Int J Legal Med (1997) 110:22–26

ORIGINAL ARTICLE

N. Harashima · C. Liu · Y. Katsuyama · M. Ota
H. Fukushima

Sequence variation of allele 27 at the D1S80 locus

Received: 7 February 1996 / Received in revised form: 29 October 1996

Abstract In 180 unrelated Japanese individuals 18 examples of allele 27 were detected at the locus D1S80 (MCT118). On 6% polylacrylamide gels 5 out of these 18 alleles were found to migrate between allele 26 and allele 27, but closer to allele 27, and thus were labelled variants of allele 27. All 18 examples of allele 27 were sequenced and the results were compared. Although all had the same number of base pairs (578 bp) the five variants could be subdivided into three types V1, V2 and V3. The variants and the standard allele were composed of the same kinds of repeat units, but the order of arrangement was different. We investigated whether it was possible to distinguish the standard allele 27, and the variants V1, V2, and V3 by PCR-RFLP. EcoRII and MspI which have restriction sites within the repeat units were adopted as restriction enzymes. The variants could be discriminated from each other after treatment of the PCR fragments with EcoRII or MspI, followed by PAG electrophoresis.

Key words D1S80 · Allele variants · Gene sequencing · MspI · EcoRII · Cleavage site

Introduction

A revolution in forensic science has occurred over the past few years as a result of the development of many new methods for the isolation and subsequent analysis of DNA. In particular, genotyping of D1S80 (MCT118) (Cosso and Reynolds 1995), HLA-DQα (Sato et al.1995) and five different loci (LDLR, GYPA, HBGG, D7S8 and GC) (Budowle et al. 1995a) has been greatly facilitated by the development of PCR-based kits. As a result, forensic identification based on DNA analysis by these commercially available kits, has been used as physical evidence in the investigation of crimes such as assault, sexual assault, and homicide, as well as for paternity testing.

There have been many reports concerning amplified fragment length polymorphism (Amp-FLP) (Boerwinkle et al.1989; Horn et al.1989; Kasai et al.1990; Rand et al. 1992) and one of the most important and widely studied systems is located at the D1S80 locus.

D1S80 is a variable number of tandem repeat (VNTR) locus generally with a repeat size of 16 base pairs (Kasai et al. 1990). Because of its good distribution in almost all populations and its high heterozygosity index and discrimination power values (Sugiyama et al. 1993; Budowle et al. 1995b; Sepulchre et al. 1995), this system has been applied in forensic laboratories worldwide.

Recently, it has been reported that there are large alleles (Kadasi and Bohusova 1995) and variants (or interalleles) (Skowasch et al. 1992; Alonso et al. 1993; Huang et al. 1994; Nagai et al. 1994) that do not match the commercially supplied allelic ladder. Skowasch et al. (1992) found five interalleles located near alleles 17, and 22–25, while Nagai et al. (1994) found interalleles of alleles 22, 27 and 31. These microvariations could be due to sequence variations or slight differences in the size of alleles. However, no sequencing data for these interalleles has yet been reported.

Allele 27 has a higher frequency in Japanese (Sugiyama et al. 1993) and other Oriental populations compared with Caucasians, Hispanics and African Americans(Budowle et al. 1995b). Therefore, our initial attempt was to analyse the allele 27 fragment in detail since little information is available.

In this paper, we will report on the variants of allele 27 and their sequencing data, based on our previous investigation of D1S80 polymorphism in the Japanese population.

N. Harashima · C. Liu · Y. Katsuyama · M. Ota
H. Fukushima (✉)
Department of Legal Medicine,
Shinshu University School of Medicine,
Matsumoto, Nagano 390, Japan

Materials and methods

DNA extraction

Genomic DNA from unrelated Japanese individuals was isolated from blood using proteinase K/SDS lysis and the phenol extraction method.

PCR amplification

PCR was performed in accordance with our previous method (Sugiyama et al. 1993). The samples were amplified in a 25 µl reaction mixture using Promega Taq DNA polymerase and corresponding buffer in a DNA Thermal Cycler 480 (Perkin Elmer). Amplification conditions were 94° C for 1 min, 65° C for 1 min, 72° C for 2 min for 30 cycles.

