EXCLI Journal 2007;6:177-182 – ISSN 1611-2156
Received: December 14, 2006, accepted: June 21, 2007, published: June 26, 2007

Original article:

DXS10079, DXS10074 and DXS10075: new alleles and SNP occurrence

S. Hering[a,*], H. Rodig[b], C. Augustin[c], J. Edelmann[d], M. Heidel[a], J. Dressler[a], R. Szibor[e]

[a]Institut für Rechtsmedizin, Technische Universität Dresden, Fetscherstraße 74, 01307 Dresden, Germany, Fax: +49-351-4584325; E-mail: Sandra.Hering@tu-dresden.de (*corresponding author); [b]Biotype AG, Dresden; [c]Institut für Rechtsmedizin, Universitätsklinikum Hamburg-Eppendorf; [d]Institut für Rechtsmedizin, Universität Leipzig; [e]Institut für Rechtsmedizin, Otto-von-Guericke-Universität Magdeburg, Germany

ABSTRACT

The number of established X-chromosomal STR markers suitable for forensic usage has risen continuously during recent years. The observation of X-chromosomal transmission lines can significantly contribute to the solving of complex kinship cases. The highly polymorphic tetranucleotide markers DXS10079, DXS10074 and DXS10075 are located within a 280 kb region at Xq12 and provide stable haplotypes. Most of these haplotypes occur at low frequencies in the German population, which would qualify this marker cluster as a useful tool in pedigree analysis. For routine use it is necessary to investigate the allele structure and check for variations in the repeat flanking region in samples of different ethnic populations. The information on SNP occurrence may help to minimise pitfalls caused by partial primer mismatching. We sequenced a variety of samples from Germans, Asians and Africans with respect to different STR alleles. For all three marker systems SNPs were detected in the repeat flanking regions. Some alleles found in the marker systems DXS10074 and DXS10075 exhibited typical repeat structures and SNP patterns found only in Africans and differing from Germans and Asians. The highest SNP diversity for DXS10079 was present in samples of all three ethnic groups. Further population data are needed to confirm this observation.

Key words: X chromosome; STR cluster; Xq12; SNP

INTRODUCTION

Typing of X-chromosomal STRs can significantly contribute to the solving of complex kinship cases (Bini et al., 2005; Edelmann & Szibor, 2001; Shin et al., 2005; Szibor et al., 2003, 2005; Tabbada et al., 2005; Wiegand et al., 2003). Recently, we introduced the three X-chromosomal markers DXS10079, DXS10074 and DXS10075 as an STR cluster located in a 280-kb region of Xq12 (Hering et al., 2005). Inheritance of these closely linked markers as stable haplotypes was proven by a study with three generation pedigrees (Hering et al., 2006). About 74 % of haplotypes of this newly established cluster occur at a frequency of less than 2 % in the German population.

Before the X-chromosomal STRs can be routinely used in forensic practice, it is necessary to investigate the allele structure in different populations. Sometimes, pitfalls can occur when null alleles are caused by variability in the primer binding regions (Gusmao et al., 1996). For this reason we checked the sequencing results for SNPs in the repeat flanking region. We performed a data base search including *Human SNPView* (http://www.ensembl.org, v36) and compared the results with our own sequencing results from DNA samples of Germans and some non-Caucasoid populations. This paper presents additional sequence information for the markers DXS10079, DXS10074 and DXS10075.

MATERIALS AND METHODS

Blood samples and buccal swabs were collected predominantly from German Caucasoids and some Africans and Asians. Samples were taken from students and their families and from cases of routine kinship testing. DNA extraction was carried out using the QIAamp DNA Blood Kit (Qiagen GmbH, Hilden, Germany). For each STR marker different alleles were selected for sequence analysis resulting in a total of 44 sequenced alleles for DXS10079 and DXS10074, respectively, and 36 sequenced alleles for DXS10075. Additionally, cell line DNAs K562, 9947A (Promega), NA 9948 and NA 3567 (Coriell) were included in the investigation. The amplification was carried out in a 25 µl PCR reaction volume containing approximately 1 ng DNA, 200 µM each dNTP, 1.5 mM $MgCl_2$, 0.5 µM of each primer, 1U Taq polymerase (Applied Biosystems, Foster City, CA) and 1×PCR buffer. The cycle conditions in a PTC-200 cycler (MJ Research Inc, Watertown, MA, USA) were as follows: 95°C – 3 min soak, 94°C - 30 sec, 60°C - 1 min, 72°C - 1 min, 30 cycles, 72°C - 10 min final extension. Sequencing procedure of long amplicons (surrounding the primer binding regions for fragment analysis) was done as previously described (Hering et al., 2005). Direct Taq-cycle sequencing procedure using the BigDye-Terminator Kit v1.1 and the ABI Prism 310 sequencer, Sequencing Analysis v3.7 (Applied Biosystems, Foster City/CA, USA) was performed according to the manufacturers instructions.

