Int J Legal Med (1998) 112 : 50–54

ORIGINAL ARTICLE

M. Klintschar · B. Glock · E. M. Dauber · W. R. Mayr

Genetic variation and sequence studies of a highly variable short tandem repeat at the D17S976 locus

Received: 7 October 1997 / Received in revised form: 7 May 1998

Abstract The STR locus D17S976 was investigated by PCR amplification and native polyacrylamide gel electrophoresis in 158 unrelated Austrian Caucasians. No deviations from Hardy-Weinberg expectations were observed. The mean exclusion chance was 0.792, the discriminating power was 0.980 and the observed heterozygosity rate was 0.873. Moreover two alternative denaturing electrophoretic protocols are proposed. An allelic ladder consisting of 14 sequenced alleles (236–288 bp) was constructed. Sequence analysis revealed that the locus contained three different repeat motifs: ATCA, ATCT and ACCT, all of which vary in number between alleles. The aggregate number of the three tetrameric repeat types was used for allele designation. As a repeat with a single base deletion (ATC) was found in both the smallest and the largest alleles, a ".3" was added to the allele designation in those cases. Therefore the smallest allele is designated 19.3, and the largest allele is designated 32.3. To evaluate the exact extent of sequence variation more extensive sequence studies are necessary.

Key words Short tandem repeat · D17S976 · DNA Sequencing · Population data · Forensic DNA typing

Introduction

Short tandem repeat systems (STRs) are DNA polymorphisms consisting of di- to pentameric repeats [1, 2]. They are highly sensitive and allow typing of stains which are severely degraded as the amplified fragments are usually shorter than 300 bp [3]. On the other hand, most STRs are distinctly less polymorphous than RFLPs with only 3–6 common alleles [1]. Therefore a larger number of systems has to be typed for comparable results and some of the advantages of the STR approach, e.g. the saving of time and material [3], are affected. One possibility to overcome this problem is typing highly polymorphic STRs such as ACTBP2 [4]. But these STRs are more difficult to type and are less suitable for inter-laboratory controls [5]. As STR loci are abundant in the human genome [2] and only few of them have yet been evaluated for forensic purposes, it might be possible to find new, more suitable STRs. Therefore, the tetrameric repeat locus D17S976 (GDB entry number G00–250–233) was selected as the forensic utility has not yet been investigated. The aim of this study was to evaluate the genetic variation of this STR in an Austrian Caucasian population sample, to test its sensitivity and to construct a sequenced allelic ladder.

M. Klintschar (✉)
Institut für Gerichtliche Medizin, Karl Franzens Universität, Universitätsplatz 4, A-8010 Graz, Austria
e-mail: michael.klintschar@kfunigraz.ac.at
Fax +42 (316) 380 9655

B. Glock · E. M. Dauber · W. R. Mayr
Institut für Blutgruppenserologie und Transfusionsmedizin, Universität Wien, Währinger Gürtel 18–20/4i, A-1090 Wien, Austria

Materials and methods

Preparation of DNA

DNA was extracted from blood samples of 158 unrelated Austrian caucasians as described [6]. For the sensitivity study K562 cell line DNA was stepwise diluted down to 10 pg per µl.

PCR protocol

Amplification was performed using 5 µl extracts without previous quantification in a 25 µl volume. The reaction mixture included 0.5 U Dynazyme DNA Polymerase (Finnzymes, Espoo, Finland), 0.4 µM each primer, 100 µM each nucleotide, 2.5 µl 10 × PCR buffer (Finnzymes, Espoo, Finland).