Allele typing on agarose gel and native PAG

Amplified products were run on a native PAG (T 6%, C 2%, 1 mm thick, 30 cm long) or a 2% agarose gel (4 mm thick, 20 cm long) together with the commercially supplied allelic ladder (Perkin Elmer: AmpliFLP D1S80 Allelic Ladder) with 0.09 M TBE (pH 8.0) as running buffer. For vertical PAG electrophoresis the running conditions were 350 V for 6 h at 35° C and for agarose gels 50 V for 20 h. Gels were stained with ethidium bromide and photographed under ultraviolet transillumination.

Cloning and sequencing

Following electrophoresis of the amplified products on PAG or agarose gels, the target alleles stained with ethidium bromide were excised from the gels and extracted by the electroelution method (Geno Technology, Inc., GeneCAPSULE). The purified PCR product was ligated to the pBlue Script II SK+ vector at an EcoRV site by the T-A cloning method (Holton and Graham1990; Marchuk et al. 1990), and then transformed into NovaBlue Competent Cell. After the cells were cultured on LB plates, individual white colonies were selected and cultured in a LB medium overnight. Plasmids were isolated from the culture by the alkaline extraction method.

For the sequencing reaction, 5 µg plasmid DNA dissolved in water was used as the template and T3 or T7 were applied as primers. The amplification conditions were 96° C for 15 s, 50° C for 15 s, 60° C for 4 min and 25 cycles using a DNA Thermal Cycler 480 (Perkin-Elmer). Sequencing was performed in 6% denaturing polyacrylamide gels in TBE buffer using a Dyedeoxy Terminator Sequencing Kit on a 373 ABI sequencer, and data analysis was automatically performed using Data Collection and SeqEd software (Applied Biosystems, Foster City, CA).

Direct sequencing

Following electrophoresis of the amplified PCR products, the target alleles were excised from PAG and extracted by electroelution. These alleles were then used directly as the template for the sequencing reaction after concentrating in Microcon-100 (Amicon, Beverley, USA). The primers were the same as those applied for PCR. The sequencing reaction conditions and the analysis method were the same as those described.

PCR-RFLP

The target alleles isolated using the above method were re-amplified and then digested with EcoRII and MspI at 37° C overnight. The digested fragments were electrophoresed on a 12% PAG (30 cm in length) in TBE buffer at 200 V for 40 min. The band pattern was viewed following ethidium bromide staining under UV. pBR322 HaeIII was used as a size standard.

Results and discussion

Allele and genotype frequencies for a highly polymorphic amplifiable VNTR locus, D1S80 have been determined in many populations by the amplified fragment length polymorphism technique. According to a recent report (Sato et al. 1995) 29 alleles were observed ranging from allele 14 to allele 42. Moreover, 250 out of 435 possible genotypes were detected, and high values were obtained for heterozygosity and the power of discrimination.

In addition to being a genetic marker for forensic testing and paternity determination, the D1S80 locus is useful for population genetic research due to its high degree of polymorphism. Allele frequencies of the D1S80 locus can be classified into three main racial groups (Sato et al. 1995), Caucasoid, Negroid, and Mongoloid by cluster analysis. Oriental and black groups can be clearly distinguished from the Caucasoid cluster.

