Int J Legal Med (2010) 124:83–87
DOI 10.1007/s00414-009-0328-9

TECHNICAL NOTE

Validation of six closely linked STRs located in the chromosome X centromere region

Jeanett Edelmann · Sandra Hering · Christa Augustin · Stefan Kalis · Reinhard Szibor

Received: 31 October 2008 / Accepted: 27 January 2009 / Published online: 20 February 2009

Abstract We propose that clusters of closely linked markers, which segregate as stable haplotypes, provide a high potential to solve complex kinship cases. It is known that the X-chromosomal centromere region shows an extremely low degree of recombination. Hence, we focused our interest on the region between 56 and 64 Mb distant from the Xp telomere and considered 6 STRs which are now registered in the Genome Data Base as DXS10161, DXS10159, DXS10162, DXS10163, DXS10164, and DXS10165. All of these markers show a tetranucleotide or pentanucleotide structure and exhibit high or medium polymorphic information content. As a peculiarity, DXS10163 is a combination of a pentanucleotide STR and an 18 bp INDEL polymorphism. We report here the primer sequences, the repeat structures, the allele distributions and parameters of forensic interest for a German population sample.

Keywords X chromosome · Centromere · Haplotype · Short tandem repeats · Indel polymorphism

Electronic supplementary material The online version of this article (doi:10.1007/s00414-009-0328-9) contains supplementary material, which is available to authorized users.

J. Edelmann
Institut für Rechtsmedizin, Universität Leipzig,
Johannisallee 28,
04103 Leipzig, Germany

S. Hering
Institut für Rechtsmedizin, Technische Universität Dresden,
Fetscherstrasse 74,
01307 Dresden, Germany

C. Augustin
Institut für Rechtsmedizin,
Universitätsklinikum Hamburg-Eppendorf,
Butenfeld 34,
22529 Hamburg, Germany

S. Kalis
Institut für Blutgruppenserologie und Genetik,
Dunantstrasse 16a,
79110 Freiburg, Germany

R. Szibor (✉)
Institut für Rechtsmedizin,
Otto-von-Guericke-Universität Magdeburg,
Leipziger Straße 44,
39120 Magdeburg, Germany
e-mail: reinhard.szibor@med.ovgu.de

Introduction

In recent years, numerous STRs spread over the whole X chromosome (ChrX) have been detected and established for forensic purposes [2, 3, 7, 13–15]. Typing of male DNA enables direct recognition of the ChrX marker haplotypes. Considering pedigrees often provides a way to deduce haplotypes also for females in an indirect manner. For example, ChrX typing of a mother and her son reveals both the male and the maternal haplotypes. Whereas most of the STRs used in the forensic field exhibit allele frequencies in the range of 0.05–0.40, the majority of haplotypes comprising 2–3 STRs exhibit frequencies between 0.001 and 0.02. Hence, in kinship testing, when two (or more) persons share such a rare STR cluster haplotype, there is a strong indication of kinship. Therefore, our group is attracted to the idea to use tightly linked STR clusters, which segregate as stable haplotypes, in kinship testing in complex cases. As has been demonstrated [6, 14, 15], this approach is indeed successful and can be used for solving complex deficiency cases. The usefulness of such STR clusters, among other things, depends on their stability against recombination. Several ChrX-linked STR clusters

