Int J Legal Med (2005) 119: 310–313
DOI 10.1007/s00414-005-0554-8

SHORT COMMUNICATION

Marielle Heinrich · Heike Felske-Zech · Bernd Brinkmann · Carsten Hohoff

Characterisation of variant alleles in the STR systems D2S1338, D3S1358 and D19S433

Received: 16 February 2005 / Accepted: 11 April 2005 / Published online: 18 June 2005

Abstract We have observed three hitherto undescribed off-ladder alleles at three widely used STR loci. These were isolated, sequenced and designated as follows: allele 10 (D2S1338, one case), allele 21 (D3S1358, two cases) and allele 6.2 (D19S433, six cases). These sequences are described in comparison to non-variant alleles, and their implications for the semi-automated STR analysis will be discussed.

Keywords STR · Rare allele · Microsatellite · Variant allele

Introduction

The short tandem repeat (STR) loci D2S1338, D3S1358 and D19S433 are commonly used (e.g., [1]) in human parentage as well as identity testing and are part of some national DNA databases (e.g., UK, Austria). Allelic ladders from kits (e.g., [9]) contain alleles 15–27 (D2S1338) and 9–18 (D19S433), whereas at D3S1358, 11 different alleles ranging from 11 to 20 [2, 3] and even up to 26 repeats [4] have been described. The NIST STRbase (http://www.cstl.nist.gov/biotech/strbase) refers to more than 250 variant alleles, but only very few of them have been characterized properly and published so far (e.g., [5, 6]).

M. Heinrich · B. Brinkmann (✉) · C. Hohoff
Institut für Rechtsmedizin,
Universitätsklinikum Münster,
Röntgenstr. 23,
48149 Münster, Germany
e-mail: brinkma@uni-muenster.de
Fax: +49-251-8355158

H. Felske-Zech
Institut für Rechtsmedizin,
Charité–Campus Benjamin Franklin,
Hittorfstr. 18,
14195 Berlin, Germany

Materials and methods

Genomic DNA was extracted from buccal scrapings as follows: a small piece of the cotton wool was cut off by sterile scissors; 50 µl of 20 mg/ml proteinase K (Roche, Mannheim, Germany) and 200 µl of 5% (w/v) chelex-100 suspension (Bio-Rad, München, Germany) were added. The mixture was incubated for 30 min at 56°C and then boiled for 8 min. D3S1358, D2S1338 and D19S433 were amplified using commercially available kits (AmpFlSTR SGMplus [7] and SEfiler [8], Applied Biosystems, Darmstadt, Germany; genRES MPX-3 [9], SERAC, Bad Homburg, Germany) and analyzed on ABI PRISM 310 and 3100 Genetic Analyzers with Genotyper 2.5.2 (Applied Biosystems). For sequencing, the samples were amplified using published primer sequences ([10]; Genbank/GDB accession nos.: G08202, 196594 and G08036). After eluting the alleles of interest from polyacrylamide gels, the sequence was determined on both strands using the BigDye Terminator kit v2.0 and an ABI PRISM 310 Genetic Analyzer (Applied Biosystems).

Results and discussion

In one out of 2,857 samples [allele frequency, 0.0002; 95% confidence interval (CI) 0–0.001], an off-ladder allele for D2S1338 was observed by using the SGM Plus kit and the MPX-3 kit. It was considered to be allele 10, which was verified by sequencing. The variable region shows the following structure that follows the consensus structure (data given in Table 1):

$(TGCC)_4(TTCC)_6$

Using the MPX-3 kit, allele 10 in D2S1338 is located within the range of D3S1358 (Fig. 1a). Using the SGM plus kit, this allele is located within the range of D16S539. In both multiplex kits, allele 10 in D2S1338 represents an integer allele in the range of appearance. Thus, the risk of mistyping of this allele is very high.

