Int J Legal Med (2007) 121: 78–81
DOI 10.1007/s00414-005-0060-z

SHORT COMMUNICATION

Dorit Becker · Dorit Vogelsang · Werner Brabetz

Population data on the seven short tandem repeat loci D4S2366, D6S474, D14S608, D19S246, D20S480, D21S226 and D22S689 in a German population

Received: 6 August 2005 / Accepted: 13 October 2005 / Published online: 19 November 2005

Abstract Allele frequencies for the autosomal tetranucleotide short tandem repeat loci D4S2366, D6S474, D14S608, D19S246, D20S480, D21S226 and D22S689 were investigated in a sample of 189 unrelated German individuals using a multiplex polymerase chain reaction approach. The loci showed no significant deviations from Hardy-Weinberg equilibrium except for D14S608. All genotyped alleles were cloned and sequenced, and an allelic nomenclature consistent with the ISFH recommendations was defined.

Keywords DNA typing · Short tandem repeat · German population · Multiplex · Polymerase chain reaction

Introduction

Polymorphic short tandem repeats (STR) are powerful genetic markers to distinguish individuals and still preferred for forensic applications and paternity testing [1]. Evaluation of new STR markers or primer redesign for conventional STR systems is a useful tool to obtain additional information and to complement conventional STR analysis [2, 3]. There is a growing interest in polymorphic markers unlinked to the current CODIS loci [4]. Multiplex polymerase chain reaction (PCR) based on tetranucleotide STRs allows fast and reliable genotyping. Here we report allele frequency data and the variable repeat sequence motifs of seven autosomal loci which were analysed in a German population using two multiplex PCR systems.

Materials and methods

Population samples were obtained from 189 unrelated German Caucasians from the area around Dresden. DNA from cell line 9,947 A was purchased from Coriell Cell Repositories (Camden, NJ) as a control. DNA was extracted from buccal swabs using the NucleoSpin Tissue kit (Macherey&Nagel, Düren, Germany). PCR was carried out with a GeneAmp PCR System 9,700 Thermal Cycler (Applied Biosystems, Foster City, CA) in a 25-µl reaction volume containing 1–2 ng DNA, 50–150 nM primer, 200 µM of each dNTP, 3 mM $MgCl_2$ and 1 U JumpStart *Taq* DNA polymerase (Sigma, Taufkirchen, Germany). One primer for each locus was labelled at the 5′-end with 6-carboxyfluorescein (FAM), 4,7,2′,4′,5′,7′-hexachloro-6-carboxy-fluorescein (HEX) or NED (proprietary to Applied Biosystems) for capillary gel electrophoresis. The PCR protocol consisted of an initial denaturation step at 94°C for 4 min, 94°C for 30 s, 58°C for 90 s, 72°C for 75 s over 30 cycles and a terminal step at 68°C for 60 min. The PCR products were analysed by capillary electrophoresis using an ABI Prism 310 Genetic Analyzer with polymer POP-4 (Applied Biosystems). PCR products obtained with unlabelled primers were cloned into *Escherichia coli* TOP10F′ (Invitrogen, Karlsruhe, Germany) using the plasmid pCR2.1-TOPO and the TOPO TA Cloning kit (Invitrogen). Sequencing of both strands of the DNA inserts was performed using the Big Dye Terminator Cycle Sequencing kit (Applied Biosystems). The STR genotypes of the population were statistically analysed by calculating the observed heterozygosity (HET), the polymorphic information content (PIC) [5], the power of discrimination (PD) [6], the mean exclusion chance (MEC) [7] and the deviation from the Hardy-Weinberg equilibrium (HWE) based on χ^2 test [8].

Results and discussion

Allele frequencies of unrelated individuals from the area around Dresden were investigated in two screening multiplexes. A triplex assay for D4S2366, D6S474 and D19S246 was developed (Fig. 1a), whereas alleles for D14S608, D20S480, D21S226 and D22S689 were amplified in a quadruplex assay (Fig. 1b). DNA sequences and fluorescence dye labels for the PCR primers are given in Table 1.

