysis software by Biotek Fischer. Markers from the MultiPlex II ladder were used

in the range spanning between fragment 203 bp – allele 15 for the D1S533 locus and 217 bp – allele 7 for locus D7S820.

Allele frequencies determined for studied populations were compared (using 2 way RXC software according to G. Carmody, Carleton University, Ottawa, Canada) with frequencies ascertained in other Polish populations and a population of US Caucasians. In order to test concordance of an allele's segregation with the Hardy-Weinberg equilibrium, statistical evaluation of results with  $\chi^2$  [15], "exact" [7] and "LR" [19] tests were performed. Coefficients defining the usefulness of these markers for forensic sciences were also calculated [8]. Mean theoretical values of paternity index *PI* calculated according to Brenner et al.'s [2] formula were compared with values of *PI* calculated in analysed paternity cases. The *PI* values were calculated by three different methods: as a geometric mean and median from the *PI* set, and mean probability as an arithmetic mean of the set of probability of paternity (PDP), which was subsequently converted to a *PI* value.

## RESULTS AND DISCUSSION

The application of the MultiPlex II kit by Lifecodes, and the technique of allelic fragment separation and silver detection allowed determination of genotypes of 4 loci in all analysed samples. In the present study, 4% denaturing gels recommended by the manufacturer were replaced by 6% denaturing gels. The separation of products in the denaturing conditions of electrophoresis causes the appearance of visible differences in migration of complementary DNA strands. In connection with this in the case of electrophoresis on 4% gels in FGA and D1S533 systems, complementary strands of DNA of one allele were weakly separated from fragments of the larger allele. This caused problems in identification of genotype (Figure 1a), even after prolonged electrophoresis separation. Significantly more readable electrophoresis separations were achieved by application of 6% denaturing gels; separation of complementary strands of FGA and D1S533 systems is so close that they almost form one fragment, while fragments in the allelic ladder are sufficiently separated from each other. The good readability of separation of remaining systems of multiplex D7S820 and D9S304 using these gels does not cause problems during genotype determination (Figure 1b).

Visually identified alleles showed migration velocity identical to alleles in the allelic ladder, with two exceptions. The first of these was allele 16 in the D1S533 system, which did not have a counterpart in the allelic ladder. In each of 9 analysed cases, the length of this allele amounted to 207 ( $\pm 0.6$  bp) and was one tetranucleotide repeat motif longer than allele 15 (position 9 and 14 on Figure 1b). The second inconsistency with the allelic ladder re-

lates to the FGA system, whose allelic ladder includes only 8 alleles with an even number of tetranucleotide repeats. Alleles with an odd number of repeat motifs lay between alleles of this ladder. Therefore there is no possibility of identifying inter-alleles in the FGA system. In some persons, inter-alleles 22.2, 23.2 and 24.2 were earlier determined in the FGA system (using the PowerPlex 2.1 kit and fluorescent detection) – these were read as alleles 23, 24 and 25 respectively by the silver technique.