Table 1 Overview of the sequenced alleles of the STR systems D2S1338, D3S1358 and D19S433

	allele	sequence of repeat region	n
D2S1338	**10**	$(TGCC)_4$ $(TTCC)_6$	1
	17	$(TGCC)_6$ $(TTCC)_{11}$	1
	18	$(TGCC)_6$ $(TTCC)_{12}$	1
	19	$(TGCC)_8$ $(TTCC)_{11}$	1
		$(TGCC)_7$ $(TTCC)_{12}$	1
	20	$(TGCC)_7$ $(TCCC)$ $(TTCC)_{12}$	3
		$(TGCC)_7$ $(TTCC)_{10}$ $(GTCC)_1$ $(TTCC)_2$	1
	21	$(TGCC)_7$ $(TTCC)_{11}$ $(GTCC)_1$ $(TTCC)_2$	1
	22	$(TGCC)_7$ $(TTCC)_{12}$ $(GTCC)_1$ $(TTCC)_2$	1
	23	$(TGCC)_7$ $(TTCC)_{13}$ $(GTCC)_1$ $(TTCC)_2$	2
	24	$(TGCC)_7$ $(TTCC)_{14}$ $(GTCC)_1$ $(TTCC)_2$	2
	25	$(TGCC)_7$ $(TTCC)_{15}$ $(GTCC)_1$ $(TTCC)_2$	2
	26	$(TGCC)_7$ $(TTCC)_{16}$ $(GTCC)_1$ $(TTCC)_2$	2
	27	$(TGCC)_7$ $(TTCC)_{17}$ $(GTCC)_1$ $(TTCC)_2$	2
D3S1358	11	$(AGAT)_7$ $(AGAC)_2$ $(AGAT)_2$	2
	12	$(AGAT)_8$ $(AGAC)_2$ $(AGAT)_2$	1
	13	$(AGAT)_9$ $(AGAC)_2$ $(AGAT)_2$	2
	14	$(AGAT)_{10}$ $(AGAC)_2$ $(AGAT)_2$	3
	15	$(AGAT)_{10}$ $(AGAC)_3$ $(AGAT)_2$	1
		$(AGAT)_{11}$ $(AGAC)_2$ $(AGAT)_2$	13
		$(AGAT)_{12}$ $(AGAC)_1$ $(AGAT)_2$	2
	15.2	$(AGAT)_{10}$ $(AG)_1$ $(AGAC)_3$ $(AGAT)_2$	2
	16	$(AGAT)_{11}$ $(AGAC)_3$ $(AGAT)_2$	5
		$(AGAT)_{12}$ $(AGAC)_2$ $(AGAT)_2$	13
		$(AGAT)_{13}$ $(AGAC)_1$ $(AGAT)_2$	2
	17	$(AGAT)_{13}$ $(AGAC)_2$ $(AGAT)_2$	13
		$(AGAT)_{12}$ $(AGAC)_3$ $(AGAT)_2$	10
		$(AGAT)_{14}$ $(AGAC)_1$ $(AGAT)_2$	1
	18	$(AGAT)_{13}$ $(AGAC)_3$ $(AGAT)_2$	12
		$(AGAT)_{12}$ $(AGAC)_4$ $(AGAT)_2$	1
		$(AGAT)_{14}$ $(AGAC)_2$ $(AGAT)_2$	4
	19	$(AGAT)_{14}$ $(AGAC)_3$ $(AGAT)_2$	3
		$(AGAT)_{15}$ $(AGAC)_2$ $(AGAT)_2$	1
	20	$(AGAT)_{16}$ $(AGAC)_2$ $(AGAT)_2$	1
	21	$(AGAT)_{16}$ $(AGAC)_3$ $(AGAT)_2$	2
D19S433	**6.2**	$(AAGG)$ (AA) $(AAGG)$ $(TAGG)$ $(AAGG)_5$	6
	9	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGA)$ $(AAGG)_6$	1
	10	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)_8$	1
	11	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)_9$	2
	11.1	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)$ A $(AAGG)_8$	1
	12	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)_{10}$	2
	12.1	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)_5$ A $(AAGG)_5$	4
	13	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)_{11}$	3
	13.2	$(AAGG)$ (AA) $(AAGG)$ $(TAGG)$ $(AAGG)_{12}$	1
	14	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)_{12}$	3
	15	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)_{13}$	3
	15.2	$(AAGG)$ (AA) $(AAGG)$ $(TAGG)$ $(AAGG)_{14}$	1
	16	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)_{14}$	4
	16.2	$(AAGG)$ (AA) $(AAGG)$ $(TAGG)$ $(AAGG)_{15}$	1
	17	$(AAGG)$ $(AAAG$ $(AAGG)$ $(TAGG)$ $(AAGG)_{15}$	1
	17.2	$(AAGG)$ (AA) $(AAGG)$ $(TAGG)$ $(AAGG)_{16}$	2
	18	$(AAGG)$ $(AAAG)$ $(AAGG)$ $(TAGG)$ $(AAGG)_{16}$	2