Table 1 Sequences and position of the STR primers used in this study

	5′ labeling	Sequence (5′–3′ direction)	Primer 5′ position
DXS10161 F	HEX	TCT GTA TAA GCA TCA TGG GAC TTC	55 999 212
DXS10161 R		CAA AAT GTC TTT GAA TTA GTG AAG G	55 999 441
DXS10161 F-Seq		GCA CTG ACC AGT TTC AAT AG	55 999 151
DXS10159 F	HEX	CAT TGC ACT CCA ACC TGG AT	56 766 054
DXS10159 R		TGA CAT GTT GCT GGA TTT GG	56 766 244
DXS10159 F-Seq		GCA GTA GAA TCA CTT GAA CC	56 765 999
DXS10162 F	FAM	ATT TTG TAA GTT GAT TCC ACC C	61 837 629
DXS10162 R		GGT TAT CCC CAG TTA AAA GC	61 837 941
DXS10162 F-Seq		TCA CCA GGC AGT TTC CCA G	61 837 480
DXS10163 F		GGG CAA TGT AGC CAA ACT CT	62 002 675
DXS10163 R-S	HEX	CAG AGA AAA AAA CAC AAC TTC C	62 002 819
DXS10163 R-L	FAM	TCC TTG AAA CAC ACA ACT TCC	No matches
DXS10163 F-Seq		TTG GGA GAT AGG CTG GGT G	62 002 574
DXS10164 F	FAM	TGA ACG GAA TGT ACT TTT CCA A	62 161 308
DXS10164 R		GGC AGT ATT AAT CAA AAC AGA ATG G	62 161 515
DXS10164 F-Seq		TCT TTT GCA TAT GGG TTT TCT G	62 161 267
DXS10164 R-Seq		AAC AAA GCT GGA GGC ATC AC	62 161 562
DXS10165 F	HEX	ACT TTT GCC CAA AGG TAG GG	62 994 016
DXS10165 R		CGT CAA AGT GTG CAT GTG TG	62 994 172
DXS10165 F-Seq		CCT TGA CAA GCA AGT GTG GA	62 993 880
DXS10165 R-Seq		AGG CTG GGT TGT AGT GAT GG	62 994 295

have been described during recent years. Information on recombination activity of the chromosomal regions has been known for more than 10 years since Nagaraja et al. [12] published an X chromosome recombination map at 75-kb STS resolution. Information on recombination activity between mapped and unmapped markers can be exactly retrieved from the second generation combined linkage-physical map of the human genome of chromosomes [10]. With respect to the susceptibility to recombination, the established clusters map to quite different regions. However, none of them can be regarded as intrinsically free of recombination. STR clusters for usage in ChrX haplotyping are described for the regions Xp22 [7], Xp11.23 [1], Xq12 [6], Xq21 [15], Xq22 [4], Xq26.2 (Rodig et al., unpublished results) and Xq28 [5].

It is well known that recombination occurs extremely rarely around the centromere, and a further low recombination region lies at Xq13.3–Xq21.3 about 76–84 Mb distant from the Xp telomere [8, 9]. The study presented here establishes six ChrX microsatellites which from the theoretical point of view were the most promising STRs in the contigs NT_011669 and NT_011630. All of these markers are located in the centromere region between 56.0 Mb and 64.0 Mb from Xptel. The newly described STRs were registered in the GDB as DXS10161, DXS10159, DXS10162, DXS10163, DXS10164 and DXS10165 (Table 1). DXS10163 represents a combination of a pentanucleotide STR and an INDEL polymorphism. The latter class of markers are diallelic polymorphisms which are systematically described by Mills et al. [11]. These markers seem to be underestimated for use