D. Becker (✉) · D. Vogelsang · W. Brabetz
Biotype AG, Moritzburger Weg 67 D,
01109 Dresden, Germany
e-mail: d.becker@biotype.de
Tel.: +49-351-8838128
Fax: +49-351-8838435

All different alleles found for the loci were cloned and sequenced to establish an allelic nomenclature based on the variable tandem repeat motifs and classified according to the recommendations of the International Society for Forensic Haemogenetics (ISFH) [9]. Repeats could be divided into simple, compound and complex repeat sequences (Table 2) [9]. There was no evidence for incomplete repeat motifs. Conserved TGA and ATC motifs for D6S474 and D20S480 were not considered for the allelic nomenclature. D19S246 showed a complex repeat sequence with variable and non-variable tetranucleotide repeats as well as a conserved AT motif. If not stated otherwise, at least two clones per allele were sequenced (Table 2).

The allele frequencies for all loci were determined, and the population data were statistically analysed. The PD values of the STR loci ranged from 0.78 for D21S226 to 0.951 for D19S246 (Table 3). A comparison of our German population data with allele frequency data for Chinese populations revealed significant differences with regard to the observed alleles and allele frequencies of D4S2366 [10], D6S474 [11] and D22S689 [12]. In contrast to our data Choi et al. [13] found only eight alleles for D14S608 in a Korean population. In the German population, we typed for D14S608 the additional alleles 4, 6 and 7, the latter of which revealed even the highest allele frequency of 0.288 (see Table 3). Furthermore, we found significant deviation from HWE for this locus also after Bonferroni correction. Deviations from HWE can point either to mistyping genotypes, sampling bias or natural selection. The primer pair, which was used in PCR for D14S608, gave amplicons shorter than 230 bp, and all peaks could be clearly assigned in capillary electrophoresis. So far, we can exclude mistyping errors. Due to high degrees of HET (Table 3) and high discrimination power, most of the seven STR markers except for D21S226 can be efficiently used for kinship analyses or forensic purposes in German populations. All markers except for D21S226 and D19S246 are located on different chromosomes than or separated by at least 50 cM from STR loci, which are currently used in STR databases. Consequently, these markers are probably unlinked to common STRs and useful to obtain additional information in STR analysis for forensic casework and paternity testing. PCR with the established screening multiplexes allows fast and sensitive genotyping for these seven autosomal markers. A detection limit of at least 100 pg DNA could be achieved for both multiplex assays.

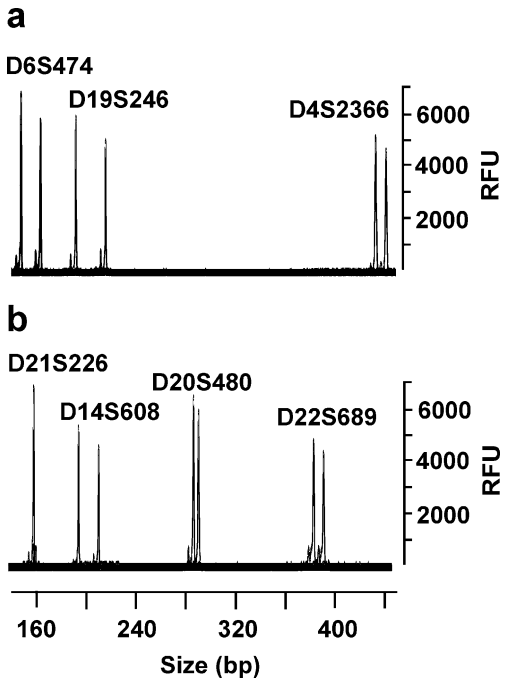

Fig. 1 PCR genotyping profiles of 1 ng DNA from cell line 9947A. The STRs were amplified using a triplex PCR (**a**) and a quadruplex PCR (**b**). Products were analysed on an ABI Prism 310 Genetic Analyzer. Dye labels corresponding to the electropherograms are shown in Table 1. The genotypes according the allelic nomenclature are as follows: D6S474 (13;17), D19S246 (30;36), D4S2366 (11;13), D21S226 (10;10), D14S608 (7;11), D20S480 (15;16), and D22S689 (16;18). (*RFU* relative fluorescence unit, *bp* base pair)