The repeat structure and the number of sequenced alleles (*n*) is indicated; the three variants from this study are shown in bold

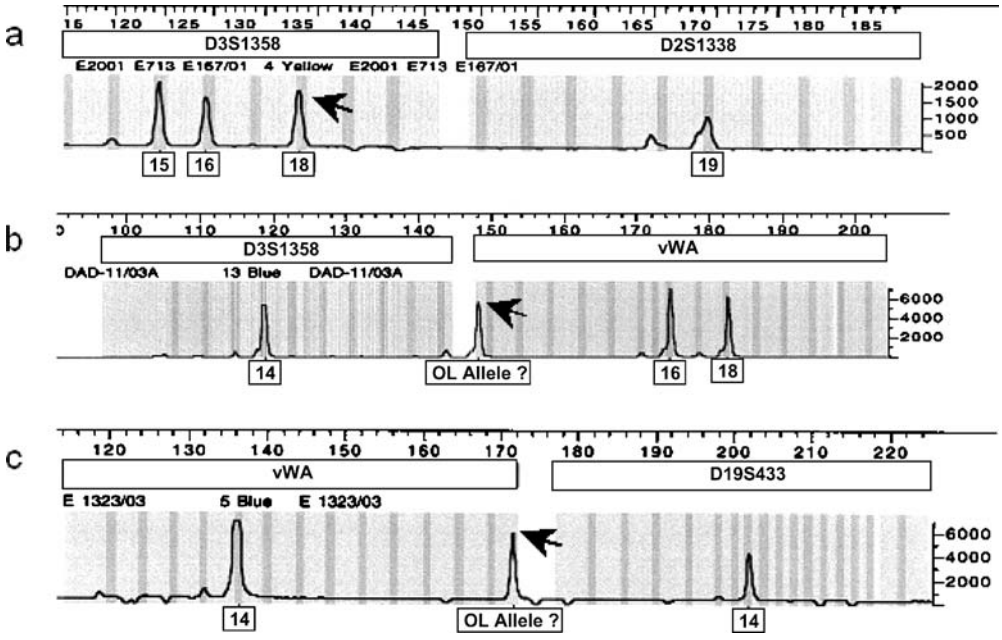

Fig. 1 **a** Allele 10 at D2S1338 (indicated by an *arrow*) appears as allele 18 in the category range of D3S1358 (genRES MPX-3). **b** Allele 21 at D3S1358 (indicated by an *arrow*) appears as "OL Allele" in the category range of vWA (AmpFlSTR SEfiler). **c** Allele 6.2 at D19S433 (indicated by an *arrow*) appears as "OL Allele" in the category range of vWA (genRES MPX-3)

After multiplex PCR analysis, a peak outside of the Genotyper range for D3S1358 was detected in two out of 11,070 samples (allele frequency, 0.0001; 95% CI 0–0.0003; Fig. 1b). According to the apparent amplicon size, we considered it as allele 21. This was confirmed by sequencing leading to the following structure of the variable region, which is similar to the other alleles at D3S1358 with three variable AGAY blocks (see Table 1):

$$(AGAT)_{16}(AGAC)_3(AGAT)_2$$

Allele 21 at D3S1358 appears in a different category range, e.g., in vWA using SEfiler (Fig. 1b).

In four—as far as we know—unrelated persons in a set of 4,181 samples (allele frequency, 0.0005; 95% CI 0.0002–0.0013), a peak outside the range of D19S433 appeared in the electropherogram. According to the apparent amplicon size, we considered it as allele 6.2, which was verified by sequencing. The variable region shows the following structure that is similar to those of longer, non-integer alleles, e.g., 15.2, 16.2, 17.2 (data shown in Table 1):

$$(AAGG)(AA)(AAGG)(TAGG)(AAGG)_5$$

Using the MPX-3 multiplex kit, the allele 6.2 in D19S433 appears in the category range of the STR system vWA (Fig. 1c). Using the SEfiler or the SGM Plus kit, the allele 6.2 may not be observed because the default settings of the Genotyper macro show a zoom window size starting at 95 bp. Thus, the peak indicating allele 6.2 (apparent size, 91.5 bp) is outside of this range. This may lead to mistyping of samples as homozygote in D19S433.