Table 2 Sequence structure of the ChrX marker alleles

Marker	Sequence structure	Alleles
DXS10161	Pr $I_{(24)}$ -N_{40}-$(TATC)_2$-TATT-$(TATC)_{7\text{-}13}$-$(ATCT)_5$-ATG-$(TATC)_5$-$(ATCT)_3$-N_{34}-Pr $II_{(25)}$	22–28
DXS10159	Pr $I_{(20)}$ -N_{27}-$(AAAG)_4$-$(AGAA)_2$-AG-$(AGAA)_{5\text{-}6}$-$(AAAG)_{10\text{-}17}$-N_{11}-Pr $II_{(20)}$	21–29
DXS10162	Pr $I_{(21)}$ -N_{39}-(TCTT)-TT-$(TCTT)_3$-T-$(TCTT)_{10\text{-}18}$-N_9-Pr $II_{(22)}$	14–23
DXS10163	*S-type*: Pr $I_{(20)}$-N_7-$(AAATA)_{14\text{-}19}$-A-$(AAATA)_2$- N_{16}-N_{89}-Pr $II_{(25)}$	S16–S21
	L-type: Pr $I_{(20)}$-N_7-$(AAATA)_{11\text{-}20}$-A-$(AAATA)_2$-N_{16}-**N_{18}**-N_{89}-Pr $II_{(25)}$	L13–L23
DXS10164	Pr $I_{(22)}$ -N_{64}-$(ATTCT)_{8\text{-}13}$-N_{42}-Pr $II_{(25)}$	8–13
DXS10165	Pr $I_{(20)}$ -N_{13}-$(AAGA)_6$-AAA-$(AAAG)_{11\text{-}18}$-N_{21}-Pr $II_{(20)}$	17–24

Table 3 Allele frequencies in a German population

Allele	DXS10161 n=541	DXS10159 n=538	DXS10162 n=575	DXS10163 n=574	DXS10164 n=667	DXS10165 n=668
8					0.007	
9					0.014	
10					0.692	
11					0.227	
12					0.055	
13					0.006	
14			0.0014			
15			0.0056			
16			0.0677			
17			0.1748			0.003
18			0.4147			0.033
19			0.2271			0.184
20			0.0733			0.366
21		0.001	0.0310			0.283
22	0.003	0.006	0.0028			0.105
23	0.047	0.070	0.0014			0.024
24	0.205	0.241				0.002
25	0.422	0.236				
26	0.309	0.296				
27	0.013	0.106				
28	0.002	0.036				
29		0.006				
S16				0.0028		
S17				0.0028		
S18				0.0071		
S19				0.0042		
S21				0.0042		
L13				0.0127		
L14				0.0056		
L15				0.0381		
L16				0.0325		
L17				0.1582		
L18				0.2684		
L19				0.2867		
L20				0.1299		
L21				0.0240		
L22				0.0028		
L23				0.0014		

Table 4 Statistical parameters

	DXS10161	DXS10159	DXS10162	DXS10163	DXS10164	DXS10165
PIC	0.628	0.745	0.687	0.766	0.421	0.693
MEC	0.628	0.745	0.687	0.766	0.421	0.693
PD male	0.687	0.779	0.727	0.794	0.471	0.735
PD female	0.844	0.917	0.885	0.929	0.670	0.885
Het obs.	0.746	0.823	0.726	0.794	0.480	0.752
Het exp.	0.689	0.782	0.730	0.797	0.472	0.737
HWE (p)	0.719	0.665	0.391	0.860	0.297	0.307

PIC Polymorphism information content, *PD* power of discrimination, *MEC* mean exclusion chance, *Het obs.* observed heterozygosity, *Het exp.* expected heterozygosity

Table 5 STR features and control DNA patterns

Alleles	DXS10161 *n*=7	DXS10159 *n*=9	DXS10162 *n*=10	DXS10163 *n*=16	DXS10164 *n*=6	DXS10165 *n*=8
Length (bp)	218–242	164–196	150–186	121–171	193–218	145–173
Nomenclature	22–28	21–29	14–23	S16–S21 L13–L23	8–13	17–24
K562	25	27	18	L17	10	21
9947A	25/26	24/25	19/19	L18/L19	10/10	20/21

in forensic science and more attention should be paid to them.

Materials and methods

In this study, we investigated a German population sample of unrelated individuals for the six markers and the number of X chromosomes investigated was 538 at least. These samples were drawn from cases of routine kinship testing concerning female children. The people involved gave their consent for the investigation of STRs of forensic significance. Additionally, we typed the reference cell line DNA samples K562 and 9947A. For sequencing, we selected two to five amplicons of male DNA specimens for every fragment length of each marker. ChrX typing of 354 male DNA specimens directly provided the haplotype data comprising all of the six STR loci. We investigated 109 female meioses with regard to the possible occurrence of crossing-over.

