Gotaro Watanabe,[1] *M.Sc.; Kazuo Umetsu,*[1] *Ph.D.; Isao Yuasa,*[2] *Ph.D.; and Tsuneo Suzuki,*[1] *M.D.*

Simultaneous Determination of STR Polymorphism and a New Nucleotide Substitution in Its Flanking Region at the CD4 Locus

REFERENCE: Watanabe G, Umetsu K, Yuasa I, Suzuki T. Simultaneous determination of STR polymorphism and a new nucleotide substitution in its flanking region at the CD4 locus. J Forensic Sci 1998;43(4):733–737.

ABSTRACT: In the course of the investigation of a pentanucleotide repeat polymorphism at the human CD4 locus, a C-A transversion was found at the position corresponding to the 3′ end of the original forward primer presented by Edwards et al. (1). In the present study, the simultaneous determination of the new sequence polymorphism and the pentanucleotide repeat polymorphism at the CD4 locus was attempted. To achieve this purpose, we adopted amplified product length polymorphism (APLP) analysis and designed some new allele-specific forward primers tagged with non-complementary nucleotides differing in length. A total of 646 DNA samples from peripheral blood of Japanese, Chinese and German populations were investigated. Although the C-A transversion was restricted to CD4*5, a new subtype allele with A and 5 repeats, designated CD4*5A, was observed at polymorphic frequencies in the three populations. The simultaneous genotyping by APLP analysis resulted in dramatically increased heterozygosity and discriminating power of the human CD4 locus.

KEYWORDS: forensic science, DNA typing, short tandem repeat, CD4 locus, population study, C-A transversion, pentanucleotide repeat

Short tandem repeat (STR) polymorphisms are microsatellite loci containing repeat motifs 2 to 5 nucleotides long. Since STR sequences may be amplified from small amounts of highly degraded template DNA, they are becoming more widely used for genetic characterization of samples. The pentanucleotide repeat polymorphism at the human CD4 locus was first described by Edwards et al. (1). This is a highly polymorphic STR on chromosome 12p, and many reports on allele frequency in various populations have been published (2–8). Although some alleles displayed one repeat unit sequence CTTTC instead of the consensus sequence TTTTC (6,9), nucleotide substitutions in the 5′ flanking region to STR have not been reported thus far.

The present paper describes a novel C-A nucleotide substitution at the position corresponding to the 3′ end of the original forward primer (1). An amplified product length polymorphism (APLP) analysis was used for the simultaneous determination of the new sequence polymorphism and the conventional STR polymorphism and it was applied to the investigation of CD4 allele frequencies in Japanese, Chinese and German populations.

[1] Postgraduate student, associate professor, and professor, respectively, Department of Forensic Medicine, Yamagata University School of Medicine, Yamagata 990–9585, Japan.

[2] Assistant professor, Department of Legal Medicine, Tottori University School of Medicine, Yonago 683-8503, Japan.

Received 29 May 1997; and in revised form 6 Oct. 1997; accepted 8 Dec. 1997.

Materials and Methods

Isolation of DNA

Genomic DNAs were isolated from peripheral blood of 333 unrelated Japanese (Yamagata), 143 Chinese (Shenyang) and 170 German (Munich) individuals. DNA was prepared according to the standard phenol-chloroform extraction procedure (10).

Preparation of Primers and PCR

Five oligonucleotide primers were newly designed for the nucleotide sequences of the CD4 gene (GenBank M86525). Oligonucleotides were synthesized with the DNA synthesizer Gene Assembler Plus (Pharmacia, Freiburg, GER) and cleaned with NAP-10 columns (Pharmacia). Among them two allele-specific primers (CD4-A and CD4-C) were prepared to have an allele-specific nucleotide (A or C) at 3′ terminus and non-complementary nucleotide sequences at 5′ terminus according to the strategy of APLP analysis (11). The nucleotide sequences of primers used and PCR conditions for each primer set are shown in Tables 1 and 2, respectively. PCR was performed in a total reaction volume of 20 μL containing 5 to 20 ng genomic DNA, 10 mM Tris-HCl buffer (pH = 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM each dNTP, each primer set and 0.5 U Taq polymerase (Takara, Kyoto, Japan). Each mixture was subjected to 30 amplification cycles using a DNA Thermal Cycler model PJ2000 (Perkin-Elmer, Norwalk, CT). After the last extension, the samples were stored at 4°C until analysis.