Interestingly, all four persons carrying an allele 6.2 in D19S433 are known to have a Turkish origin. Two additional samples carrying 6.2 in D19S433 were sent to us from other laboratories for sequence analysis: one from Turkey and one with an unknown origin but with a Turkish forename. In Turkish population genetic studies, the observation of an allele 6.2 has not been mentioned so far (e.g., [11]).

All these rare alleles have implications in the (semi-automated) analysis: the zoom of the Genotyper window should start at ~85 bp in case of using the SEfiler and/or the SGM Plus kits to assure that short alleles in D19S433 are indeed displayed. Additionally, one should not solely rely on the category ranges given in a Genotyper macro, but rather keep in mind the possibility of off-ladder variant alleles.

Finally, understanding STR particularities such as primer binding site mutations [12] and off-ladder alleles should assist the future development of improved reagents to type STR loci correctly.

Acknowledgements The authors would like to thank Anke Junge (Institute of Legal Medicine, Bonn, Germany) and Günter Braunschweiger (genFor, Grafschaft Gelsdorf, Germany) for submitting a sample and cloned alleles, respectively, as well as Marianne Schürenkamp, Karin Rauße and Jennifer Bartsch for their excellent technical assistance.

This work is part of the Ph.D. thesis of M.H.

References

1. Grubwieser P, Mühlmann R, Parson W (2003) New sensitive amplification primers for the STR locus D2S1338 for degraded casework DNA. Int J Legal Med 117:185–188
2. Mornhinweg E, Luckenbach C, Fimmers R, Ritter H (1998) D3S1358: sequence analysis and gene frequency in a German population. Forensic Sci Int 95:173–178
3. Lazaruk K, Wallin J, Holt C, Nguyen T, Walsh PS (2001) Sequence variation in humans and other primates at six short tandem repeat loci used in forensic identity testing. Forensic Sci Int 119:1–10
4. Grubwieser P, Mühlmann R, Niederstätter H, Pavlic M, Parson W (2005) Unusual variant alleles in commonly used short tandem repeat loci. Int J Legal Med 119:164–166

5. Bhoopat T, Hohoff C, Steger HF (2003) Identification of DYS385 allele variants by using shorter amplicons and Northern Thai haplotype data. J Forensic Sci 48:1108–1112
6. Klein R, Braunschweiger G, Junge A, Wiegand P (2003) A very long ACTBP2 (SE33) allele. Int J Legal Med 117:235–236
7. Cotton EA, Allsop RF, Guest JL, Frazier RR, Koumi P, Callow IP, Seager A, Sparkes RL (2000) Validation of the AMPFlSTR SGM plus system for use in forensic casework. Forensic Sci Int 112:151–161
8. Coticone RS, Oldroyd N, Philips H, Foxall P (2004) Development of the AmpflSTR SEfiler PCR amplification kit: a new multiplex containing the highly discriminating ACTBP2 (SE33) locus. Int J Legal Med 118:224–234
9. Schlenk J, Seidl S, Braunschweiger G, Betz P, Lederer T (2004) Development of a 13-locus PCR multiplex system for paternity testing. Int J Legal Med 118:55–61
10. Li H, Schmidt L, Wei M-H, Hustad T, Lerman MI, Zbar B, Tory K (1993) Three tetranucleotide polymorphisms for loci: D3S1352, D3S1358, D3S1359. Hum Mol Genet 2:1327
11. Hadi Cakyr A, Simsek F, Katyrcy N, Tasdelen B (2004) STR data for the AmpFlSTR SGM Plus from the eastern and western sections of Mediterranean region of Turkey. Forensic Sci Int 142:55–57
12. Heinrich M, Müller M, Rand S, Brinkmann B, Hohoff C (2004) Allelic drop-out in the STR system ACTBP2 (SE33) as a result of mutations in the primer binding region. Int J Legal Med 118: 361–363