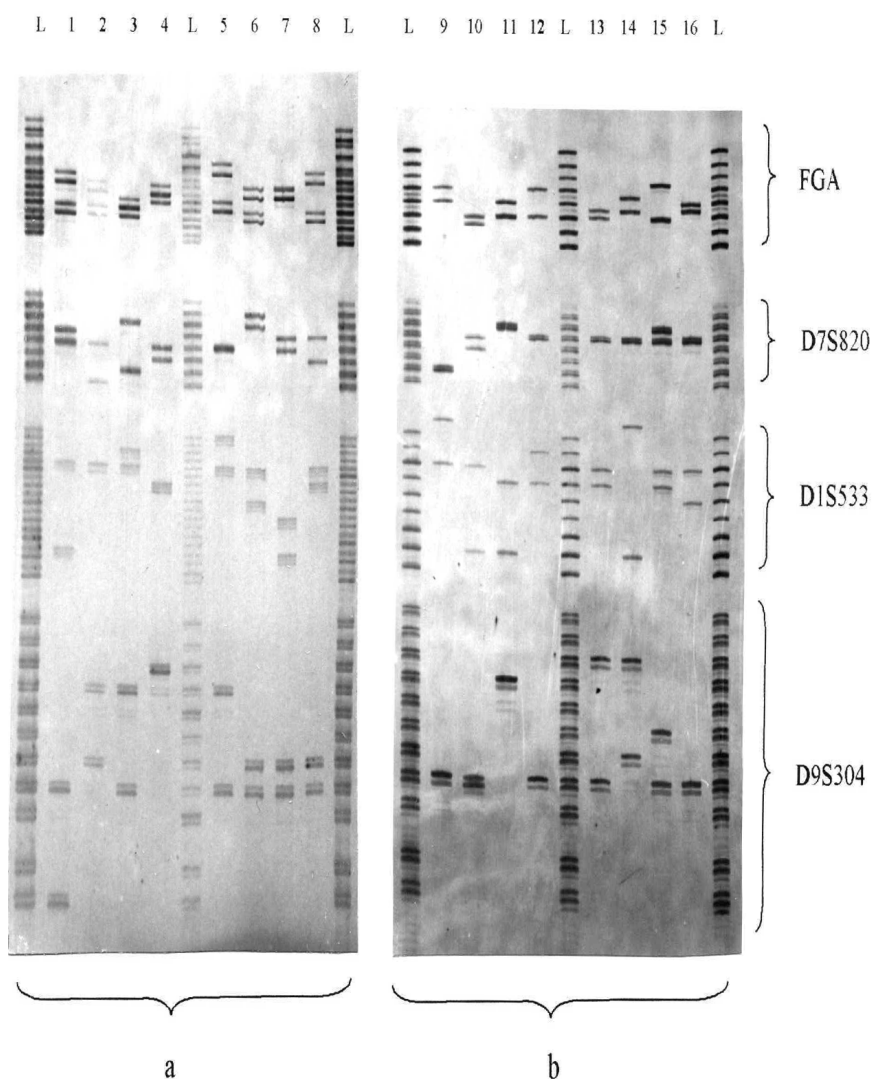


Fig. 1. The separation of multiplex PCR products on a 4% (a) and 6% (b) denaturing PAA gel. Lanes 1–16: allelic fragments in the nine DNA samples; lanes L – allelic ladder for four loci.

TABLE I. DISTRIBUTION OF GENOTYPE AND ALLELE FREQUENCIES OF 4 STR SYSTEMS (FGA, D7S820, D1S533, D9S304) IN A POPULATION SAMPLE FROM SOUTH-EAST POLAND

FGA, N = 602											
Allele	18	19	20	21	22	23	24	25	26		Frequency
18	0										0.010
19	0	1									0.095
20	1	15	4								0.150
21	1	6	16	10							0.169
22	2	12	21	26	13						0.219
23	1	11	12	9	18	4					0.126
24	1	4	12	12	21	11	7				0.140
25	0	7	4	11	6	6	7	3			0.081
26	0	0	1	1	0	0	2	2	0		0.010
D7S820, N = 602											
Allele	7	8	9	10	11	12	13	14			Frequency
7	0										0.005
8	0	10									0.171
9	1	12	5								0.138
10	1	41	23	32							0.319
11	0	19	23	36	9						0.201
12	1	11	8	23	22	6					0.134
13	0	0	6	4	3	3	1				0.030
14	0	0	0	0	0	1	0	0			0.002
D1S533, N = 602											
Allele	7	8	9	10	11	12	13	14	15	16	Frequency
7	0										0.002
8	0	4									0.120
9	0	5	1								0.081
10	0	2	2	1							0.025
11	0	1	2	2	4						0.068
12	1	13	11	1	7	11					0.161
13	0	23	14	5	14	18	31				0.307
14	0	16	8	1	5	18	30	7			0.164
15	0	3	4	0	2	4	16	5	0		0.057
16	0	1	1	0	0	2	3	2	0	0	0.015
D9S304, N = 602											
Allele	4	7	8	9	10	11	12	13	14	15	Frequency
4	1										0.103
7	2	0									0.017
8	28	3	40								0.377
9	2	1	29	6							0.109
10	2	1	13	2	3						0.055
11	4	0	16	4	3	2					0.071
12	11	1	31	10	5	0	7				0.135
13	8	0	20	6	1	9	9	4			0.105
14	2	2	4	0	0	2	0	2	0		0.020
15	1	0	3	0	0	1	0	0	0	0	0.008

Moreover it was found that the amplification reaction using MultiPlex II requires a considerable quantity of DNA template (4 ng), and that a small decrease in amount of DNA template causes weaker amplification of D1S533 fragments.

Genotyping results and genome frequencies are presented in Table I. The results reveal that in the FGA locus 37 genotypes created by 9 alleles were identified with the most frequent being allele 22 (0.219). In locus D7S820, 8 alleles (24 genotypes) were found, of which the most common alleles were 10 (0.319) and 11 (0.201), while two alleles, 7 and 14, have frequencies below 0.5%. In the analysed population the presence of 10 alleles and 40 genotypes of the D1S533 and D9S304 loci were observed. In the D1S533 system the most frequently occurring alleles were 18 (0.307) and 19 (0.2126) and only one allele, 7, had a frequency below 0.5%. The most frequent in the D9S304 system was allele 8 (0.377) and the least, allele 5 (0.008).

Data presented in Table II indicate that the allele distribution of the analysed STR loci in the population sample from south-east Poland shows no deviation from the Hardy-Weinberg equilibrium, and there is no statistically important divergence between the observed and expected heterozygosity. It should be emphasised that the “exact” test and the “LR” test performed directly on results of analysis of 301 persons for the D9S304 system gave a low probability value ( $P = 0.06$  and  $P = 0.05$ ), while calculation using Monte Carlo methods with 10000 permutation [7] gave a high probability value ( $P = 0.992$  and  $P = 0.984$ ).