In general, the typing of unknown samples can be performed on PAG by comparing their bands with those of an "allelic ladder", which is a combination of the common alleles of a particular VNTR (or STR) locus. However, as has been reported, many variants (or interalleles) exist in

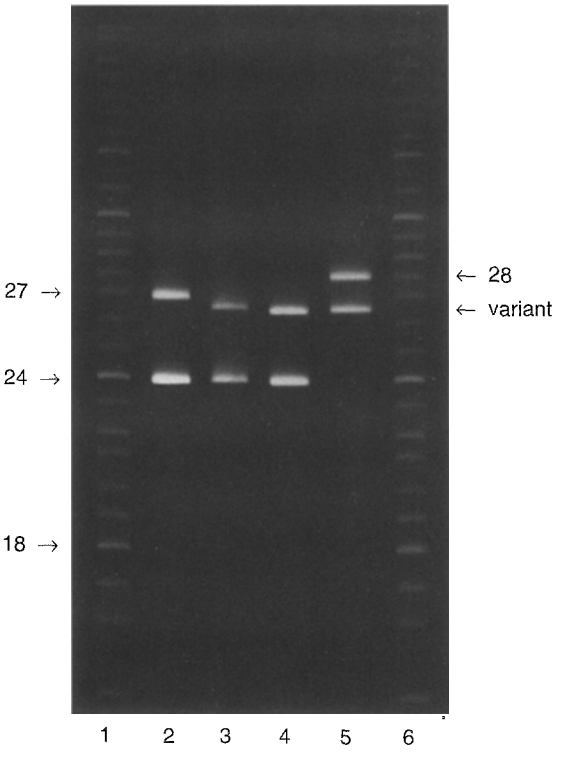

Fig. 1 Electrophoretic band patterns of allele 27 and variants on a 6% polyacrylamide gel. Lanes 1 and 6: AmpliFLP D1S80 allelic ladder, lane 2: standard 27/24, lane 3: variant 1/24, lane 4: variant 2/24, lane 5: 28/variant 3

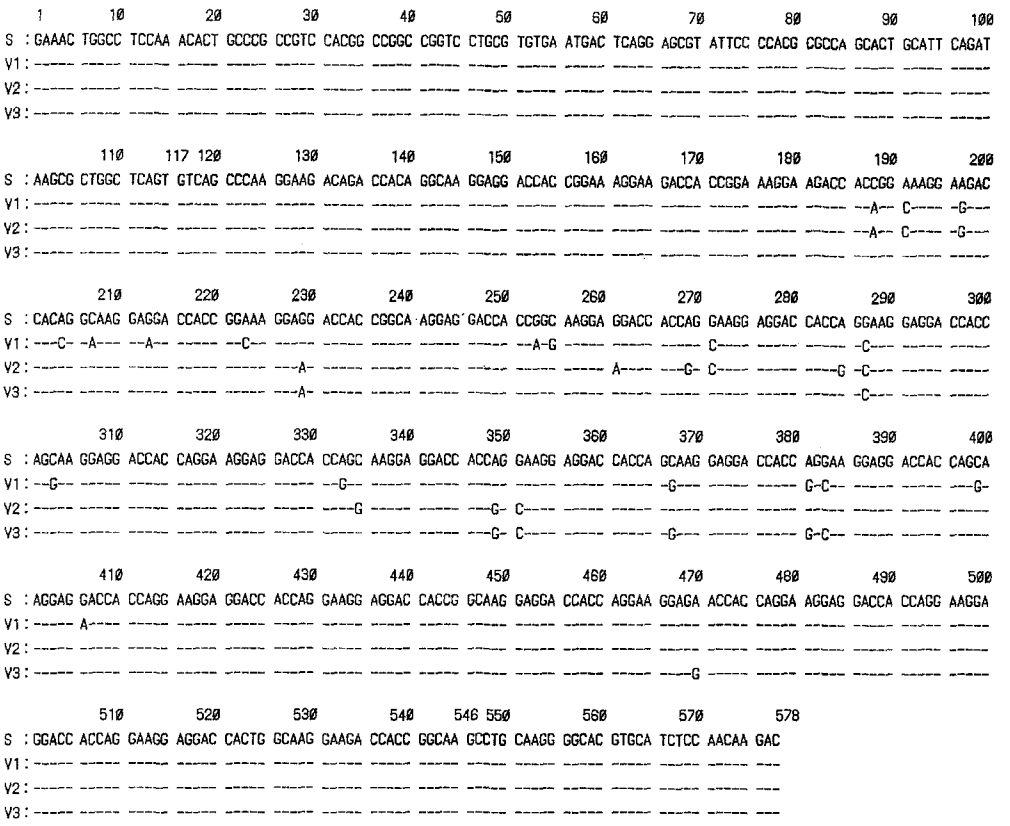

Fig. 2 The sequences of allele 27 and three kinds of variants. bp 1–116: 5′ flanking region, 117–546: repeating units, 547–578: 3′ flanking region. S: standard allele 27, V1–V3: variants