RESULTS AND DISCUSSION

DXS10079

The sequence for DXS10079 found in Ensembl v36 part (+) is demonstrated in Fig. 1. The data base search for this genomic region indicated two SNP positions. The T-C transition in nucleotide position (np) 66499227 (known as rs5919390 and 12399215) was found in 11 alleles with different ethnic origin (Table 1). This is in accordance with the genotype frequencies given from the HapMap project (http://www.hapmap.org). We could not detect the C-T transition in the np 66499146 in our sequencing study comprising 49 alleles. However, 2 additional SNPs were found i.e. C/A in np 66498889 and T/G in np 66499054. Since our primers involving the 66498889 SNP were used in fragment analysis procedure, we observed a considerable loss in the yielded PCR product when the A allele is present.

Table 1 Chromosomes found with differences from the GenBank sequence for DXS10079 allele 21. Ethnic origin of the sequenced alleles: 31 Germans, 9 Asians, 4 Africans. In addition, the five alleles from the cell line samples K562, NA 9948, NA 3567 and 9947A were sequenced.

Nucleotide position	SNPView Ensembl v36	Variation	Allele	Ethnic origin (number)
66498889	unknown	C/A	20	German (1)
66499054	unknown	T/G	14 19 20	German (1) NA 9948 9947A
66499146	rs11798277	C/T	not found	
66499227	rs5919390 rs12399215	T/C	18 - 24	German (5) Asian (2) African (4)

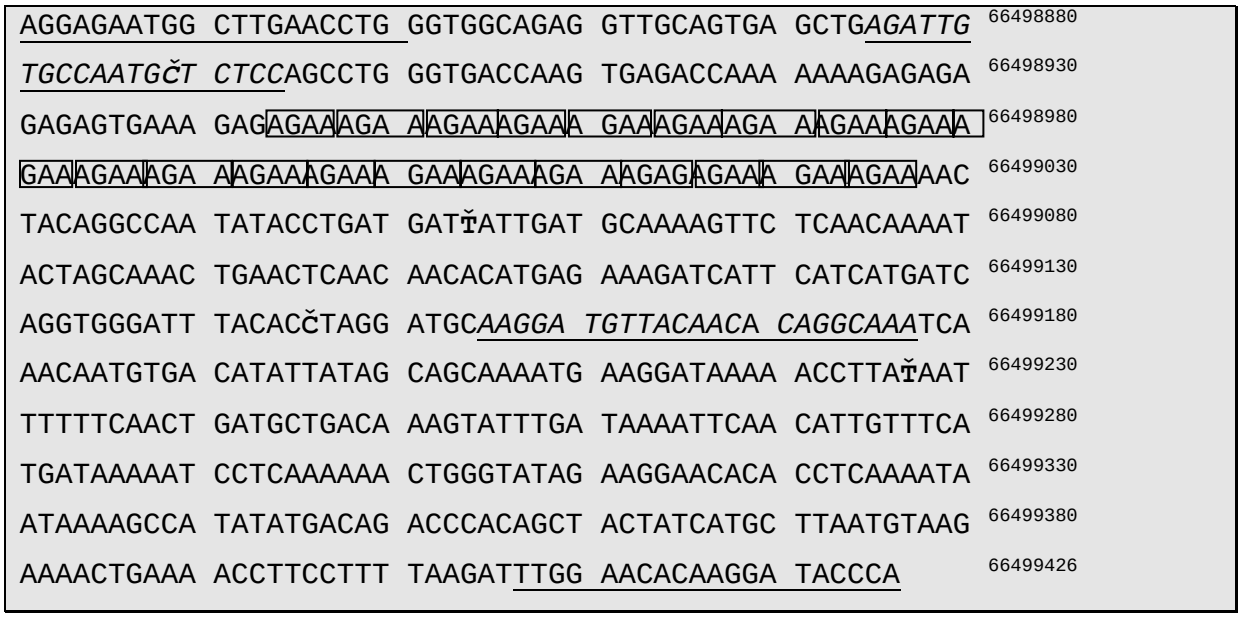

Fig. 1 DXS10079 sequence structure for allele 21 (Ensembl v36): primer binding regions are underlined (long amplicons for sequencing; short amplicons in the multiplex PCR in italics). Tetra nucleotide repeats are boxed. SNP positions are shown as bold letters with accent.

