Int J Legal Med (2000) 114 : 45–49 © Springer-Verlag 2000

ORIGINAL ARTICLE

R. Pawlowski · A. Maciejewska

Forensic validation of a multiplex containing nine STRs – population genetics in Northern Poland

Received: 15 September 1999 / Accepted: 14 January 2000

Abstract This paper presents the allele frequency distributions for the nine loci (D3S1358, VWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317 and D7S820) present in the commercially available Profiler Plus kit. DNA samples of 202 individuals from Northern Poland were amplified in a multiplex reaction with subsequent automatic detection using capillary electrophoresis. All loci met the Hardy-Weinberg equilibrium conditions. The calculated probability of identity was 2.26×10^{-11} giving an average probability of identity of 1 in 44 billion. Considerable improvement of analysis precision was observed after substituting the GeneScan 500 for the fluorescent ladder 60–400 bp. The influence of DNA concentration on stutter and artefact formation and the ratio of heterozygote alleles was analysed.

Key words Profiler Plus · Validation studies · Capillary electrophoresis · Population genetics · Poland

Introduction

Short tandem repeat (STR) loci are widely used in forensic genetics for identification and paternity testing [1]. The introduction of multiplex PCR techniques allowed simultaneous, rapid and robust amplification of several DNA loci from minute biological stains or fresh blood [2, 3, 4]. One of the commercially available multiplex kits is the Profiler Plus from Perkin-Elmer (PE) consisting of the nine polymorphic loci D3S1358, VWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317 and D7S820 and amelogenin. Using the fluorescent detection technology simultaneous automatic analysis of all ten amplified loci can be carried out [5]. The correct interpretation of DNA profiles requires an identification of the characteristics of individual loci and of the influence of various factors which can cause artefacts. The first aim of our work was an analysis of samples from a Northern Poland population using the Profiler Plus kit and the second was a validation study from the point of view of forensic usefulness and credibility, with particular reference to the principles formulated by TWGDAM.

This paper is dedicated to Professor Bernd Brinkmann on the occasion of his 60th birthday

R. Pawlowski · A. Maciejewska
Institute of Forensic Medicine, Medical University of Gdansk, 80-210 Gdansk, Debinki 7, Poland
e-mail: Richard@amg.gda.pl; Fax: +48-58-3410485

Materials and methods

DNA was isolated from 202 blood samples from unrelated individuals living in Northern Poland. The extraction was carried out as described previously [6]. DNA was quantified using the QuantiBlot kit (PE, USA) with chemiluminescence detection. Amplification of the ten loci included in the Profiler Plus kit (PE, USA) was carried out in accordance with the manufacturer's instructions on 2400 or 877 Thermal Cyclers (PE, USA). Detection of PCR products using capillary electrophoresis was carried out in an ABI Prism 310 Genetic Analyser. From the PCR product 1 µl was mixed with 12 µl of deionised formamide and 1 µl of internal standard. The two internal DNA standards used were the Perkin-Elmer GS500 labelled with ROX and Promega fluorescent ladder 60–400 bp labelled with CXR (FL-CXR). Before electrophoresis the samples were denatured for 3 min at 95 °C and subsequently snap-cooled on ice. The samples were run on a capillary 47 cm in length filled with the denaturing polymer POP4. The separation was conducted for 24 min at 15 kV, 9 mA and 10 µW. The analysis of the fragments was carried out using the computer programme GeneScan v. 2.1.

For sequencing the rare FGA alleles the FGA primers described by Mills et al. [7] were used and sequencing was done in both directions. The allele to be sequenced was excised from the silver stained polyacrylamide gel, eluted and purified with Centricon100 (Amicon) spin column [8]. Sequencing using BigDye Terminator sequencing kit (PE) was performed in 20 µl containing 3.2 pmol of forward or reverse sequencing primer according to the manufacturer's instruction. Sequencing products were precipitated with absolute ethanol and 3 M sodium acetate. Electrophoresis of the sequencing products was carried out on ABI310 [9]. The data were analysed using ABI Prism sequencing analysis software version 2.1 (PE).

Comparison of analysis precision using different internal DNA standards

The Profiler Plus allelic ladder was injected many times on different capillaries and using different batches of POP4 gels. The GS500 (PE) and FL-CXR (Promega) were used as internal standards and standard deviations of analysis precision for 20 randomly chosen runs with GS500 and FL-CXR were calculated using the Excel 6 Microsoft computer programme. For some validation studies the DNA standard from the Quantiblot kit (PE) was used and the minimal cut-off level was established at 150 relative fluorescence units (RFU).