TABLE II. RESULTS OF TESTING HARDY-WEINBERG EQUILIBRIUM FOR 4 STR SYSTEMS IN THE POPULATION OF SOUTH-EAST POLAND

Test		FGA ( $N = 602$ )	D7S820 ( $N = 602$ )	D1S533 ( $N = 602$ )	D9S304 ( $N = 602$ )
Heterozygosity	Observed	0.860	0.791	0.804	0.791
	Expected	$0.851 \pm 0.034$	$0.792 \pm 0.039$	$0.824 \pm 0.036$	$0.799 \pm 0.038$
Tests for Hardy-Weinberg equilibrium	Homozygosity	$\chi^2 = 0.285, df=1$ $P = 0.59$	$\chi^2 = 0.001, df=1$ $P = 0.99$	$\chi^2 = 0.848, df=1$ $P = 0.36$	$\chi^2 = 0.106, df = 1$ $P = 0.75$
	Exact	$\chi^2 = 32.76, df = 35$ $P = 0.58$	$\chi^2 = 30.71, df = 27$ $P = 0.26$	$\chi^2 = 39.32, df = 44$ $P = 0.67$	$\chi^2 = 60.02, df = 44$ $P = 0.06$
	LR	$\chi^2 = 34.39, df = 35$ $P = 0.49$	$\chi^2 = 29.57, df = 27$ $P = 0.34$	$\chi^2 = 39.39, df = 44$ $P = 0.67$	$\chi^2 = 60.97, df = 44$ $P = 0.05$

Allele frequencies of 4 STR systems analysed in this study were compared (using Carmody’s test) with frequencies of American Caucasian population data given by the MultiPlex II manufacturer [18]. The comparison

shows statistically important differences for all 4 STR systems, which indicates the heterogeneity of these two populations.

Population data for FGA and D7S820 systems were compared to published Polish population data from: the northern region [16], Pomerania-Kujawy [13, 14] and Lower Silesia [5]. The comparison indicates that the south-east population is homogenous with other Polish populations in terms of the D7S820 locus. For the FGA locus the comparison shows existence of significant differences with the Pomerania-Kujawy population –  $P(\chi^2) = 0.0001$ ,  $P(G) = 0.0001$  and homogeneity with the remaining Polish populations.

The usefulness of the four systems of the MultiPlex II (FGA, D7S820, D1S533, D9S304) in cases concerning paternity was assessed on the basis of comparison with earlier forensic genetic expert studies carried out in 87 selected cases. Out of a total of 41 cases where paternity had been excluded, analysis by the MultiPlex II kit confirmed exclusion in 40 cases (97.56%). Confirmation of exclusion required study of only one of the STR systems in 9 of these cases, and in the remainder – two or more systems. The observed percentage of paternity exclusion (97.56%) was therefore very close to the theoretical chance of exclusion, which was 98%.

The value of the paternity index was analysed in 46 cases with confirmed paternity. It was ascertained that the highest combined value of  $PI$  for the MultiPlex II amounted to 2924 (corresponding to probability of paternity = 0.9996) and the lowest 6 (corresponding to probability of paternity = 0.8571).  $PI$  values close to one (i.e. those which have a minimal influence on combined  $PI$  value, and therefore also on the MultiPlex II) were observed twice for D7S820 and D1S533 systems, three times in the D9S304 system and were not noted in the case of the FGA system.

TABLE III. MEAN VALUES OF PATERNITY INDEX ( $PI$ ) OBTAINED IN 46 CASES WITH CONFIRMED PATERNITY FOR 4 ANALYSED STR SYSTEMS

DNA System	Mean $PI$ values			
	Theoretical mean acc. [2]	In analysed paternity cases		
		Mean probability	Geometric mean	Mean medians
FGA	3.516	2.792	2.609	3.337
D7S820	2.392	2.278	2.088	2.917
D1S533	2.541	2.562	2.320	3.099
D9S304	2.392	2.685	2.858	3.709
MultiPlex II	51	44	36	112

Mean  $PI$  values were calculated from individual  $PI$  values of STR systems and combined  $PI$  value of MultiPlex II in 46 cases of confirmed paternity. Mean  $PI$  values were calculated by three different methods: as a geometric mean, mean probability and mean median. Calculation results presented in Table III indicate the significant influence of the calculation

method on the mean value of  $PI$ . The lowest variability (depending on the number of analysed cases) is shown by the mean  $PI$  value calculated from the set of probability of paternity ( $PDP$ ) and is similar to the theoretical mean  $PI$  value obtained from Brenner et al.'s [2] formula.