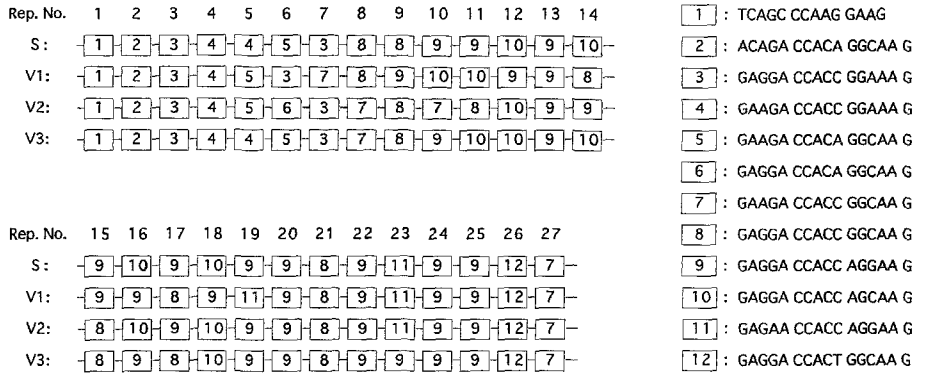

Fig. 3 Sequence structure of the repeat units of the standard allele 27 and the three variants. S: standard allele 27, V1–V3: variants

the D1S80 locus that do not match the allelic ladder, which makes genotype identification difficult.

Clear identification of the sequences of the variants is helpful in screening out these variants from population samples by various techniques, e.g. RFLP (Ota et al. 1991). Recently, a few reports have been made regarding sequencing data on variants of STR systems (Möller and Brinkmann 1994; Barber et al. 1996).

In this study 18 examples of allele 27 were detected in 180 unrelated Japanese individuals. On agarose gels, all of these alleles showed the same migration rate, but on 6% PAG, five alleles migrated differently from the majority of alleles 27. These alleles were found to migrate between allele 26 and allele 27, but closer to allele 27, as shown in Fig. 1, and thus were labelled variants of allele 27. The genotypes of the samples containing the five variants were 24–27 (four samples) and 27–28 (one sample). Sequencing following cloning and direct sequencing of the 18 alleles 27 were carried out in order to determine the sizes and sequences of the standard and variant alleles 27. For each sample, at least three clones were tried and direct sequencing was also repeated at least three times. All the clones from one sample showed the same result, and the sequences obtained from cloned fragment sequencing and direct sequencing appeared to be the same. The sequences of the 13 alleles 27, which were exactly the same as the standard allele 27, were consistant with the data appearing in the manufacturer's manual (Perkin Elmer) for allele 18 and the data of Kasai et al. (1990). The five variants had the same number of base pairs (578 bp) as the standard al-

Table 1 Predicted fragment sizes following restriction endonuclease digestion

Predicted fragment length (bp)	EcoRII				MspI			
	S	V1	V2	V3	S	V1	V2	V3
317			+					
269	+			+				
253		+						
192					+			
115					+	+	+	+
96					+	+	+	+
70	+	+	+	+				
64						+	+	
48		+	+	+		+	+	
39					+	+	+	+
38					+	+	+	+
32	+	+	+	+	+	+		+
16	+	+	+	+	+	+	+	+
4					+	+	+	+