Primer sequences

Primer 1: ATA TGC CAC CAC ACC TGG TT
Primer 2: TGG TAG CAT GCA TCT GTA GTC C
(GDB entry number G00–250–233)

Amplification protocol

94 °C 1 min, 66 °C 1 min, 72 °C 2 min for 6 cycles
93 °C 1 min, 65 °C 1 min, 72 °C 2 min for 6 cycles
93 °C 1 min, 63 °C 1 min, 72 °C 2 min for 18 cycles

Electrophoresis

Non-denaturing polyacrylamide gel electrophoresis (PAGE) (5%T, 3%C-750 μm) was performed at a separation distance of 16 cm with piperazine diacrylamide as crosslinker, 80 mM formate, 28 mM CHES and 2% agarose plugs. Running conditions were at an initial 1000 V, 40 mA and 5 W with ramping for 5 W every 90 min until the bromphenol blue marker front reached the anodal plug. Bands were visualized by silver staining [7]. This native electrophoretic setup was used for typing all 158 persons included in the present study. Additionally, two different types of denaturing PAGE were tested.

Firstly the alleles were separated on a SG-600–20 vertical gel apparatus (CBS, Del Mar, Ca.) using a 5% denaturing gel and visualized by silver staining (STR systems manual #TMD004, Promega Corporation, Madison, WI). Secondly, the PCR products were detected by fluorescence laser detection on an A.L.F. sequencer (Pharmacia Biotech, Sweden): electrophoresis was carried out using 1 μl of the amplification product, 7.5 fmol of the size standards 200 bp and 300 bp (Pharmacia Biotech) and 3 μl loading buffer (Pharmacia Biotech). After denaturating, the samples were run on an 0.5 mm thick, 6% denaturing polyacrylamide gel at 50 °C, 1500 V, 38 mA, 34 W in 0.6 × Tris-Borat EDTA for 200 min.

Using these denaturing protocols 40 samples which were already typed with the native setup were retyped.

Sequencing

For sequencing, alleles eluted from polyacrylamide gels using the QIAexII kit (Qiagen, FRG) were reamplified with 0.05 μM biotinylated forward primer and 0.05 μM reverse primer. Strand separation and single stranded sequencing (solid phase) was conducted on an A.L.F. automatic sequencer (Pharmacia, Sweden) as previously described [8].

Allelic designation and allelic ladder

Allelic designation was made according to the recommendations of the DNA Commision of the International Society for Forensic Haemogenetics [9–11] based on the number of repeats in the repeating unit and was therefore defined as (AYCW)n where Y denotes C/T and W denotes A/T [12]. To the designation of fragments showing an ATC trinucleotide insertion, i.e. an incomplete repeat, a .3 was added. Of the alleles sequenced in this study 14 were used to construct an allelic ladder.

Statistical evaluation

For evaluating Hardy-Weinberg expectations, an exact test [13] using the computer program GENEPOP, version 1.2 (M. Raymond and F. Rousset, Montpellier) was performed. The mean exclusion chance (ME) was calculated according to Krüger et al. [14] and the discriminating power was calculated as $1-\Sigma$ (expected phenotype frequencies)^2 [15].

Results

Sequence structure of D17S976 alleles

In 158 Austrian individuals 15 different alleles were found using native PAGE, 18 alleles were sequenced and 14 were used to construct the allelic ladder. As the GDB entry of the D17S976 locus includes no sequencing data, we propose the sequence given in Table 1 as a consensus sequence. Comparisons of the 26 bp 5′ flanking region and the 112 bp 3′ flanking region showed absolute length and sequence conformity in the alleles sequenced. The repeat composition of the D17S976 alleles with fragment lengths ranging between 236 and 288 bp is displayed in Fig. 1 and Table 2. D17S976 is a compound STR [16] consisting of ATCA, ATCT, and ACCT repeats, all varying in number between alleles. Moreover, in the smallest and the largest alleles an ATC trinucleotide was found. In the repeat regions of all alleles an isolated T was observed after 10–17 repeats and an ATCCT unit was observed after a single ATCT repeat. A basic sequence structure of $(ATCA)_{1–3}$ $(ATCT)_{2,10–13}$ $(ATC)_{0–1}$ $(ATCT)_{0,5,9–12}$ T (ATCT) ATCCT $(ATCT)_{2,8–11,13}$ $(ACCT)_{0–1}$ $(ATCT)_{0.8–11}$ was found.