DXS10074

Fig. 2 depicts the sequence structure for allele 14 with three SNP positions. In this STR, two types of tetra nucleotide variation exist (Hering et al., 2005). The "short" alleles from 4 to 17 showing regular AAGA are in agreement with the GenBank sequence. Alleles 4, 11, 13–15 and 17 have not been described so far and seem to be mainly of African origin. The "long" alleles varying from 12 to 21 exhibit one irregular AAGG repeat near the 3' area of the polymorphic region. Alleles 12 and 19.3 have also not been found in the German population up to now.

Table 2 lists SNPs juxtaposed to the DXS10074 repeat region. We found a frequent T-C transition in np 66760436 which is published as rs5918767 in Ensembl v36. In this SNP only 4 African chromosomes exhibited the T allele according to the GenBank sequence. A second SNP in np 66760376 was associated only with "short" STR alleles in German and African chromosomes.

A third SNP was detected in np 66760142 and the very rare G allele is associated with the STR allele 21. This localisation affects the forward primer binding region.

Table 2 lists variations found in the 5' repeat flanking region in alleles 16, 16.2 and 19.3. Sequence structure for one of the allele 16.2 found earlier (Hering et al., 2005) was revised. Finally, the K562 cell line DNA showed a G-T transversion in the polymorphic repeat region.

DXS10075

Only one SNP was found in the region np 6781041 to 66781505 at np 66781224 (Fig. 3). We sequenced the mentioned range of 21 Caucasoids (20 Germans, 1 Russian), 9 Asians (Vietnamese), 6 Africans (4 Mozambiquean, 2 Ghanaian) and the five cell line chromosomes comprising the alleles 12 to 21.3. Predominantly we found a T to A transversion. Only 4 African chromosomes exhibited the T allele as shown in the GenBank sequence. This SNP, i.e. the T-A transversion, is already known and was registered as rs945048.

Our results were in accordance with the *Human SNPView* of Ensembl v36. Available population data from the HapMap project listed the T allele only in Africans, whereas all Asians and Caucasoids exhibited an A at this position.