Statistical calculations

Compliance with the Hardy-Weinberg equilibrium was determined using the likelihood ratio test [10] and exact test [11]. The distribution homogeneity of alleles of the Profiler Plus loci among different populations, the expected frequency of heterozygotes (H),

Table 1 Allele frequencies and statistical calculations for nine loci included in the Profiler Plus in the population sample from Northern Poland ($n = 202$)

Allele	D3S1358	VWA	FGA	D8S1179	D21S11	D18S51	D5S818	D13S317	D7S820
6									0.0025
7							0.0025		0.0074
8				0.0074				0.1188	0.1361
9				0.0124			0.0520	0.0743	0.1708
10				0.0644		0.0124	0.0668	0.0421	0.2871
11				0.0718		0.0050	0.3069	0.4035	0.1931
12				0.1856		0.0941	0.4183	0.2277	0.1782
13	0.0025	0.0049		0.3193		0.0891	0.1411	0.0842	0.0228
14	0.1361	0.1188		0.2203		0.1460	0.0124	0.0470	0.0025
15	0.2574	0.1312		0.1040		0.2079		0.0025	
16	0.2351	0.1980	0.0050	0.0124		0.1634			
17	0.1931	0.2549		0.0025		0.1337			
18	0.1609	0.1856	0.0124			0.0767			
19	0.0149	0.0842	0.0941			0.0445			
20		0.0173	0.1634			0.0149			
21		0.0050	0.1881			0.0050			
21.2			0.0025						
22			0.1980			0.0074			
22.2			0.0124						
23			0.1139						
23.2			0.0074						
24			0.0941						
25			0.0569						
25.2			0.0148						
26			0.0371						
27					0.0173				
28					0.1535				
29					0.2203				
29.2					0.0025				
30					0.2252				
30.2					0.0569				
31					0.0965				
31.2					0.0693				
32					0.0099				
32.2					0.1040				
33					0.0025				
33.2					0.0322				
34.2					0.0050				
38					0.0050				
PD	0.919	0.940	0.961	0.931	0.950	0.963	0.864	0.899	0.932
PIC	0.765	0.798	0.848	0.766	0.830	0.851	0.656	0.724	0.771
MEC	0.594	0.646	0.725	0.604	0.698	0.730	0.465	0.553	0.604
LR test	0.895	0.151	0.295	0.736	0.432	0.478	0.675	0.486	0.543
Exact test	0.850	0.190	0.337	0.777	0.449	0.506	0.715	0.453	0.593
$H_{obs.}$	0.851	0.822	0.876	0.777	0.896	0.871	0.703	0.792	0.762
H_{exp}	0.798	0.825	0.865	0.796	0.850	0.868	0.705	0.756	0.801

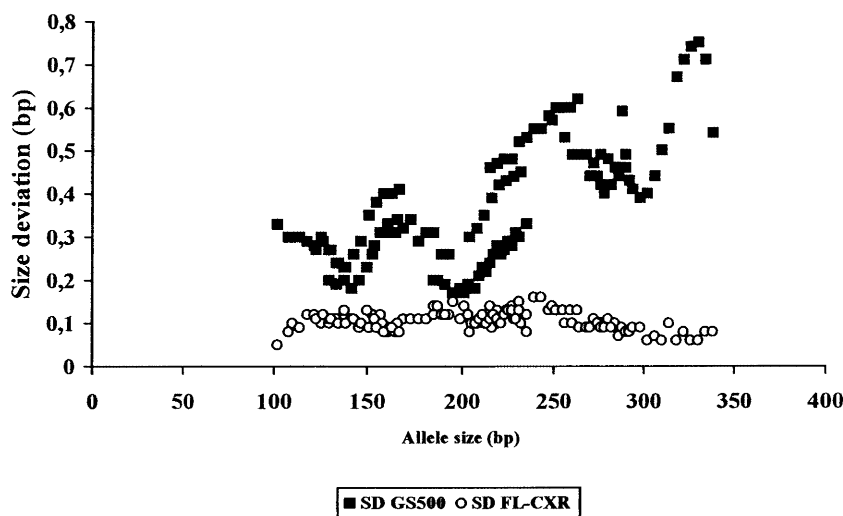

Fig. 1 Size deviation of Profiler Plus allelic ladders run with GS500 and FL-CXR 60-400. The analysis precision (SD) for two internal standards plotted as a function of length of Profiler Plus allelic ladder alleles

power of discrimination (PD) and the polymorphic information content (PIC) were also calculated [12].