Table 1 Consensus sequence of the locus D17S976 (allele 19.3). Repeat region in bold print, primer sequences underlined

```
  1                                                                  50
5′-ATATGCCACCACACCTGGTTTTTTAAATCAATCAATCAATCTATCTATCA
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3′-TATACGGTGGTGTGGACCAAAAAATTTAGTTAGTTAGTTAGATAGATAGT

  51                                                                 100
  TCTATCTATCTATCTATCTTATCTATCCTATCTATCTATCTATCTATCTA
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  AGATAGATAGATAGATAGAATAGATAGGATAGATAGATAGATAGATAGAT

  101                                                                150
  TCTATCTATCTAAGATGGAGTCTCACCCTGTAGCCCAAGCTGGAGTGCAG
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  AGATAGATAGATTCTACCTCAGAGTGGGACATCGGGTTCGACCTCACGTC

  151                                                                200
  TGGTGCAATCTTGGCTCACTGCAACCTCTGCCTCCCATGCTCAAGCGATT
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  ACCACGTTAGAACCGAGTGACGTTGGAGACGGAGGGTACGAGTTCGCTAA

  201                                    236
  CTCCCAAGTAGCTGGGACTACAGATGCATGCTACCA-3′
  ||||||||||||||||||||||||||||||||||||
  GAGGGTTCATCGACCCTGATGTCTACGTACGATGGT-5′
```

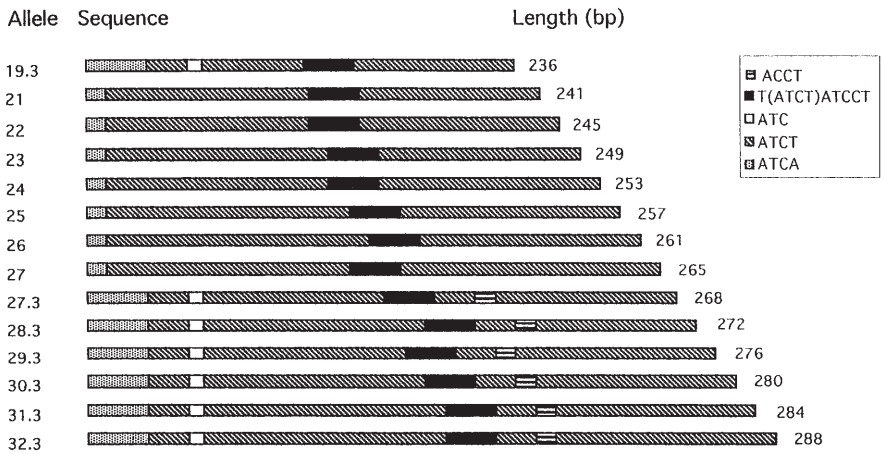


Fig. 1 Sequence composition of the repeat regions of the 14 alleles included in the D17S976 ladder

Table 2 Sequence structure of the repeat regions found in 18 D17S976 alleles (5′-3′). (n: number of alleles sequenced; Asterisk: Allele included into the allelic ladder