Analysis of PCR Products

Aliquots (3 μL) from each amplification reaction mixture were resolved on a native polyacrylamide gel (T = 10%, C = 5%, sized 8 cm × 7 cm × 1 mm) in 0.375 M Tris-HCl buffer (pH = 8.9) with 0.025 M Tris-glycine buffer (pH = 8.3) as a tank buffer for 70 min at 200 V. The gel was stained with ethidium bromide.

Direct Sequencing

DNA sequencing was performed using PCR products under the PCR condition 2 in Table 2. To obtain a single allelic product, the amplified fragments were electrophoretically separated on native polyacrylamide gels. Each allelic product purified from a gel was sequenced from both ends with a Taq Dye Deoxy Terminator Cycle

Sequencing Kit (Perkin-Elmer), using the PCR primers as sequencing primers. Fluorescent labeled reaction products were analyzed with an automated DNA Sequencer 373A (Perkin-Elmer).

Results and Discussion

Two samples of CD4 type 5–10 from Japanese individuals were amplified under the PCR condition 1 in Table 2, i.e., using the original primer set of Edwards et al. (1) at nine annealing temperatures differing by 2°C from 52°C to 68°C. One sample showed that both 5 and 10 alleles were equally amplified at all annealing temperatures tested (Fig. 1A). In another sample, allele 5 was fainter than allele 10 and it dropped out when annealing temperatures were higher than 64°C (Fig. 1B). We sequenced these two DNA samples using PCR products under the PCR condition 2 in Table 2 to characterize the difference between two types of allele 5 (Fig. 2). As shown in Fig. 3, we found a nucleotide substitution at the position corresponding to the 3′ end of the forward primer for CD4 locus. The normally amplified type of allele 5 was a cytosine (C) in accordance with the sequence data reported, while the poorly amplified type had an adenine (A). Thus, CD4 locus was divided into type A and type C alleles. In PCR amplification used Edwards's primer set (1), type A allele is more difficult to be amplified than type C allele. Therefore, when type A allele was not amplified in PCR amplification, we might mistake heterozygous for homozygous due to the absence of type A allele band. In fact, Casarino et al. (8) reported that the homozygotes were generally more frequent than expected.

To determine the new nucleotide substitution polymorphism, an APLP analysis was carried out. The PCR was performed using two new allele-specific forward primers, CD4-A and CD4-C, on condition 3 in Table 2. This method detected the nucleotide substitution type (A or C) and the conventional pentanucleotide repeat polymorphism at the human CD4 locus simultaneously, because the primer set amplifies a region including STR. A total of nine alleles were observed in the Japanese, Chinese and German populations. Fig. 4 shows that each allele was clearly and unambiguously distinguished by the corresponding allele-specific band. Since the two allele-specific primers differed by 3 nucleotide in length from each other, the PCR products of type A were 3 base pairs shorter than those of type C allele. Type A was observed only in the allele 5, and allele 5 could be classified into alleles 5A and 5C. The other alleles 6–12 were type C. There is strong linkage disequilibrium between the repeat polymorphism and the C-A transversion. A C-A transversion had occurred in the allele 5 before the divergence of the Europeans and the Asians. The allele 5A increased in ancestral populations by the random genetic drift. However, C-A transversion in alleles other than 5A, crossing-over between the allele 5A and the other alleles, and the tandem repeat expansions of the allele