This means that the method of calculation of mean  $PI$  has a bearing on assessment of the usefulness of the analysed MultiPlex II (and any other genetic markers) in paternity confirmation. The extreme values of  $PI$  (36–112) result in a probability of paternity values from 0.974 ( $PI = 36$ , which would signify low usefulness of the assessed multiplex) to 0.991 ( $PI = 112$ , indicating high usefulness of the assessed multiplex).

Table IV presents statistical parameters evaluating usefulness of the whole MultiPlex II kit (and each of the 4 STR systems as well as other commercially available kits with silver detection) for forensic genetic expert studies.

TABLE IV. COEFFICIENTS ASSESSING USEFULNESS OF STR MULTIPLEXES WITH SILVER DETECTION FOR GENETIC AND FORENSIC EXAMINATION (WITH 4 MULTIPLEX II STR SYSTEMS)

Commercial multiplexes	$PE$	$MPI$	$PI$	$PD$	$DI$	$PIC$
Multiplex II (Lifecodes)*	0.980	51.2	44.7	0.99995	$1.54 \times 10^{-5}$	—
FGA *	0.716	3.516	2.792	0.956	0.036	0.831
D7S820*	0.582	2.392	2.278	0.922	0.078	0.760
D1S533*	0.606	2.541	2.562	0.948	0.069	0.802
D9S304*	0.582	2.392	2.685	0.937	0.078	0.777
MultiPlex I (Lifecodes)** D12S1090, D3S1744, D18S849	0.966	37.25	32.96	0.99991	$9.1 \times 10^{-5}$	—
CTT (Promega) [12] CSF1P0, TPOX, TH01	0.884	6.896	7.343	0.9982	$2.4 \times 10^{-3}$	—
FFv (Promega) [12] F13A01, FESFPS, vWA	0.928	9.936	—	0.9988	$1.9 \times 10^{-3}$	—
SilverSTR III (Promega) [9]*** D16S539, D7S820, D13S317	0.919	9.722	9.984	0.9993	$9.7 \times 10^{-4}$	—

$PE$  = power of exclusion,  $MPI$  = mean paternity index,  $PI$  = paternity index (from mean  $PDP$ ),  $PD$  = power of discrimination,  $DI$  = discrimination index,  $PIC$  = polymorphism information content.

\* Coefficients calculated on the basis of presented study results.

\*\* On the basis of unpublished results of own study.

\*\*\* Results obtained for fluorescent multiplex Gamma calculated for the 3 enumerated STR systems i.e. without D5S818.



The high values of power of discrimination (*PD*), polymorphism information content (*PIC*) and discrimination index (*DI*) indicate that MultiPlex II kit is a useful tool in criminalistic analysis and paternity testing.

It is well known, however, that currently there is strong pressure to standardise the criminalistic research range and create a national data base: this means that use of a particular STR system, or group of STR systems in forensic analysis depends on whether it has been included in international data bases. It is worth noting that the D1S533 and D9S304 loci incorporated in the MultiPlex II kit have not as yet been included in any of the databases created so far [4].

The requirement of concordance between the research range and the database range does not concern analysis made in paternity cases; moreover results of these analyses cannot be inputted into data bases. So, the value of the chance of paternity exclusion and mean *PI* value of MultiPlex II together with the technical ease of analysis and low kit cost, confirm its high usefulness in paternity testing.

In paternity cases, in order to submit an expert opinion confirming paternity of alleged father with a probability close to certainty, it is necessary to obtain *PDP* > 99.999%. Therefore, analysis results of MultiPlex II are not sufficient basis for such an opinion, but could be the fundamental element of a genetic expert opinion based on STR systems with use of the cheap silver detection technique.

#### References:

1. Bassam B. J., Caetano-Anolles G., Gresshoff P. M., Fast and sensitive silver staining of DNA polyacrylamide gels, *Analytical Biochemistry* 1991, vol. 196, pp. 80–83.
2. Brenner C., Morris J. W., Paternity index calculations in single locus hypervariable DNA probes: validation and other studies, in: Proceedings from the International Symposium on Human Identification, Promega Corporation 1989, pp. 21–53.
3. Brinkmann B., The STR approach, *Advances of Forensic Haemogenetics* 1996, vol. 6, pp. 41–51.
4. Butler J. M., Forensic DNA Typing. Biology & technology behind STR markers, Academic Press, London 2001.
5. Dobosz T., Częstości wybranych genów warunkujących polimorfizm genetyczny w obrębie populacji dolnośląskiej, Akademia Medyczna we Wrocławiu, Wrocław 2000.
6. Edwards M. C., Gibbs R. A., Multiplex PCR: advantages, development and applications, *PCR Methods and Applications* 1994, vol. 3, pp. 565–575.
7. Guo S. W., Thompson E. A., Performing the exact test of Hardy-Weinberg proportion for multiple alleles, *Biometrics* 1992, vol. 48, pp. 361–372.