Allele	Sequence	Length (bp)	n
19.3	$(ATCA)_3$-$(ATCT)_2$-ATC-$(ATCT)_5$-T-(ATCT)-ATCCT-$(ATCT)_8$	236	2*
21	$(ATCA)_1$-$(ATCT)_{10}$-T-(ATCT)-ATCCT-$(ATCT)_9$	241	1*
21.3	$(ATCA)_3$-$(ATCT)_2$-ATC$(ATCT)_5$-T-(ATCT)-ATCCT-$(ATCT)_{10}$	244	1
22	$(ATCA)_1$-$(ATCT)_{10}$-T-(ATCT)-ATCCT-$(ATCT)_{10}$	245	1*
23	$(ATCA)_1$-$(ATCT)_{11}$-T-(ATCT)-ATCCT-$(ATCT)_{10}$ $(ATCA)_1$-$(ATCT)_{10}$-T-(ATCT)-ATCCT-$(ATCT)_{11}$	249	1* 1
24	$(ATCA)_1$-$(ATCT)_{11}$-T-(ATCT)-ATCCT-$(ATCT)_{11}$	253	1*
25	$(ATCA)_1$-$(ATCT)_{12}$-T-(ATCT)-ATCCT-$(ATCT)_{11}$	257	1*
26	$(ATCA)_1$-$(ATCT)_{13}$-T-(ATCT)-ATCCT-$(ATCT)_{11}$	261	1*
27	$(ATCA)_1$-$(ATCT)_{12}$-T-(ATCT)-ATCCT-$(ATCT)_{13}$	265	1*
27.3	$(ATCA)_3$-$(ATCT)_2$-ATC-$(ATCT)_9$-T-(ATCT)-ATCCT-$(ATCT)_2$-(ACCT)-$(ATCT)_9$	268	1*
28.3	$(ATCA)_3$-$(ATCT)_2$-ATC-$(ATCT)_{11}$-T-(ATCT)-ATCCT-$(ATCT)_2$-(ACCT)-$(ATCT)_8$	272	1*
29.3	$(ATCA)_3$-$(ATCT)_2$-ATC-$(ATCT)_{10}$-T-(ATCT)-ATCCT-$(ATCT)_2$-(ACCT)-$(ATCT)_{10}$	276	1*
30.3	$(ATCA)_3$-$(ATCT)_2$-ATC-$(ATCT)_{11}$-T-(ATCT)-ATCCT-$(ATCT)_2$-(ACCT)-$(ATCT)_{10}$	280	1*
31.3	$(ATCT)_3$-$(ATCT)_2$-ATC-$(ATCT)_{12}$-T-(ATCT)-ATCCT-$(ATCT)_2$-(ACCT)-$(ATCT)_{10}$	284	2*
32.3	$(ATCA)_3$-$(ATCT)_2$-ATC-$(ATCT)_{12}$-T-(ATCT)-ATCCT-$(ATCT)_2$-(ACCT)-$(ATCT)_{11}$	288	1*

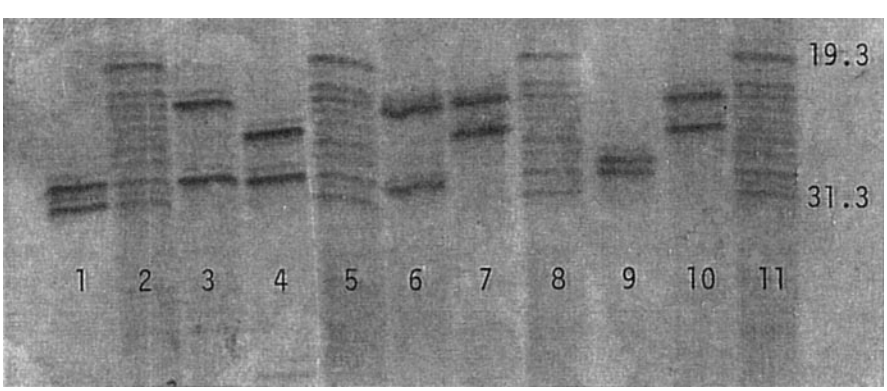


Fig. 2 Native electropherogram of the locus D17S976. Lanes 2, 5, 8, 11: allelic ladder (19.3–32.3) lane 1: 29.3, 31.3, lane 3: 23, 29.3, lane 4: 25, 29.3, lane 6: 23, 29.3, lane 7: 23, 26, lane 9: 28.3, 29.3, lane 10: 23, 26

Population data

Figure 2 shows a typical electropherogramm of the D17S976 system on a horizontal, native PAGE while Fig. 3 shows the separation of the allelic ladder on an automatic Sequencer (A.L.F., Pharmacia, Sweden) under denaturing conditions. All 40 samples which were typed using both native and denaturing conditions gave the same results for both methods. Between allele 21 and 22 an interallele was observed once. Sequencing confirmed that it included two additional ATCT repeats compared to allele 19.3 and it was thus designated 21.3 (Table 2). No other interalleles have been observed so far. Allele frequencies of the D17S976 locus in Austria are shown in Fig. 4. A to-