TABLE 1—*Nucleotide sequences of primers used.*

Name	Sequence
Forward primers (CTTTT strand)	
CD4-1*	5′-TTGGAGTCGCAAGCTGAACTAGC-3′
CD4-2	5′-CTATATGCTACAGATGAAGT-3′
CD4-A	5′-cGaCGCAAGCTGAACTAGA-3′
CD4-C	5′-gcataTCGCAAGCTGAACTAGC-3′
Reverse primers (AAAAG strand)	
CD4-3*	5′-GCCTGAGTGACAGAGTGAGAACC-3′
CD4-4	5′-TTGAGCCCAGGAAGTTGAG-3′
CD4-5	5′-AGTGACAGAGTGAGAACCT-3′

Small letters are non-complementary nucleotides.
* Original primers reported by Edwards et al. (1).

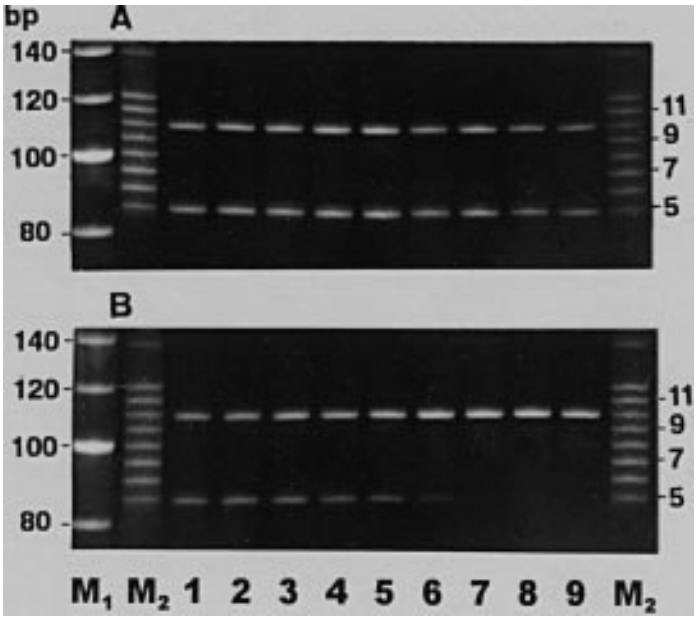

FIG. 1—*Comparison of PCR amplified products of two samples by using Edwards's primer set (1) and annealing temperature at 52°C to 68°C. 1: 52°C, 2: 54°C, 3: 56°C, 4: 58°C, 5: 60°C, 6: 62°C, 7: 64°C, 8: 66°C, 9: 68°C. M_1: size marker, M_2: allelic ladder marker. A: HumCD4 normal type 5–10, B: HumCD4 abnormal type 5–10.*

TABLE 2—*PCR conditions for each primer set.*

No.	Primer Set (Amount)		PCR		
			Denaturation	Annealing	Extension
1.	CD4-1 (8pmol)	CD4-3 (8pmol)	94°C, 0.5 min	52–68°C, 0.5 min	68°C, 2 min
2.	CD4-2 (8pmol)	CD4-4 (8pmol)	94°C, 0.5 min	53°C, 0.5 min	72°C, 2 min
3.	CD4-A (10pmol) CD4-C (5pmol)	CD4-5 (8pmol)	95°C, 0.5 min	54°C, 0.5 min	72°C, 0.5 min