Table 1 Loci, chromosomal location, GenBank accession number and primer sequences

Locus	Chromosomal location	Accession number	Primer sequence
D4S2366	4p16–15.2	G08339	fw: NED-TGCTTGCTT TGTTTTGTGGTA rv: GCTGATGGGAAC GTAAAATGA
D6S474	6q21–22	G08540	fw: FAM-TGTACAAAAG CCTATTTAGTCAGG rv: TCATGTGAGCCAA TTCCTCT
D14S608	14q11.1–11.2	G09052	fw: NED-TAAAGGTTTAT CCATGCTGTAGC rv: ACGTGGTACAGGTA GATAAATGG
D19S246	19q13.3–q13.4	L13118	fw: HEX-AGAGTGAGA TTCCACCTTTC rv: GAAACACATCATT TACCCAC
D20S480	20q	G08049	fw: FAM-GTGGTGAACA CCAATAAATGG rv: AAGCAAATAAAA CCAATAAACTCG
D21S226	21q22.1	M93147	fw: HEX-GGAAACCAC TCTAAGACATA rv: AGCTAAATGTCTG TAGTTATT
D22S689	22q11.2–22q12.1	G08087	fw: NED-CCCATCAC CACCAGTCCTACT rv: CAGGAAGTCAA GGCTGTAGGG

fw forward, *rv* reverse (orientation according to sequence accession number)

Table 2 Sequencing data of different alleles for all loci

Locus	Allele	Repeat structure
Simple repeats		
D14S608	4–15	$(TCTA)_{4–15}$
Simple repeats with sequence variants		
D21S226	7,8,11	$(TAAA)_{7,8,11}$
D21S226	9/10	$(TAAA)_8(TAGA)/(TAAA)_9CAAA$
Compound repeat sequences		
D4S2366	9–15	$(ATAG)_{5–11}(ATTG)_{2–3}(ATAG)_{1–2}$
D6S474	13–18	$(TAGA)_5TGA(TAGA)_{8–13}$
D20S480*	12–19	$(TATC)_{10–17}ATC(TATC)_2$
D22S689*	12–20	$(TAGA)_{8–16}(CAGA)_3(TAGA)_{0–1}$
Complex repeat sequences		
D19S246	27	$(ATAG)_4ATAC(ATAG)_2TTAG(ATAG)_2(ATAC)_7$ $ACAGATAC(ATAG)_3ACAG(ATAG)_2AT(TAGA)_2$
	30	$(ATAG)_4ATAC(ATAG)_7(ATAC)_8ACAGATAC(ATAG)_3$ $ACAG(ATAG)_2AT(TAGA)_2$
	32	$(ATAG)_4ATAC(ATAG)_6(ATAC)_{11}ACAGATAC(ATAG)_3$ $ACAG(ATAG)_2AT(TAGA)_2$
	33	$(ATAG)_4ATAC(ATAG)_7(ATAC)_{11}ACAGATAC(ATAG)_3$ $ACAG(ATAG)_2AT(TAGA)_2$
	34	$(ATAG)_4ATAC(ATAG)_8(ATAC)_{11}ACAGATAC(ATAG)_3$ $ACAG(ATAG)_2AT(TAGA)_2$
	35	$(ATAG)_4ATAC(ATAG)_{11}(ATAC)_8ACAGATAC(ATAG)_3$ $ACAG(ATAG)_2AT(TAGA)_3$
	36	$(ATAG)_4ATAC(ATAG)_{12}(ATAC)_8ACAGATAC(ATAG)_3$ $ACAG(ATAG)_2AT(TAGA)_3$
	37	$(ATAG)_4ATAC(ATAG)_{11}ACAG(ATAC)_9ACAGATAC(ATAG)_3$ $ACAG(ATAG)_2AT(TAGA)_3$
	38	$(ATAG)_4ATAC(ATAG)_{13}(ATAC)_9ACAGATAC(ATAG)_3$ $ACAG(ATAG)_2AT(TAGA)_3$
	39	$(ATAG)_4ATAC(ATAG)_{13}(ATAC)_{10}ACAGATAC(ATAG)_3$ $ACAG(ATAG)_2AT(TAGA)_3$