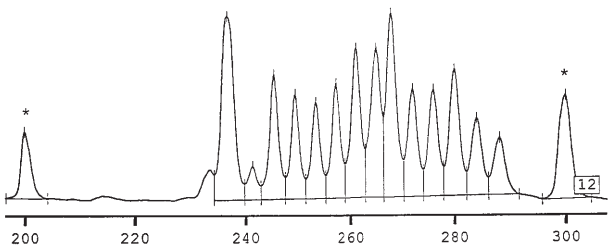


Fig. 3 Denaturing electropherogram of the allelic ladder for D17S976 (alleles 19.3 through allele 32.3) detected by fluorescence laser detection on the A.L.F. sequencer. The asterisks indicate the 200 bp and 300 bp size markers

tal of 67 genotypes were found in the 158 subjects tested. Only three alleles were more frequent than 10%: 23 (15.8%), 19.3 (13.6%) and 22 (10.8%). No significant deviations from Hardy Weinberg expectations could be observed using the exact test ($p = 0.24$). The heterozygosity rate was 0.873, the mean exclusion chance (ME) 0.792 and the discriminating power (PD) 0.980.

Sensitivity studies

Correct typing was possible with as little as 100 pg DNA per 25 µl assay.

Primer design

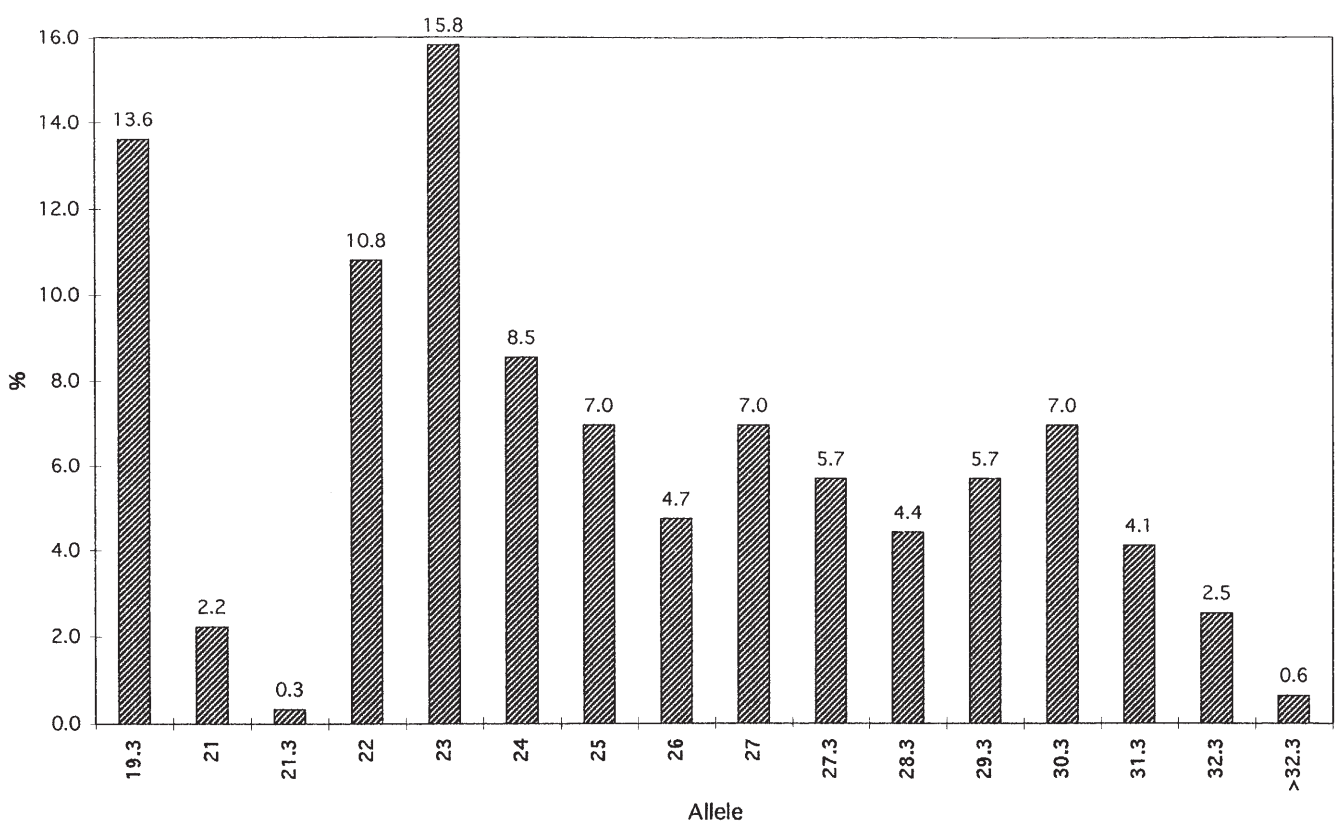
In this study the original primers given in the GDB entry were used. In an attempt to improve the sensitivity or to reduce the length of the amplified fragments new primers were constructed using the PRIMER program. However, none of these primers gave satisfactory results.