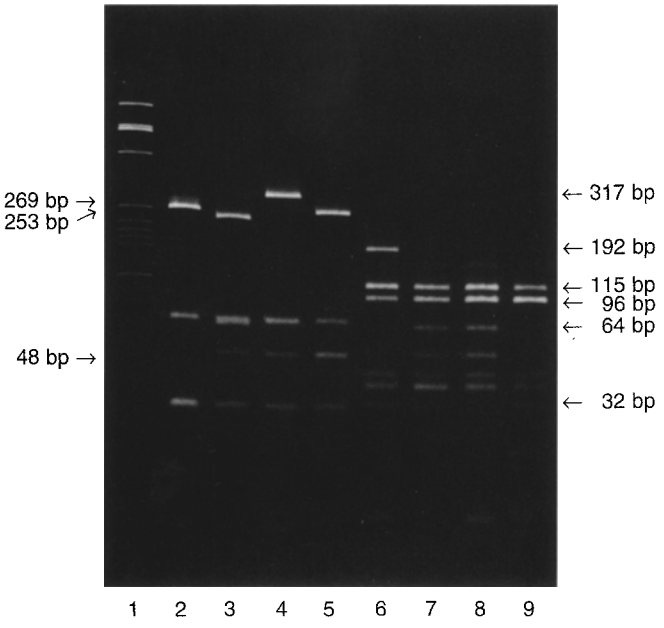

Fig. 4 Electrophoretic band patterns of the standard allele 27 and the three variants following treatment with EcoRII or MspI. Lane 1: pBR322-HaeIII, lane 2–5: EcoRII treatment, lane 6–9: MspI treatment

lele 27, but the sequences were different. According to their sequence differences, the five variants could be subdivided into three types (Fig. 2): V1 (two samples), V2 (two samples) and V3 (one sample). The base changes of these variants were found only in the repeating units. The flanking regions of the five variants were the same as those of the standard allele 27. Compared with the standard allele, base substitutions were observed at 18, 12, and 8 sites in V1, V2, and V3, respectively. Figure 3 summarizes the patterns of 12 kinds of 16 bp repeat units and their position in the standard allele 27 and variants. Twelve repeating units of V1, seven of V2 and six of V3 were found to be different from those of the standard allele. From these observations, it is predicted that the variants (or interalleles) of some other allele in the D1S80 locus, e.g. the variants of alleles 17, 22, 23, 24 , 25, or 31, may also be composed of the same kinds of repeating units, although the order of arrangement may be different.

The PCR-RFLP method is an efficient and convenient technique. In this study, we investigated whether it was possible to distinguish the standard allele 27, V1, V2, and V3 from each other by PCR-RFLP. The differences among the EcoRII or MspI restriction sites of these alleles were observed only in the repeating units. The predicted fragment length of the allele 27 digested by these two kinds of enzymes is shown in Table 1.

Following treatment with EcoRII, four fragments were predicted to appear in the standard allele 27, while five fragments were predicted to be seen in the variants (Table 1). From the PAG band patterns (Fig. 4), the migration differences among bands of 317 bp, 269 bp and 253 bp can be used to distinguish the variants from each other, and to distinguish V1 and V2 from the standard. Although a 269 bp band appeared in both the standard allele and V3, the standard and V3 could still be distinguished by the absence of a 48 bp band in the standard allele.

Following treatment with MspI, nine fragments were predicted to be seen in V1, while eight and seven fragments were predicted to be observed in the standard allele, V2, and V3, respectively (Table 1). From the PAG band pattern (Fig. 4), the 192 bp band was seen only in the standard allele. V3 could be distinguished from V1 and V2 by the absence of the 64 bp band and the 48 bp band while V2 could be distinguished from V1 by the absence of the 32 bp band. These electrophoretic band patterns after treatment with EcoRII or MspI, are not compatible with the predicted fragment lengths shown in Table 1.

To summarize the preceding results, treatment of the PCR fragments with EcoRII and MspI followed by PAG electrophoresis could be used to discriminate between the variants of allele 27. In the D1S80 locus, the variants (or interalleles) of some other alleles have also been reported. Although it is necessary to understand the full sequences of these variants, the identification of these variants can be easily made by PCR-RFLP. Moreover, PCR-RFLP can provide an effective and easy means of subtyping D1S80 alleles.