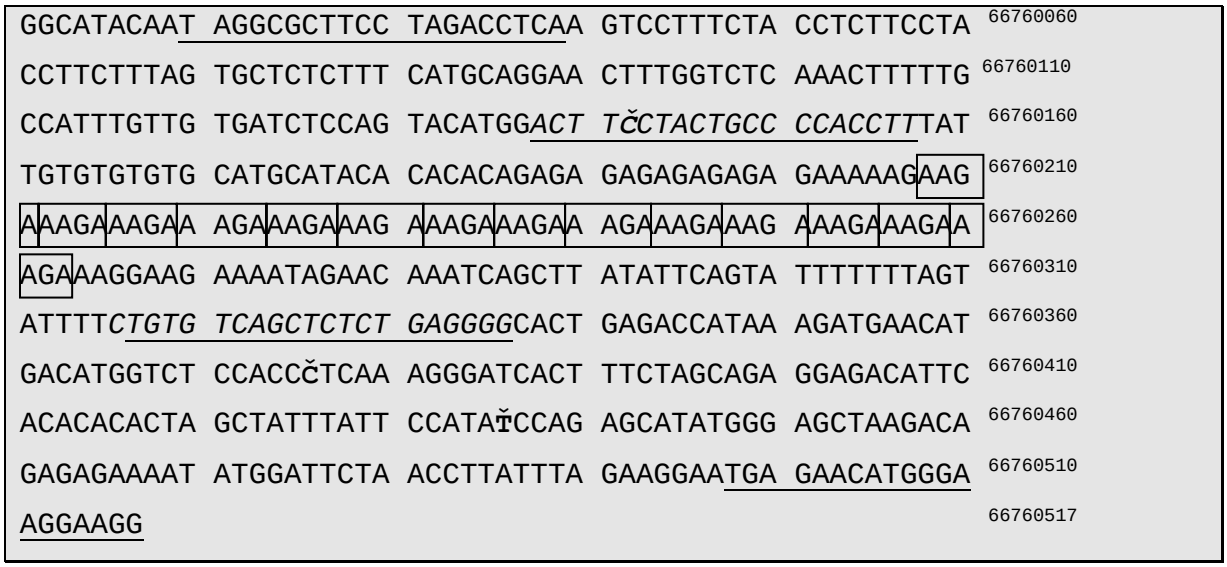

Fig. 2 DXS10074 sequence structure for allele 14 (Ensembl v36): primer binding regions are underlined (long amplicons for sequencing; short amplicons in the multiplex PCR in italics). Tetra nucleotide repeats are boxed. SNP positions are shown as bold letters with accent.

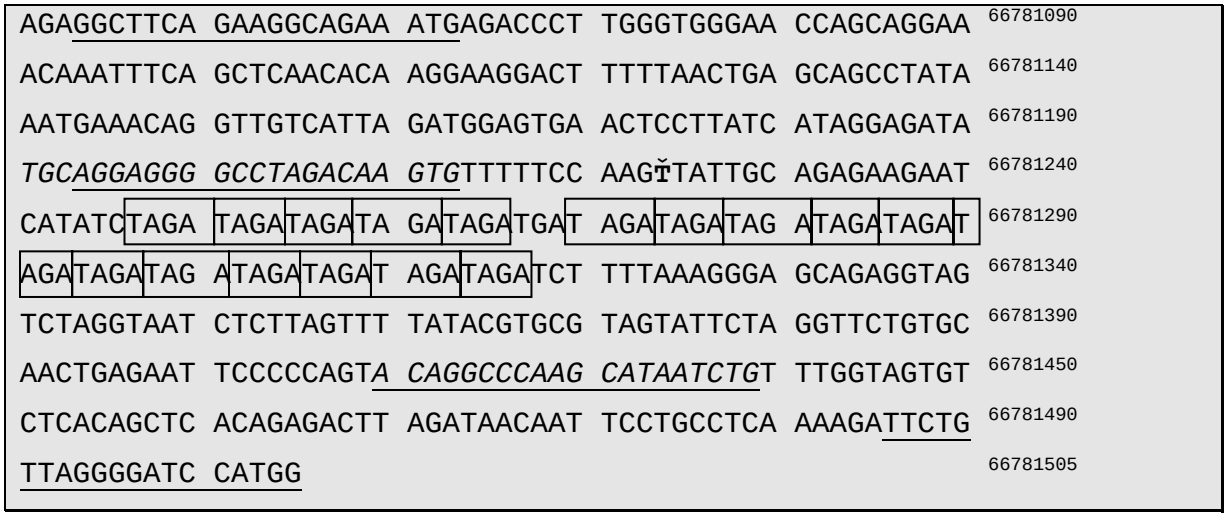

Fig. 3 DXS10075 sequence structure for allele 17 (Ensembl v36): primer binding regions are underlined (long amplicons for sequencing; short amplicons in the multiplex PCR in italics). Tetra nucleotide repeats are boxed. The SNP position is shown as bold letter with accent.

Table 2 DXS10074: SNP combinations for GenBank np 16760376 and 16760436 in relation to different alleles for 31 Germans, 1 Algerian, 6 Asians, 6 Africans, and five cell line chromosomes. Additional variations are printed in bold letters.