8. Koziół P., Czerski T., Mądro R., Population genetic data for HumF13B, HUM LPL and HumPRTB in southeast Poland, *International Journal of Legal Medicine* 1999, vol. 113, pp. 55–57.
9. Koziół P., Ciesielka M., Mądro R., Analysis of the short tandem repeat systems D5S818, D7S820, D16S539 and D13S317 in a population sample from south-east Poland [in press].
10. Koziół P., Ciesielka M., Mądro R., Genetic data on 9 complex STR systems in the population of south-east Poland and their usefulness in parentage testing, *Problems of Forensic Sciences* 2001, vol. 45, pp. 66–80.
11. Lins A. M., Micka K. A., Sprecher C. J. [et al.], Development and population study of an eight-locus short tandem repeat(STR)multiplex system, *Journal of Forensic Sciences* 1998, vol. 43, pp. 1–13.
12. Mądro R., Ciesielka M., Monies D. [i in.], Przydatność tripleksów CTT i FFv w sprawach spornego ojcostwa, *Postępy Medycyny Sądowej i Kryminologii* 1999, t. 5, s. 365–374.
13. Miścicka-Śliwka D., Czarny J., Berent J. A. [et al.], Population genetics of the STRs vWA, D3S1358 and FGA in the Pomerania-Kujawy region of Poland, *International Journal of Legal Medicine* 1999, vol. 112, pp. 391–392.
14. Miścicka-Śliwka D., Czarny J., Grzybowski T. [et al.], Population genetics of the STRs vWA, TH01, TPOX, CSF1PO, D5S818, D13S317, D7S820, D16S539, LPL, F13B, FESFPS, F13A01 and ACTBP2 in the Pomerania-Kujawy region of Poland, *Forensic Science International* 2000, vol. 119, pp. 119–122.
15. Nei M., Sampling variance of heterozygosity and genetic distance, *Genetics* 1974, vol. 76, pp. 379–390.
16. Pawłowski R., Dettlaff-Kąkol A., Jezierski G. [i in.], Genetyka populacyjna dziewięciu loci typu STR z zestawu Profiler Plus w próbie populacyjnej z obszaru Polski, *Archiwum medycyny sądowej i kryminologii* 2000, t. L, s. 207–213.
17. Turowska B. M., Sanak M., Opolska-Bogusz B., Częstość alleli systemów STR: D13S317, D7S820 i D16S539 w populacji Polski Południowej, *Archiwum medycyny sądowej i kryminologii* 2000, t. L, s. 193–197.
18. Quick-Type™ user's manual, Lifecodes Corporation, USA 1997.
19. Weir B. S., Independence of VNTR alleles defined as fixed bins, *Genetics* 1992, vol. 130, pp. 873–887.

## POLIMORFIZM 4 UKŁADÓW STR (D9S304, D1S533, D7S820, FGA) W POPULACJI POLSKI POŁUDNIOWO-WSCHODNIEJ

Piotr KOZIOŁ, Sylwia CHOCHOLSKA, Roman MADRO

### WSTĘP

Polimorfizm DNA typu STR znalazł powszechne zastosowanie w genetyce sądowej po opracowaniu techniki powielania wielu *loci* STR w jednej reakcji PCR i genotypowania w jednym procesie rozdzielania elektroforetycznego oraz detekcji [3, 4, 6]. Równoczesne genotypowanie kilkunastu układów STR wymaga jednak użycia kosztownej aparatury do automatycznej analizy fragmentów allelicznych znakowanych fluorochromami [11]. Alternatywną technikę detekcji produktów PCR stanowi tymczasem tania metoda barwienia srebrem, której zastosowanie pozwala na zróżnicowanie w jednej elektroforezie od 3 do 4 *loci*. Wykonanie kilku reakcji multipleksowych pozwala zatem uzyskać ilość informacji porównywalną z tą, którą daje technika fluoroscencyjna. Przykładem może być użycie (do amplifikacji kilku *loci* i detekcji srebrzej) komercyjnych zestawów produkowanych przez firmę Promega pozwalających w trzech reakcjach PCR uzyskać powielenie alleli 9 układów STR (CSF1P0, TPOX, TH01, F13A01, FESFPS, vWA, D16S539, D7S820, D13S317), które były przedmiotem publikacji [12, 17]. Pozwalają na to również takie zestawy firmy Lifecodes, jak MultiPlex I (D18S849, D3S1744 i D12S1090) oraz MultiPlex II (D9S304, D1S533, D7S820, FGA), umożliwiające genotypowanie 7 układów STR. Nie były one jednak dotychczas testowane w populacji polskiej.