Discussion

The D17S976 locus is highly polymorphic and easy to type on both native and denaturing polyacrylamide gels (Fig. 2). The sensitivity of this locus is comparable to that of other STRs [17]. The polymorphicity, however, is higher than that of most STRs commonly used in forensic practice, ranging between that of ACTBP2 [18] and FGA [19] and comparable to that of D12S391, another recently introduced STR [20, 21]. An additional advantage of the locus D17S976 might be that no other STRs on chromosome 17 are routinely used for forensic purposes and that thus linkage disequilibrium is unlikely.

Given the complicated structure of the D17S976 locus, it is surprising that only one interallele (21.3) has so far been observed. Although only 18 alleles have been sequenced up to now, these results suggest the occurrence of both sequence variation (e.g. allele 23: different distribution of the 3 types of tetrameric repeats) and length variation concerning the occurrence of the TCA trimer (e.g. allele 22/21.3) (Table 2). The combination of tetrameric repeats and a single trimeric repeat unit is well known for TH01 [22], for which a reliable resolution of fragments

Fig. 4 Histogram of the D17S976 frequencies in 158 Austrian Caucasians



differing for 1 bp is possible [23] and seems to be especially common in STRs featuring ATCT repeats [16]. Although the alleles of the locus D17S976 are larger than those for THO1, the detection of differences in migration due to 1 bp deletions should be possible using any of the electrophoretic setups proposed. As we have only once observed such differences in 158 persons, it seems probable that, similar to TH01 [22], the occurrence of the triplet is for the most part limited to the shortest and the longest alleles, e.g. the allele 19.3 seems to be common, while 20 should be extremely rare. For confirming these assumptions, however, the number of alleles sequenced is not suffficient and further sequencing studies have to be performed.

The rather unusual amplification protocol with three different annealing temperatures was chosen because the PCR proved to be most sensitive and least prone to artifacts.

Although the compound repeat structure and the reliability of this system in both forensic stain analysis and paternity testing has to be investigated further, this locus promises to be a useful tool since it is highly polymorphic and sufficiently sensitive. In particular, the fact that this locus is superior to most STRs in respect to its polymorphicity while still amenable to typing on native gels should make this locus especially interesting for smaller laboratories.

Acknowledgements The authors are indebted to Christa Nöhammer and Maria Langer for useful hints. This work was partially supported by grant 6523/1 of the Jubilee Fonds of the Austrian National Bank to M. Klintschar.