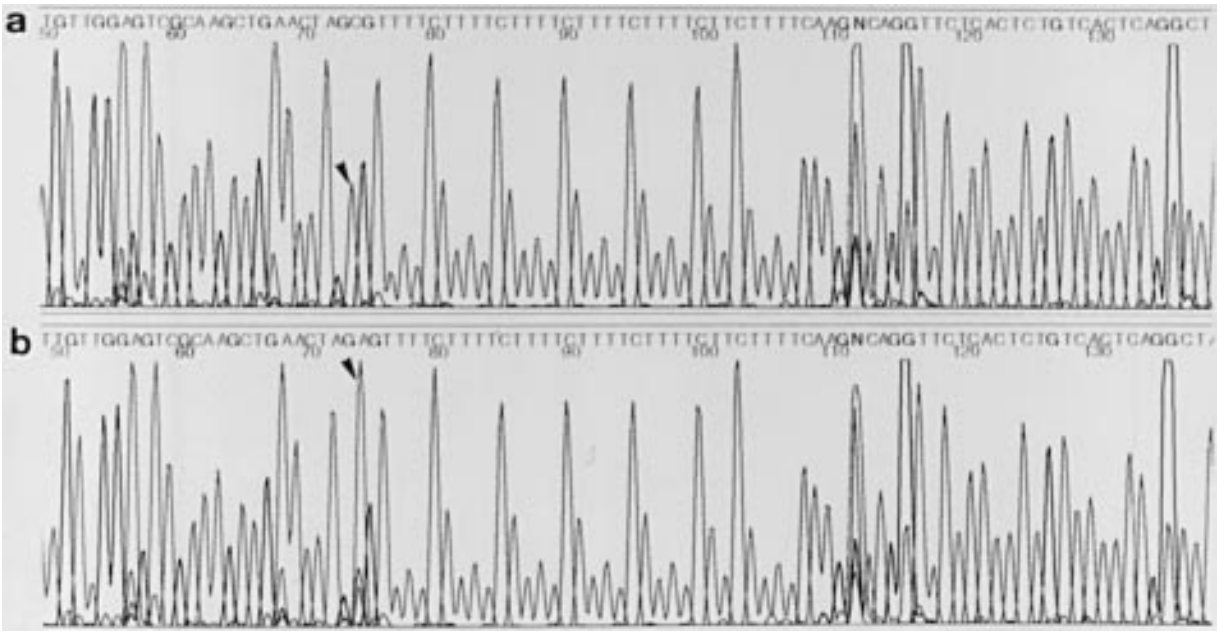

FIG. 2—*Sequence analysis of CD4 5A allele (a) and CD4 5C allele (b). Arrowheads indicate mutation sites.*

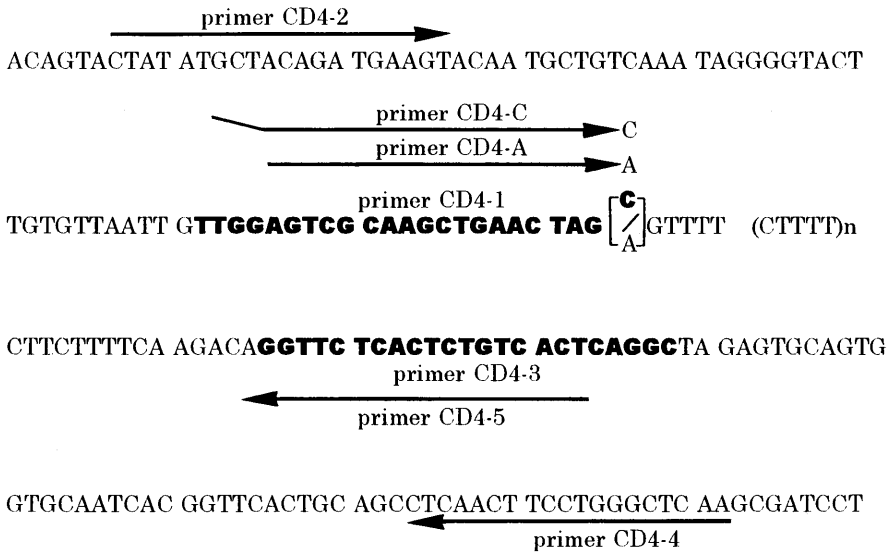

FIG. 3—*Sequence of the CD4 locus and annealing regions for primers. The sequences in bold type correspond to primers presented by Edwards et al. (1).*

5A by aberrations such as slippage during DNA replication have not occurred. The changes of the repeat number in CD4 locus may be regulated by the C-A transversion in the flanking region (12,13). Further studies on Africans and non-human primates are needed to elucidate the mechanism of the strong linkage disequilibrium.