To check for deviation from the Hardy–Weinberg equilibrium (HWE), we determined the genotypes of at least 202 females.

All primer data are shown in Table 1. Primer sequences were established by checking the contigs NT_011669 and NT_011630 (Genome Systems Human BAC Library) and the exact map positions were retrieved using the UCSC in silico PCR tool (http://www.genome.ucsc.edu/).

A special three-primer system was introduced for the amplification of the combined STR-INDEL polymorphism DXS10163. One forward primer was combined with two differently labeled reverse primers: a HEX-labeled primer matching the INDEL deletion allele and a FAM-labeled primer is appropriate for the INDEL insertion allele. Thus, the STR alleles linked to the INDEL short allele occur in the green channel and all STR alleles combined with the INDEL long allele appear in blue.

PCR amplification was carried out in a 15-µl PCR reaction volume containing approximately 0.1–1 ng DNA, 200 µM of each dNTP, 1.5 mM $MgCl_2$, 0.5 µM of each primer, 1 U Taq polymerase (AmpliTaq-Gold, Applied Biosystems, Foster City, CA) and 1/10 volume of the appropriate Taq polymerase buffer. The following PCR cycle protocol was used: 95°C–10 min soak; 94°C–45 s, 59°C–1 min, 72°C–1 min, 30 cycles, 72°C–10 min final extension in a T3 (Biometra, Göttingen, Germany).

The same conditions were used to generate amplicons for the cycle sequencing procedure.

The resulting PCR products were analyzed in the denaturing polymer POP4 on ABI Prism®310 and ABI Prism®3130 Genetic Analyzers (Applied Biosystems) and the Genotyper and Genemapper software (Applied Biosystems) were used. Amplicon sizing was based on the 550 size standard (Biotype AG, Dresden, Germany).

To analyze the variability of the microsatellite repeat structures and the adjacent regions, we produced appropriate amplicons of male DNA samples using primers given in Table 1 and performed the cycle sequencing procedure. The Big Dye Cycle Sequencing kit (Applied Biosystems) was used as recommended by the manufacturer.

HWE analysis was done using the exact test. Parameters of forensic interest were calculated using formulas as reviewed earlier [14].

Table 6 Distribution of the STR haplotypes DXS10161-DXS10159-DXS10162-DXS10163-DXS10164-DXS10165 in a sample of 354 X chromosomes

	German sample (*n*)	Proportion of the sample (%)
Number of investigated chromosomes	354	–
Number of different observed haplotypes	299	–
Most frequent haplotypes (24-26-18-L19-10-19 and 24-26-18-L20-10-20; *f*~1.13%)	8	2.26
Number of haplotypes with frequencies *f*~0.8%	30	8.47
Number of haplotypes with frequencies *f*~0.5%	58	16.38
Number of unique haplotypes (f~0.28% each)	258	72.88

Results and discussion

All markers showed robust amplification properties and appear suitable for forensic purposes. Allele sequence structure and nomenclature of the six markers is shown in Tables 2. Five markers exhibited a regular repeat structure. The repeat flanking regions analyzed in this study are in accordance with the GenBank sequences, and no SNPs were detected. DXS10163 is a combined marker consisting of a pentanucleotide STR and an INDEL polymorphism. The INDEL element exhibiting 18 nucleotides (gtttcaaggaatttaccc) begins 16 bp downstream of the repeat region. This situation results in two series of haplotypes. The (long) L-type alleles contain the 18 bp INDEL element and a variable number of STR repeats. The (short) S-type represents the variable STR alleles in combination with the deletion of the INDEL element. As described, a three-primer system using allele specific reverse primers enables a joint analysis of the STR and the INDEL polymorphism revealing the DXS10163 haplotypes.