Celem pracy było więc oznaczenie genotypów 4 lokus DNA (D9S304, D1S533, D7S820 i FGA) w próbie 301 niespokrewnionych mieszkańców Polski południowo-wschodniej, ustalenie częstości ich alleli, wykonanie obliczeń statystycznych w zakresie segregacji tych alleli, porównanie wyników badań autorów tej pracy z częstościami uzyskanymi dla innych populacji oraz ocena przydatności multipleksu w sprawach spornego ojcostwa.

### MATERIAŁ I METODY

Do badań użyto 301 próbek DNA, które wyizolowano z krwi pobranej w Zakładzie Medycyny Sądowej AM w Lublinie od niespokrewnionych mieszkańców Polski południowo-wschodniej obojga płci w związku z wykonywaniem ekspertyz genetycznych w sprawach o ustalenie ojcostwa. Przebadano również próbki DNA wyizolowane od dzieci w 87 sprawach alimentacyjnych, w których rutynowe oznaczenia 4 układów VNTR (D4S139, D5S110, D7S21, D11S12) i 9 układów STR (PowerPlex 2.1) w 46 sprawach potwierdziły ojcostwo pozwanego z prawdopodobieństwem graniczącym z pewnością, a w 41 przypadkach wykluczyły ojcostwo pozwanego.

DNA wyizolowano metodą organiczną z użyciem mieszaniny fenol-chloroform, a następnie oznaczano jego stężenie metodą slot-blot z zastosowaniem sondy D17Z1 (Gibco-BRL, Gathersburg, MD, Stany Zjednoczone). Do jednoczesnej amplifikacji

czterech układów STR (D9S304, D1S533, D7S820 i FGA) użyto zestawu MultiPlex II kit (Lifecodes Corporation) [18]. Reakcję PCR wykonywano w objętości 20  $\mu$ l zgodnie z warunkami producenta zestawu, używając cyklera Trio-block firmy (Biometra). Produkty PCR były rozdzielane w denaturujących 6% żelach poliakryloamidowych (4% T, 6% C, 8 M mocznik) w buforze 0.5  $\times$  TBE, pH = 8,3 przez 1,5 godziny przy mocy prądu 60 W. Fragmenty DNA wybarwiono następnie metodą srebrową [1], a genotypy odczytywano przez porównanie do wzorca w postaci drabiny alleli. Prawidłowość genotypowania układu FGA i D7S820 tą metodą sprawdzono przez porównanie jej rezultatów z wynikami oznaczeń uzyskanymi u 170 osób techniką detekcji fluoroscencyjnej przy pomocy zestawu PowerPlex 2.1 firmy Promega oraz skanera FMBIO II firmy Hitachi [10]. Wielkość allela 16 z lokus D1S533 ustalono metodą lokalnej korelacji przy użyciu wideokamery i programu do analizy obrazu BioImage firmy Biotek-Fischer. Zastosowano przy tym markery drabiny MultiPlex II, ograniczając pomiar do obszaru między fragmentem 203 pz – allel 15 lokus D1S533 i fragmentem 217 pz – allel 7 lokus D7S820.

Ustalono częstości alleli w badanej populacji i porównano je (przy użyciu programu 2-way R  $\times$  C wg G. Carmody'ego (Carleton University, Ottawa, Kanada) z częstościami stwierdzonymi w innych populacjach polskich oraz populacji amerykańskiej (rasy białej). W celu sprawdzenia równowagi segregacji alleli zgodnie z regułą Hardy'ego-Weinberga wykonano także analizę statystyczną wyników badań testami:  $\chi^2$  [15], „exact” [7] i „LR” [19]; obliczono także parametry określające przydatność tych markerów do badań sądowych [8]. Średnie „teoretyczne” wartości indeksu ojcostwa *PI* obliczone ze wzoru Brennera i in. [2] porównano ze średnimi *PI* uzyskanymi w badanych sprawach alimentacyjnych, przy czym obliczono je trzema różnymi sposobami, tzn. średnią geometryczną i medianę ze zbioru *PI*, a średnią prawdopodobieństwa jako średnią arytmetyczną ze zbioru prawdopodobieństwa ojcostwa (*PDP*), którą następnie zamieniono na wartość *PI*.

#### WYNIKI I DYSKUSJA

Zestaw MultiPlex II firmy Lifecodes oraz zastosowana technika rozdziału fragmentów allelicznych i srebrowa detekcja pozwoliły na ustalenie genotypów w zakresie czterech *loci* we wszystkich badanych próbkach. Producent MultiPlex II zaleca wprawdzie użycie 4% denaturujących żeli poliakryloamidowych, ale w opisanych tu badaniach zastąpiono je żelami 6%. Rozdział produktów amplifikacji w denaturujących warunkach elektroforezy powoduje bowiem, że widoczne są różnice w migracji komplementarnych nici DNA. W przypadku użycia żeli 4% w układzie FGA i D1S533 komplementarne nici DNA jednego allela były w związku z tym słabo odseparowane od fragmentów większego allela. Utrudniało to identyfikację genotypu (rycina 1a), a problemu tego nie udało się rozwiązać przez wydłużanie rozdziału elektroforetycznego. Zdecydowanie czytelniejsze rozdziały uzyskano, stosując 6% żele denaturujące, na których rozdział komplementarnych nici FGA i D1S533 jest tak bliski, że tworzą nieomal jedną frakcję, przy czym alleliczne fragmenty drabiny są wystarczająco dobrze odseparowane od siebie. Czytelność rozdziału pozostałych układów multipleksu D7S820 i D9S304 na tych żelach nie sprawia natomiast problemów przy odczycie ich genotypów (rycina 1b).