66760376 C/T	66760436 T/C	Allele	5'Repeat flanking region	Variable repeat structure	Ethnic origin (number)
T	C	4	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_4$	German (1)
T	C	7	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_7$	German (3), NA 3567
T	C	8	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_8$	German (2)
T	C	9	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_9$	German (2), Algerian (1)
T	C	10	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{10}$	German (1)
C	T	11	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{11}$	African (1)
C	T	13	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{13}$	African (1)
C	T	14	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{14}$	African (1)
T	C	15	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{15}$	African (1)
T	C	17	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{17}$	African (1)
C	C	12	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_9$ AAGG $(AAGA)_2$	German (1)
C	C	13	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{10}$ AAGG $(AAGA)_2$	German (1)
C	C	14	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{11}$ AAGG $(AAGA)_2$	German (2)
C	T	15	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{12}$ AAGG $(AAGA)_2$	African (1)
C	C	15	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{12}$ AAGG $(AAGA)_2$	German (2)
C	C	16	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{13}$ AAGG $(AAGA)_2$	German (1), Asian (1), 9947A
C	C	16	T $(AC)_{\mathbf{3}}$ $(AG)_{\mathbf{9}}$ AA AAAG	$(AAGA)_{13}$ AAGG $(AAGA)_2$	Asian (1)
C	C	16.2	T $(AC)_{\mathbf{3}}$ $(AG)_{\mathbf{7}}$ AA **AA** AAAG	$(AAGA)_{\mathbf{14}}$ AAGG $(AAGA)_2$	German (2)
C	C	17	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_9$ AA**T**A $(AAGA)_4$ AAGG $(AAGA)_2$	K562
C	C	17	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{14}$ AAGG $(AAGA)_2$	German (4), Asian (4)
C	C	18	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{15}$ AAGG $(AAGA)_2$	German (2), NA 9948
C	C	19	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{16}$ AAGG $(AAGA)_2$	German (2), NA 9947A
C	C	19.3	T $(AC)_4$ $(AG)_8$ AA AAAG **AAA**	$(AAGA)_{16}$ AAGG $(AAGA)_2$	German (1)
C	C	20	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{17}$ AAGG $(AAGA)_2$	German (1)
C	C	21	T $(AC)_4$ $(AG)_8$ AA AAAG	$(AAGA)_{18}$ AAGG $(AAGA)_2$	German (3)

Our sequencing results in the examined Xq12 region demonstrate some differences between Germans and Africans in the STR structure and in SNP alleles. Further population data are needed to use such ethnic information in the marker cluster in forensic casework. The information on SNP occurrence in the repeat flanking regions of DXS10079, DXS10074 and DXS10075 may help to minimise pitfalls caused by disturbed primer binding and can assist the development of new X-chromosomal multiplexes.

REFERENCES

Bini C, Ceccardi S, Ferri G, Pelotti S, Alu M, Roncaglia E, Beduschi G, Caenazzo L, Ponzano E, Tasinato P, Turchi C, Buscemi L, Mazzanti M, Tagliabracci A, Toni C, Spinetti I, Domenici R, Presciuttini S. Development of a heptaplex PCR system to analyse X-chromosome STR loci from five Italian population samples. A collaborative study. Forensic Sci Int 2005;153: 231-6.

Edelmann J, Szibor R. DXS101: a highly polymorphic X-linked STR. Int J Legal Med 2001;114: 301-4.

Gusmao L, Amorim A, Prata MJ, Pereira L, Lareu MV, Carracedo A. Failed PCR amplifications of MBP-STR alleles due to polymorphism in the primer annealing region. Int J Legal Med 1996;108: 313-5.

Hering S, Augustin C, Edelmann J, Heidel M, Dressler J, Rodig H, Kuhlisch E, Szibor R. 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 2005;120: 337-45.

Hering S, Edelmann J, Heidel M, Dreßler J, Szibor R. A cluster of six closely linked STR markers: recombination analysis in a 3.6 Mb region at Xq12–13.1. In: Amorim A, Corte-Real F, Morling N (eds.). Progress in forensic genetics 11: proc. of the 22st International ISFG Congress held in Ponta Delgada, The Azores, Portugal, 13-16 Sept. 2005. Amsterdam: Elsevier, 2006, pp. 289-91. (International congress series, Vol. 1288).

Shin SH, Yu JS, Park SW, Min GS, Chung KW. Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. Forensic Sci Int 2005;147: 35-41.

Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D. Use of X-linked markers for forensic purposes. Int J Legal Med 2003;117: 67-74.

Szibor R, Hering S, Kuhlisch E, Plate I, Demberger S, Krawczak M, Edelmann J. Haplotyping of STR cluster DXS6801-DXS6809-DXS6789 on Xq21 provides a powerful tool for kinship testing. Int J Legal Med 2005;119: 363-9.

Tabbada KA, De Ungria MC, Faustino LP, Athanasiadou D, Stradmann-Bellinghausen B, Schneider PM. Development of a pentaplex X-chromosomal short tandem repeat typing system and population genetic studies. Forensic Sci Int 2005;154: 173-80.

Wiegand P, Berger B, Edelmann J, Parson W. Population genetic comparisons of three X-chromosomal STRs. Int J Legal Med 2003;117: 62-5.