References

1. Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49:746–756
2. Weber JL (1990) Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. Genomics 7:524–530
3. Brinkmann B (1992) The use of STRs in stain analysis. In: Proceedings from the third international symposium on human identification. Promega, Medison, USA, pp 357–373
4. Polymeropoulos MJ, Rath DS, Xiao H, Merril CR (1992) Tetranucleotide repeat polymorphism at the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2). Nucleic Acids Res 20:1432
5. Gill P, Kimpton C, D'Aloja E, Andersen JF, Bär W, Brinkmann B, Holgersson S, Johnsson V, Kloosterman AD, Lareu MV, Nellemann L, Pfitzinger H, Phillips CP, Schmitter H, Schneider PM, Stenersen M (1994) Report of the European DNA profiling group (EDNAP) – towards standardisation of short tandem repeat (STR) loci. Forensic Sci Int 65:51–59
6. Klintschar M, Kozma Z, Al Hammadi N, Abdull Fatah M, Nöhammer C (1998) A study on the short tandem repeat loci HumCD4, HumTHO1, and HumFIBRA in a Yemenian and and Egyptian population sample. Int J Legal Med 111:107–110
7. Allen RC, Graves G, Budowle B (1989) Polymerase chain reaction amplification products separated on rehydratable polyacrylamide gels and stained with silver. Biotechniques 7:736–744
8. Glock B, Schwartz DWM, Schwartz-Jungl EM, Mayr WR (1996) Allelic ladder characterization of the short tandem repeat polymorphism in intron 6 of the lipoprotein lipase gene and its application in an Austrian Caucasian population study. J Forensic Sci 41:579–581
9. DNA recommendations – 1992 report concerning recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of PCR-based polymorphisms. Int J Legal Med 105:63–64
10. DNA recommendations – 1994 report concerning further recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of PCR-based polymorphisms. Int J Legal Med 105:159–160
11. Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr W, Olaisen B (1997) DNA recommendations - further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. Int J Legal Med 110:175–176
12. Nomenclature Committee of the International Union of Biochemistry (1985) Nomenclature for incompletely specified bases in nucleic acid sequences. Recommendations 1984. Eur J Biochem 150:1–5
13. Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372
14. Krüger J, Fuhrmann W, Lichte KH, Steffens C (1968) Zur Verwendung des Polymorphismus der sauren Erythrocytenphosphatase bei der Vaterschaftsbegutachtung. Dtsch Z Gerichtl Med 64:127–146
15. Fisher R (1951) Standard calculations for evaluating a blood group system. Heredity 5:95–102
16. Urquhart A, Kimpton CP, Downes TJ, Gill P (1994) Variation in short tandem repeat sequences – a survey of twelve microsatellite loci for use as forensic identification markers. Int J Legal Med 107:13–20
17. Klintschar M, Kubat M (1995) A study of the short tandem repeat systems HUMVWA and HUMTH01 in an Austrian population sample. Int J Legal Med 107:329–330
18. Klintschar M, Neuhuber F (1998) A study on the short tandem repeat system ACTBP2(SE33) in an Austrian population sample. Int J Legal Med 111:46–48
19. Neuhuber F, Klintschar M, Radacher M (1998) A collaborative genetic study on the STR system FGA in two Austrian population samples. Forensic Sci Int 91:1–6
20. Lareu MV, Pestoni C, Schürenkamp M, Rand S, Brinkmann B, Carrancedo A (1996) A highly variable STR at the D12S391 locus. Int J Legal Med 109:134–138
21. Klintschar M, Ricci U, Al Hammadi N, Reichenpfader B, Ebner A, Giovanucci Uzielli ML (1998) Genetic variation at the STR loci D12S391 and CSF1PO in four populations from Austria, Italy, Egypt and Yemen. Forensic Sci Int in press)
22. Puers C, Hammond HA, Jin L, Caskey CT, Schumm JW (1993) Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTHO1 [AATG]n and reassignment of alleles in population analysis by using a locus-specific allelic ladder. Am J Hum Genet 53:953–958
23. Wiegand P, Budowle B, Rand S, Brinkmann B (1993) Forensic validation of the STR systems SE33 and TC11. Int J Legal Med 105:15–320