Wizualnie identyfikowane allele wykazywały szybkość migracji identyczną do alleli wzorcowej drabiny z dwoma wyjątkami. Pierwszy z nich dotyczy allele 16 w układzie D1S533, który nie miał w drabinie wzorcowej swojego odpowiednika. Wielkość tego fragmentu w każdym z 9 analizowanych przypadków wynosiła  $207 (\pm 0,6)$  pz i była większa od allele 15 o czteronukleotydomową sekwencję powtarzalną (pozycja 9 i 14 na rycinie 1b). Drugi problem dotyczył natomiast układu FGA, którego drabina alleli zawiera wzorce jedynie dla 8 alleli przyporządkowanych allelom z parzystą liczbą powtórzeń, przy czym allele oznaczone liczbami nieparzystymi układają się między allelami tej drabiny. Nie istnieje więc możliwość identyfikacji interalleli w układzie FGA. Osoby, u których wcześniej oznaczono w układzie FGA (przy użyciu zestawu PowerPlex 2.1 i fluoroscencyjnej techniki detekcji) interallele 22.2, 23.2 i 24.2, w technice srebrowej odczytano je odpowiednio jako allele 23, 24 oraz 25.

Stwierdzono ponadto, że reakcja amplifikacji przy użyciu MultiPlex II wymaga znacznej ilości matrycy (4 ng) oraz, że niewielkie zmniejszenie stężenia DNA prowadzi do słabszej amplifikacji fragmentów w lokus D1S533.

Wyniki genotypowania i obliczone częstości genowe przedstawiono w tabeli I. Wynika z niej, że w lokus FGA zidentyfikowano 37 genotypów utworzonych przez 9 alleli, z których najczęściej występował allel 22 (0,219), w lokus D7S820 znajdowało się 8 alleli (24 genotypy), z których najczęstsze to allele 10 (0,319) i 11 (0,201), natomiast dwa allele – 7 i 14 – miały częstość poniżej 0,5%. W loci D1S533 i D9S304 stwierdzono natomiast po 10 alleli i 40 genotypów. W układzie D1S533 najczęstsze były allele 18 (0,307) i 19 (0,2126), a częstość poniżej 0,5% miał tylko jeden allel 7; natomiast w układzie D9S304 najczęściej występował allel 8 (0,377), a najniższą częstość miał allel 15 (0,008).

Z tabeli II wynika, że rozkład alleli przebadanych loci STR w populacji Polski południowo-wschodniej jest zgodny z prawem Hardy'ego-Weinberga oraz że w układach tych nie stwierdzono istotnych rozbieżności między obserwowaną i oczekiwaną heterozygotycznością. Należy przy tym podkreślić, że testy „exact” i „LR” wykonane bezpośrednio w oparciu o wyniki badań 301 osób w układzie D9S304 dały niskie prawdopodobieństwo ( $P = 0,06$  i  $P = 0,05$ ), natomiast w rezultacie obliczeń metodą Monte Carlo przy liczbie 10 000 permutacji [7] wykazały wysokie prawdopodobieństwo ( $P = 0,992$  i  $P = 0,984$ ).

Uzyskane w tej pracy częstości alleli czterech układów STR zostały porównane (przy pomocy testu Carmody'ego) z częstościami podanymi przez producenta MultiPlexu II [18] dla populacji amerykańskiej (rasy białej) i stwierdzono istotne statystycznie różnice we wszystkich 4 układach, co świadczy o heterogenności tych populacji. Populacje polskie porównano jedynie w zakresie układów FGA i D7SS820, dla których dotychczas opublikowano dane z regionów: północnego [16], Pomorza-Kujaw [13, 14] i dolnośląskiego [5]. W ten sposób ustalono, że populacja Polski południowo-wschodniej jest homogenna z innymi populacjami polskimi w układzie D7S820 oraz że w układzie FGA różni się ona istotnie z populacją Pomorza-Kujaw –  $P(\chi^2) = 0,0001$ ,  $P(G) = 0,0001$ , a jest homogenna z pozostałymi.