The distribution of genotypes and allele frequencies is shown in Table 3. In each population good agreement was found between the observed and expected values according to the Hardy-Weinberg equilibrium. The proportions of types C and A in allele 5 were as follow: C = 0.65, A = 0.35 in Japanese; C = 0.64, A = 0.36 in Chinese; and C = 0.33, A = 0.67 in German. In the Japanese and Chinese populations, the proportion of the type C to A in allele 5 was much higher than that in the German populations. The allele frequencies in the Japanese population were similar to

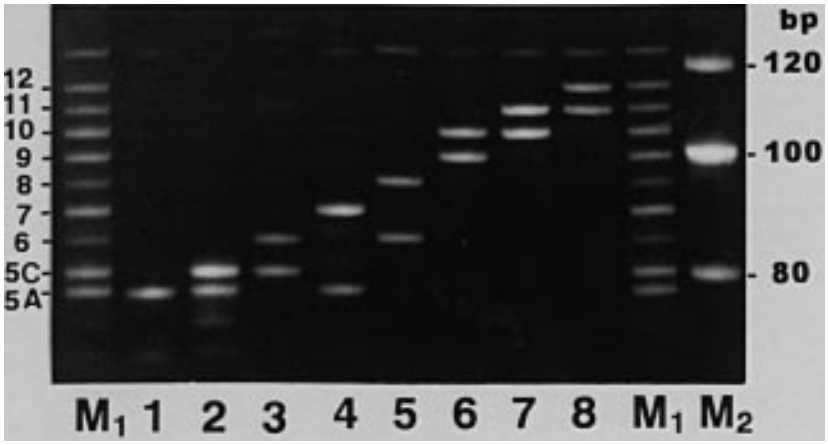

FIG. 4—*Banding patterns of PCR products by using primer set (CD4-A, CD4-C and CD4-5). 1: 5A-5A, 2: 5A-5C, 3: 5C-6, 4: 5A-7, 5: 6–8, 6: 9–10, 7: 10–11, 8: 11–12, M_1: allelic ladder marker, M_2, size marker.*

TABLE 3—*Distribution of humCD4 allele frequencies in Japanese, Chinese and German populations.*

Allele	(Size)	Japanese	Chinese	German
5A	(77bp)	0.2432	0.2098	0.2157
5C	(80bp)	0.4505	0.3706	0.1078
6	(85bp)	0.0135	0.0070	0.3471
7	(90bp)	0.0030	. . .	0.0088
8	(95bp)	. . .	. . .	0.0029
9	(100bp)	0.0030	0.0070	0.0029
10	(105bp)	0.2763	0.4056	0.2853
11	(110bp)	0.0105	. . .	0.0235
12	(115bp)	. . .	. . .	0.0059
n		333	143	170
χ^2		4.32	3.47	7.52
df		6	6	10
p		0.5–0.7	0.7–0.8	0.5–0.7
Heterozygosity				
observed		0.63	0.68	0.75
expected		0.66	0.65	0.74
PD		0.83	0.80	0.88

those in the Chinese population. In the Asian populations, allele 5 showed the highest frequency with approximately 70%, and therefore the STR variability in the Asian populations was much lower than in the African and European populations (2,3). The application of the simultaneous genotyping method to the CD4 locus presented here increased the heterogeneity from 0.44 to 0.66 in the Japanese and from 0.50 to 0.65 in the Chinese. It also increased from 0.70 to 0.74 in the German. Similarly, an average probability of discrimination between genotypes increased from 0.61 to 0.83 in the Japanese, from 0.64 to 0.80 in the Chinese, and from 0.84 to 0.88 in the German. In addition, the method is very simple and rapid, because it can be performed on a small native polyacrylamide gel requiring a short time. Since the value of the CD4 locus evidently has been enhanced, it is more useful for forensic identification and paternity testing in many populations.