Allele frequencies and statistical parameters of forensic interest are given in Tables 3 and 4. Of the six STRs, five showed high PIC values in the range of 0.628–0.745, solely DXS10164 exhibited a low PIC of 0.421. Nevertheless, this STR may contribute sufficiently to the individualisation of the ChrX in the context of haplotyping the centromere region.

A summary of the features of the STRs and of the cell line DNA typing patterns, which can be used as intra-laboratory and interlaboratory standards, is shown in Table 5.

Table 6 presents a review of the distribution and the frequencies of haplotypes from our sample of 354 chromosomes of which 72.88% are unique with frequencies lower than 0.003. All haplotypes are given in detail as ESM 1.

In our first family study, we checked 109 female meioses. However, only 61 meioses were informative in DXS10161 and DXS10165, which are the both outer markers of the investigated centromere region. No recombinations were found.

Whether or not this centromere STR cluster is truly free of recombination will become clear in future investigations involving several hundred meioses. With regard to the assumed lack of recombination, this X-chromosomal cluster might provide a counterpart to Y-chromosomal STR haplotypes and may complement the current Y-chromosomal and mitochondrial studies for population genetics and human migration.

References

1. Augustin C, Cichy R, Hering S, Edelmann J et al (2006) Forensic evaluation of three closely linked STR markers in a 13 kb region at Xp11.23. Int Congr 1239:311–314
2. Becker D, Rodig H, Augustin C et al (2008) Population genetic evaluation of eight X-chromosomal short tandem repeat loci using Mentype Argus X-8 PCR amplification kit. Forensic Sci Int Genet 2:69–74
3. Edelmann J, Hering S, Michael M et al (2001) 16 X-chromosome STR loci frequency data from a German population. Forensic Sci Int 124:215–218
4. Edelmann J, Hering S, Kuhlisch E, Szibor R (2002) Validation of the STR DXS7424 and the linkage situation on the X-chromosome. Forensic Sci Int 125:217–222
5. Edelmann J, Hering S, Augustin C, Szibor R (2008) Characterisation of the STR markers DXS10146, DXS10134 and DXS10147 located within a 79.1 kb region at Xq28. Forensic Sci Int Genet 2:41–46
6. Hering S, Augustin C, Edelmann J et al (2006) DXS10079, DXS10074 and DXS10075 are STRs located within a 280-kb region of Xq12 and provide stable haplotypes useful for complex kinship cases. Int J Legal Med 120:337–345
7. Hundertmark T, Hering S, Edelmann J et al (2008) The STR cluster DXS10148-DXS8378-DXS10135 provides a powerful tool for X-chromosomal haplotyping at Xp22. Int J Legal Med 122:489–492
8. Laan M, Wiebe V, Khusnutdinova E et al (2005) X-chromosome as a marker for population history: linkage disequilibrium and haplotype study in Eurasian populations. Eur J Hum Genet 13:452–462
9. Latini V, Sole G, Doratiotto S et al (2004) Genetic isolates in Corsica (France): linkage disequilibrium extension analysis on the Xq13 region. Eur J Hum Genet 12:613–619
10. Matise TC, Chen F, Chen W et al (2007) A second-generation combined linkage physical map of the human genome. Genome Res 17:1783–1786
11. Mills RE, Luttig CT, Larkins CE et al (2006) An initial map of insertion and deletion (INDEL) variation in the human genome. Genome Res 16:1182–1190
12. Nagaraja R, MacMillan S, Kere J et al (1997) X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. Genome Res 7:210–222
13. Szibor R (2007) X-chromosomal markers: past, present and future. Forensic Sci Int Genet 1:93–99
14. Szibor R, Krawczak M, Hering S et al (2003) Use of X-linked markers for forensic purposes. Int J Legal Med 117:67–74
15. Szibor R, Hering S, Kuhlisch E et al (2005) Haplotyping of STR cluster DXS6801-DXS6809-DXS6789 on Xq21 provides a powerful tool for kinship testing. Int J Legal Med 119:363–369