Results and discussion

In the great majority of the samples analysed using the Profiler Plus kit no problems with amplification were encountered. In a few cases it was necessary to reinject samples or allelic ladders because of problems with the allele designations, especially in the case of interalleles of the D18S51, FGA and D21S11 loci. It was observed that the three loci labelled with NED (D5S818, D13S317 and D7S820) gave the lowest amplification signals in comparison to JOE and 5-FAM labelled loci. The highest amplification signals were observed for D3S1358, VWA and FGA labelled with 5-FAM (data not shown). Analysis of fresh DNA samples, as in the case of population studies, gave no problems but for degraded biological stains, and especially mixtures of biological fluids with a low admixture of one of the components there was no signal for the yellow (NED-labelled) loci which creates an interpretation problem (data not shown).

Table 1 presents allele frequencies observed for the nine STR loci in a population sample from Northern Poland ($n = 202$). Comparison of the VWA and FGA allele distribution using monoplex reactions [6, 12] and Profiler Plus typing, for two different population samples from the same region of Poland showed no statistical differences ($p = 0.2190$ for FGA and $p = 0.1020$ for VWA).

A comparison of the homogeneity of allele frequency distribution among a population from Andalusia (Spain) [5] and our population, showed statistically significant differences for FGA, D21S11, D18S51, D5S818, D13S317 and D8S1179 ($p < 0.05$) and similar allele distributions for VWA ($p > 0.7$), D7S820 ($p > 0.35$) and D3S1358 ($p > 0.25$).

As Table 1 shows, all loci meet Hardy-Weinberg equilibrium (HWE) conditions using the likelihood ratio and exact tests. The most polymorphic and discriminating loci in our population were D18S51, FGA and D21S11 and the least informative D5S818 which is similar to the results from Spain [5]. The calculated probability of identity was 2.26×10^{-11} giving an average probability of identity of 1 in 44 billion which is slightly lower than for the Andalusian population sample (average 1 in 66 billion).

To summarise, a Polish population database was established for loci included in the Profiler Plus system. The high cumulative discrimination index of Profiler Plus makes this a very useful system for biological stain identification and paternity testing.

Sequencing of the rare FGA*16 allele

During population tests the presence of an allele 208 bp in size was confirmed which corresponds to the FGA 16 allele. This allele was sequenced in both directions using the BigDye Terminator cycle sequencing kit and the sequence confirmed as $(TTTC)_3$ TTTT TTCT $(CTTT)_8$ CTCC $(TTCC)_2$ which corresponds to the FGA allele 16 which had not previously been found in the Polish population. The sequence is consistent with the sequence of repeat region of shorter FGA alleles reported by Barber et al. [13]

Comparison of the precision of the allele size determination for the Profiler Plus with the application of the internal standards GS 500 and the FL-CXR 60–400 bp

Some precision problems during analysis of FGA alleles on the ABI310 sequencer using the monoplex reaction were shown earlier [6]. The precision of the determination of allele size after separation using the capillary electrophoresis method is one of the critical parameters significantly affecting correct phenotyping. An attempt was made to improve the precision by applying a different internal standard in the form of the fluorescent ladder CXR