Przydatność czterech układów STR MultiPlexu II (FGA, D7S820, D1S533, D9S304) w sprawach dotyczących spornego ojcostwa oceniono na podstawie porównania z wcześniej przeprowadzonymi ekspertyzami genetycznymi w 87 wybranych sprawach. W ten sposób stwierdzono, że wykluczenie ojcostwa zostało potwierdzone

w 40 przypadkach na 41 spraw z „ojcostwem wykluczonym” (tj. w 97,56%). W 9 przypadkach potwierdzenie wykluczenia stanowiły przy tym rezultaty badań tylko jednego układu STR, a w pozostałych na dwóch lub więcej układach. Obserwowany odsetek wykluczeń ojcostwa (97,56%) był zatem bardzo zbliżony do teoretycznej szansy wykluczenia, która wynosiła 98%.

W 46 sprawach z „ojcostwem potwierdzonym” przeanalizowano natomiast wartości indeksu ojcostwa i stwierdzono, że najwyższa łączna wartość  $PI$  dla MultiPlexu II wynosiła 2924 (co odpowiada prawdopodobieństwu ojcostwa = 0,9996), a najniższa 6 (co daje prawdopodobieństwo ojcostwa = 0,8571). Wartości  $PI$  bliskie jedności (tj. takie, które mają minimalny wpływ na łączną wartość współczynnika ojcostwa, czyli także dla MultiPlexu II) wystąpiły dwukrotnie w układach D7S820 i D1S533; trzykrotnie w układzie D9S304, nie stwierdzono natomiast takich wartości w przypadku układu FGA.

Z wartości  $PI$  dla poszczególnych układów w 46 sprawach z „ojcostwem potwierdzonym” oraz z łącznej wartości  $PI$  dla MultiPlexu II obliczono następnie średnie wartości  $PI$  trzema różnymi sposobami, tzn. średnią po prawdopodobieństwie, średnią geometryczną i średnią medianą. Rezultaty tych obliczeń przedstawiono w tabeli III, z której wynika, że sposób obliczania średniej ma znaczny wpływ na średnią wartość  $PI$ , przy czym najmniejszą zmienność (zależną od liczby przebadanych spraw) wykazuje średnia  $PI$  liczona z prawdopodobieństwa ojcostwa  $PDP$  i jest zbliżona do wartości średniej  $PI$  obliczonej teoretycznie wg wzoru Brennera i in. [2]. Oznacza to, że sposób obliczeń średniej  $PI$  rzutuje na ocenę przydatności analizowanego MultiPlexu II (i każdego innego markera genetycznego) do potwierdzania ojcostwa. Skrajne wartości  $PI$  (36–112) kreuja bowiem prawdopodobieństwa ojcostwa od 0,974 ( $PI = 36$ , co oznaczałoby małą przydatność ocenianego multipleksu) do 0,991 ( $PI = 112$ , która to wartość świadczy o jego dużej przydatności).

W tabeli IV podano statystyczne parametry oceniające przydatność całego zestawu MultiPlex II (i każdego z czterech badanych układów STR oraz innych multipleksów komercyjnych z detekcją srebrową) dla genetycznych ekspertyz sądowych. Za dużą przydatnością MultiPlexu II w badaniach kryminalistycznych i przy ustalaniu ojcostwa przemawiają przedstawione w niej wysokie wartości współczynnika dyskryminacji ( $PD$ ), współczynnika informacji o polimorfizmie ( $PIC$ ) i indeksu dyskryminacji ( $DI$ ).

Wiadomo jednak, że obecnie istnieje silna presja na unifikację zakresu badań kryminalistycznych i tworzenie narodowych baz danych, z czego wynika, że zastosowanie konkretnego układu lub zespołu układów STR do badań zależy od tego, czy został on uwzględniony w międzynarodowych bazach danych. Tymczasem *loci* D1S533 i D9S304, które wchodzi w skład zestawu MultiPlex II, nie zostały dotychczas umieszczone w żadnej z tworzonych baz danych [4].

Wymóg zgodności zakresu badań z zakresem bazy danych nie dotyczy ekspertyz wykonywanych w sprawach spornego ojcostwa, ponadto wyniki z tych badań nie mogą być przenoszone do tych baz. W tej sytuacji wartość szansy wykluczenia ojcostwa i średnia wartość  $PI$  dla MultiPlexu II oraz techniczna łatwość genotypowania układów tego multipleksu, jak również względnie niski koszt tego zestawu, wskazują na dużą jego przydatność w badaniach spornego ojcostwa. W sprawach spornego ojcostwa do wydania opinii potwierdzającej ojcostwo pozwanego z prawdopodobieństwem graniczącym z pewnością konieczne jest jednak uzyskanie  $PDP > 99,999\%$ .

Wnioskowanie wyłącznie na podstawie MultiPlexu II jest zatem niewystarczające do wydania takiej opinii, może jednak stanowić podstawowy element ekspertyzy genetycznej wykonywanej w oparciu o układy STR z zastosowaniem taniej detekcji srebrowej.