Generally, at least two clones per allele were sequenced except for the loci indicated by an asterisk (*)

Table 3 Allele frequencies of 189 unrelated German Caucasians

Allele	D4S2366	D6S474	D14S608	D19S246	D20S480	D21S226	D22S689
4			0.013				
5							
6			0.004				
7			0.288			0.005	
8			0.011			0.201	
9	0.347		0.074			0.241	
10	0.179		0.177			0.550	
11	0.074		0.198			0.003	
12	0.147		0.122		0.005		0.005
13	0.168	0.255	0.069		0.032		0.019
14	0.074	0.211	0.003		0.122		0.034
15	0.011	0.153	0.005		0.220		0.082
16		0.279			0.307		0.360
17		0.100			0.257		0.357
18		0.003			0.056		0.111
19					0.003		0.024
20							0.008
21							
22							
23							

Table 3 (continued)

Allele	D4S2366	D6S474	D14S608	D19S246	D20S480	D21S226	D22S689
24							
25							
26							
27				0.249			
28				0.003			
29							
30				0.071			
31							
32				0.021			
33				0.153			
34				0.077			
35				0.034			
36				0.185			
37				0.146			
38				0.056			
39				0.005			
PIC	0.760	0.740	0.800	0.820	0.740	0.530	0.680
HET	0.795	0.737	0.820	0.857	0.815	0.492	0.735
PD	0.919	0.918	0.938	0.951	0.901	0.780	0.877
MEC	0.758	0.745	0.758	0.825	0.739	0.535	0.678
HWE (p value)	0.867	0.997	0.000	0.997	0.986	0.229	0.979

PIC Polymorphic information content, *HET* heterozygosity, *PD* power of discrimination, *MEC* mean exclusion chance, *HWE* Hardy-Weinberg equilibrium

Acknowledgements The authors thank K. Thürasch for excellent technical assistance. Part of this work was financially supported by a grant of the German Federal Ministry of Education and Research (BMBF, 03i4008).

References

1. Gill P, Werrett DJ, Budowle B, Guerrieri R (2004) 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 (SWGDAM). Sci Justice 44:51–53
2. Grubwieser P, Mühlmann R, Berger B, Niederstätter H, Pavlic M, Parson W (2005) A new "miniSTR-multiplex" displaying reduced amlicon lengths for the analysis of degraded DNA. Int J Legal Med (in press)
3. Asamura H, Uchida R, Takayanagi K, Ota M, Fukushima H (2005) Allele frequencies of the six miniSTRloci in a population from Japan. Int J Legal Med (in press)
4. Coble MD, Butler JM (2005) Characterization of new MiniSTR loci to aid analysis of degraded DNA. J Forensic Sci 50:43–53
5. Botstein D, White RL, Skolnich M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331
6. Jones DA (1972) Blood samples: probability of discrimination. J Forensic Sci Soc 12:355–359
7. Krüger J, Fuhrmann W, Lichte KH, Steffens C (1968) Zur Verwendung der sauren Erythrocytenphosphatase bei der Vaterschaftsbegutachtung. Dtsch Z Gesamte Gerichtl Med 64: 127–146
8. Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372
9. Bär W, Brinkmann B, Budowle B et al (1997) DNA recommendations. Further report of the DNA commission of the ISFH regarding the use of short tandem repeat systems. Forensic Sci Int 87:179–184
10. Zhang WJ, Li YB, Zhou XP et al (2004) Population genetics of two STR loci D4S2366 and D6S1281 in a Chinese population. J Forensic Sci 49:183
11. Zhang WJ, Wu J, Jia ZJ et al (2004) Population genetics of two STR loci D2S1396 and D6S474 in a Chinese population. J Forensic Sci 49:415
12. Jia ZJ, Wu J, Hou YP et al (2004) Distributions of allelic frequencies and haplotypes of two new STR loci in a Chinese Han population. J Forensic Sci 49:409
13. Choi M, Kim JH, Lee DH, Lee SH, Rho HM (2000) Frequency data on four tetrameric STR loci D18S1270, D14S608, D16S3253 and D21S1437 in a Korean population. Int J Legal Med 113:179–180