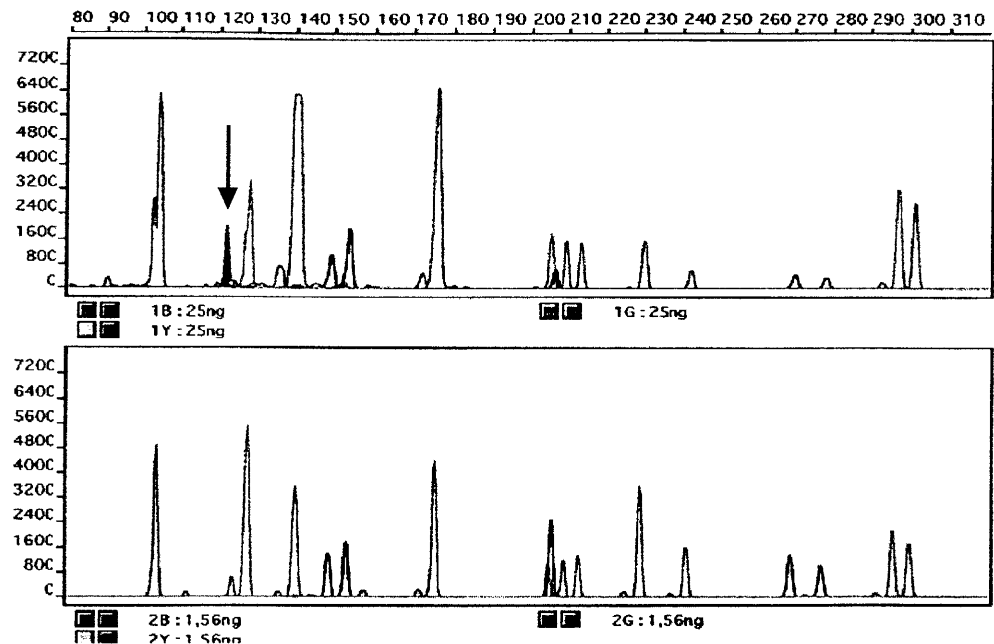

Fig. 2 Influence of DNA quantity on amplification of Profiler Plus (*top panel* 25 ng of DNA, *bottom panel* 1.56 ng DNA in a PCR mixture. In the upper panel an artificial peak of 122 bp is indicated by an arrow)

Promega (FL-CXR). Figure 1 presents the comparison of the precision of analysis with the application of GS500 and FL-CXR obtained by multiple injections on various capillaries (SD between capillaries, $n = 20$) of the allele ladders originating from the Profiler Plus kit. High SD values for GS500 were confirmed, ranging from 0.18 to 0.75 bp and a better precision was obtained with the application of FL-CXR. All the SD values were located within the range 0.06–0.16 bp for between-capillaries runs which are similar to those from Profiler Plus (User's Manual for Profiler Plus kit, PE), but for within capillary runs. Such high precision with the application of FL-CXR allows not only differentiation of alleles differing by 4 or 2 nucleotides but also those differing by 1 nucleotide (± 3SD). The high precision obtained for FL-CXR as opposed to GS500 might result from the 14 equally distributed fragments of DNA within the range 100–400 bps giving a calibration curve which enables a very precise definition of the size of the DNA fragments.

Stutter

Ascertaining the size range of stutters and the factors which influence formation allows the correct interpretation of DNA mixtures. In order to determine the proportion of stutter-type fractions, tests were carried out on over 80 samples amplified with optimal DNA concentrations (0.5–1.25 ng; 150-4500RFU). Only heterozygotes with alleles at least 8 bp apart were analysed. For the individual loci, stutter was measured in each system for the allele ranges D3S1385 14–18, VWA 14–19, FGA 16–26, D5S818 9–13, D13S317 8–12, D7S820 8–12, D8S1172 10–15, D21S11 28–33.2 and for D18S51 12–22. The level of stutter was calculated by dividing the height or area of stutter peak by the height or area of the main peak and no differences were observed (data not shown). In general it was found that an increase of the size of stutters was observed with increasing number of repeat sequences, which is consistent with Perkin Elmer observations (AmpF*l*STR Profiler Plus amplification manual) The biggest average size of stutter was observed for D18S51 (7.32%) and the smallest were for D7S820 (3.75%) and D13S317 (4.8%). For the rest of the loci average stutter values were between 5.7–6.9%. The biggest maximal values were observed for D18S51 (13.4%) and VWA (13.2%) and the smallest for D7S820 (6.7%) and D13S317 (8.4%). For the rest of the loci the maximal values were between 12.2–12.5% and none of the values exceeded 14%.