References

1. Edwards MC, Clemens PR, Tristan M, Pizzuti A, Gibbs RA. Pentanucleotide repeat length polymorphism at the human CD4 locus. Nucl Acid Res 1991;19:4791.
2. Wall WJ, Williamson R, Petrou M, Papaioannou D, Parkin BH. Variation of short tandem repeats within and between populations. Hum Mol Genet 1993;2:1023–9.
3. Brinkmann B. The STR approach. In: Carracedo A, Brinkmann B, Bär W, editors. Advance in Forensic Haemogenetics. Proceedings of the 16th Congress of the International Society for Forensic Haemogenetics; 12–16 Sept. 1995; Santiago. Berlin: Springer-Verlag, 1996;6:41–51.
4. Glock B, Schwartz DWM, Dauber EM, Schwartz-Jungl EM, Mayr WR. AmpFLP-typing for the humCD4 STR polymorphism in an Austrian caucasoid population sample: sequence data and allele distribution. In: Carracedo A, Brinkmann B, Bär W, editors. Advance in Forensic Haemogenetics. Proceedings of the 16th Congress of the International Society for Forensic Haemogenetics; 12–16 Sept. 1995; Santiago. Berlin: Springer-Verlag, 1996;6:55–7.
5. Weispfenning R, Luckenbach C, Ritter H. Non-automatic multiplex analysis and detection of six STR-loci: HPRT, FABP2, CD4, F13A1, CYP19 and LPL. In: Carracedo A, Brinkmann B, Bär W, editors. Advance in Forensic Haemogenetic. Proceedings of the 16th Congress of the International Society for Forensic Haemogenetics; 12–16 Sept. 1995; Santiago. Berlin: Springer-Verlag, 1996; 6:223–5.
6. Glock B, Schwartz DWM, Schwartz-Jungl EM, Mayr WR. Sequence determination of an allelic ladder for the STR polymorphism at the CD4 locus and application of the ladder in testing an Austrian Caucasian population sample. Forensic Sci Int 1995;78: 125–30.
7. Neuhuber F, Radacher M, Krasa B. F13B and CD4 allele frequencies in an Austrian population sample. Int J Legal Med 1996;108: 227–8.
8. Casarino L, Mannucci A, Kimpton CP, Presciuttini S, Bruni G, Costa MG, et al. Forensic evaluation of HUMCD4: an Italian database. Int J Legal Med 1996;109:49–51.
9. Urquhart A, Kimpton CP, Downes TJ, Gill P. Variation in short tandem repeat sequences—a survey of twelve microsatellite loci for use as forensic identification markers. Int J Leg Med 1994;107: 13–20.
10. Smith JC, Newton CR, Alves A, Anwar R, Jenner D, Markham

AF. Highly polymorphic minisatellite DNA probes. Further evaluation for individual identification and paternity testing. J Forensic Sci 1990;30:3–18.
11. Watanabe G, Umetsu K, Yuasa I, Suzuki T. Amplified product length polymorphism (APLP): a novel strategy for genotyping the ABO blood group. Hum Genet 1997;99:34–7.
12. Jeffreys AJ, Tamaki K, Macleod A, Monckton DG, Neil DL, Armour JAL. Complex gene conversion events in germline mutation at human minisatellites. Nature Genet 1994;6:136–45.
13. Monckton DG, Neumann R, Guram T, Fretwell N, Tamaki K, Macleod A, et al. Minisatellite mutation rate variation associated with a flanking DNA sequence. Nature Genet 1994;8:162–70.

Additional information and reprint requests:
Tsuneo Suzuki
Department of Forensic Medicine
Yamagata University School of Medicine
Yamagata 990-9585, Japan