Acknowledgement We would like to thank Dr. Kamijo of the Biochemistry department Shinshu University for technical support.

References

Alonso A, Martin P, Albarran C, Sancho M (1993) Amplified fragment length polymorphism analysis of the VNTR locus D1S80 in Central Spain. Int J Legal Med 105:311–314

Barber MD, McKeown BJ, Parkin BH (1996) Structural variation in the alleles of a short tandem repeat system at the human alpha fibrinogen locus. Int J Legal Med 108:180–185

Boerwinkle E, Xiong W, Fourest E, Chan L (1989) Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction; application to the apoprotein B3′ hypervariable region. Proc Natl Acad Sci USA 86:212–216

Budowle B, Lindsey JA, DeCou JA, Koos BW, Giusti AM, Comey CT (1995a) Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8 and GC (PM loci) and HLA-DQ alpha using a multiplex amplification and typing procedure. J Forensic Sci 40:45–54

Budowle B, Baechtel FS, Smerick JB, Presley KW, Giusti AM, Parsons G, Alevy MC, Chakraborty R (1995b) D1S80 population data in African Americans, Caucasians, Southeastern Hispanics, Southwestern Hispanics, and Orientals. J Forensic Sci 40:38–44

Cosso S, Reynolds R (1995) Validation of the AmpliFLP™ D1S80 PCR amplification kit for forensic casework analysis according to TWGDAM guidelines. J Forensic Sci 40:424–434

Holton TA, Graham MW (1990) A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. Nucleic Acids Res 19:1156

Horn GT, Richards B, Klinger KW (1989) Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. Nucleic Acids Res 17:2140

Huang NE, Chakraborty R, Budowle B (1994) D1S80 allele frequencies in a Chinese population. Int J Legal Med 107:118–120

Kadasi L, Bohusova T (1995) A new extremely large allele at the D1S80 (iMCT118) locus. J Forensic Sci 40:906–907

Kasai K, Nakamura Y,White R (1990) Amplification of variable number of tandem repeat (VNTR) locus(pMCT118) by polymerase chain reaction (PCR) and its application to forensic science. J Forensic Sci 35:1196–1200

Marchuk D, Drumm M, Saulino A, Collins FS (1990) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res 19:1154

Möller A, Brinkmann B (1994) Locus ACTBP2 (SE33) sequencing data reveal considerable polymorphism. Int J Legal Med 106:262–267

Nagai A, Yamada S, Bunai Y, Ohya I (1994) Analysis of the VNTR locus D1S80 in a Japanese population. Int J Legal Med 106:268–270

Ota M, Seki T, Nomura N, Sugimura K, Mizuki N, Fukushima H, Tsuji K, Inoko H (1991) Modified PCR-RFLP method for HLA-DPB1 and -DQA1 genotyping. Tissue Antigens 38:60–71

Rand S, Puers C, Skowasch K, Wiegand P, Budowle B, Brinkmann B (1992) Population genetics and forensic efficiency data of 4 AMPFLP's. Int J Legal Med 104:329–333

Sato H, Kasai K, Seta S (1995) Assessment on frequency distribution of MCT118 and HLADQα alleles and genotypes in Japanese population along with the comparison with other population data. Proceedings of international workshop on forensic DNA analysis of evidential samples, Tokyo, December 12–13, pp 139–142

Sepulchre MA, Wiegand P, Brinkmann B (1995) D1S80 (pMCT118); analysis of 3 ethnic subpopulations living in Brussels. Int J Legal Med 108:45–47

Skowasch K, Wiegand P, Brinkmann B (1992) pMCT118: a new allelic ladder and an improved electrophoretic separation lead to the demonstration of 28 alleles. Int J Legal Med 105:165–168

Sugiyama E, Honda K, Katsuyama Y, Uchiyama S, Tsuchikane A, Ota M, Fukushima H (1993) Allele frequency distribution of the D1S80 (pMCT118) locus polymorphism in the Japanese population by the polymerase chain reaction. Int J Legal Med 106:111–114