The influence of the quantity of DNA on the amplification of Profiler Plus loci

Various quantities of the DNA standard (0.09–25 ng) from the QuantiBlot kit were amplified and it was observed that with very high concentrations of DNA the quantity of fractions of type n and $n + 1$ increased significantly, appearing even for stutter fractions. Artefacts in the shape of additional peaks, not belonging to any of the Profiler Plus loci, were also observed. An additional peak 122 bp long (labelled JOE) was observed for DNA concentrations between 25–3.12 ng (Fig. 2, upper panel). With DNA concentration below 6.25 ng type n fractions were not observed. For DNA concentrations below 0.78ng complete amplification profile of all ten loci was not obtained but individual loci (mainly homozygotes and short alleles) could be identified even at DNA concentrations of 0.143 ng (RFU>150). An increase in DNA template creates not only the artefacts described above but also significant changes of the relative peak ratios between different loci. For instance the peak area ratio for

allele 16 of D18S51 locus and allele 12 of D7S820 locus changed from 6.5 : 1 for 25 ng DNA template to 1.5 : 1 for 1.56 ng DNA. Similar changes were observed for other loci (Fig. 2). The Profiler Plus multiplex kit is very sensitive to DNA overload leading to more artefacts and finally to difficulties with phenotyping. Precise DNA quantification before amplification seems to be a very important step.

Dependence of stutter size on quantity of DNA

In order to assess the effect of the quantity of DNA in the reaction mixture on stutter, 0.39–25 ng of the DNA standard from the QuantiBlot (PE) kit was amplified. A change was confirmed in the percentage of stutter in all loci of the Profiler Plus system depending on the DNA concentration. The lowest level of stutter were obtained in all loci using 0.78–1.56 ng DNA. The highest level of stutter in this range was observed for the D3S1358 locus (7.9–12%). The highest increase in stutter for all loci tested was observed with 6.25 ng DNA and the locus with most stutter was D3S1358 (15.08%). For some loci (D3S1358, VWA and D8S1179) the percentage stutter observed for 6.25 ng DNA was twice as high as for 0.78 ng DNA template. For the other loci (FGA, D21S11 and D18S51) we did not observe such big stutter changes as a function of DNA concentration. These data are different from those presented by Perkin Elmer (AmpFlSTR Profiler Plus amplification manual), where no significant changes with quantity of DNA input were observed. For DNA concentrations below 0.78 ng and above 12.5 ng stutter percentages were not calculated because of RFU values below 150 and above 4500, respectively.

References

1. Brinkmann B (1996) The STR approach. In: Carracedo A, Brinkmann B, Bär W (eds) Advances in forensic haemogenetics 6. Springer, Berlin Heidelberg New York, pp 41–51
2. Budowle B, Smerick JB, Keys KM, Moretti TR (1997) United States population data on the multiplex short tandem repeat loci – HUMTH01, TPOX, and CSF1PO – and the variable number tandem repeat locus D1S80. J Forensic Sci 42: 846–849
3. Furedi S, Budowle B, Woller J, Padar Z (1996) 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 Legal Med 109: 100–101
4. Greenhalgh MJ (1996) Casework experiences with a multiplex STR system. In: Carracedo A, Brinkmann B, Bär W (eds) Advances in forensic haemogenetics 6. Springer, Berlin Heidelberg New York, pp 249–251
5. Entrala C, Lorente M, Lorente JA, Alvares JC, Moretti T, Budowle B, Villanueva E (1998) Fluorescent multiplex analysis of nine STR loci: Spanish population data. Forensic Sci Int 98: 179–183
6. Pawlowski R, Maciejewska A, Paszkowska R, Welz A (1997) Frequencies of the five short tandem repeat systems (STR) in a population from North Poland. Int J Legal Med 110: 10–13
7. Mills KA, Even D, Murray JC (1992) Tetranucleotide repeat polymorphism at the human alpha fibrinogen locus (FGA). Hum Mol Genet 1: 79
8. Walsh PS, Fildes NJ, Reynolds R (1996) Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acids Res 24: 2807–2812
9. Pawlowski R, Branicki W, Kupiec T (1999) Y-chromosomal polymorphic loci DYS19, DYS390, DYS393 in a population sample from northern Poland. Electrophoresis 20: 1702–1706
10. Weir BS (1992) Independence of VNTR alleles defined by fixed bins. Genetics 130: 873–878
11. Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48: 361–372
12. Pawlowski R (1999) HUMFIBRA allele distribution in Northern Poland using capillary electrophoresis. Int J Legal Med 112: 139–141
13. Barber MD, McKeown BJ, Parkin BH (1996) Structural variation in the alleles of a short tandem repeat system at the human alpha fibrinogen locus. Int J Legal Med 108: 180–185
14. Wallin JM, Buoncristiani MR, Lazaruk KD, Fildes N, Holt CL, Walsh PS (1998) TWGDAM validation of the AmpFlSTR blue PCR amplification kit for forensic casework analysis. J Forensic Sci 43: 854–870