J Forensic Sci, Jan. 2003, Vol. 48, No. 1
Paper ID JFS2002006_481
Available online at: www.astm.org

TECHNICAL NOTE

Takashi Yoshimoto,[1] *B.S.; Toshimichi Yamamoto,*[1] *Ph.D.; Masaki Mizutani,*[1] *M.D; Rieko Uchihi,*[1] *Ph.D.; Hiroyuki Ohtaki,*[1] *B.S.; Yoshinao Katsumata,*[1] *M.D.; Worawee Waiyawuth,*[2] *M.D.; and Sirirurg Songsivilai,*[3] *M.D.*

A Novel Fluorescent Quadruplex STR Typing System and the Allele Frequency Distributions in a Thai Population

ABSTRACT: We have previously reported a new triplex amplification and typing system by silver staining for three short tandem repeat (STR) loci, 9q2h2 (D2S3020), D15S233, and D14S299 without "microvariant" alleles such as .1, .2, and, .3 alleles in the Japanese population. In the present study, we established a new quadruplex system with an additional locus D7S809 using primer sets labeled with fluorescent multi-color dyes. Using this system, we genotyped 183 Thai people, found only one "microvariant" allele (allele 20.2) at D7S809, and calculated allele frequencies and some statistical properties at these four STR loci. From these allele frequencies at four STR loci, we performed three statistical analyses including a homozygosity test, a likelihood ratio test, and an exact test for Hardy-Weinberg equilibrium (HWE). Deviations from HWE ($p < 0.05$) were observed only in the two tests at the locus D7S809.

In the present study, we compared the allele frequencies at these four loci in the Thai population to those in the Japanese population described previously. Consequently, all observed heterozygosities and power of discrimination (PD) at those loci in the Thai population were higher than 0.8 and 0.9, respectively, and all statistical values for discriminating power in the Thai population were slightly higher than those in the Japanese population. The combined paternity exclusion rate (combined PE) in the Thai population (0.978) was almost the same as that in the Japanese population (0.971). Therefore, this novel PCR amplification and typing system for four STR loci would be a convenient and informative DNA profiling system in the forensic field.

KEYWORDS: forensic science, quadruplex, DNA profiling, STR, population genetics

Currently, forensic scientists are working on DNA polymorphisms involving short tandem repeats (STRs) consisting of two to six base repetitive units. In the United States, Combined DNA Index System (CODIS) core 13 STRs were adopted to construct a huge database for criminal investigation, and some multiplex PCR amplification and typing kits for those STRs are commercially available. Some countries are making their own databases including the CODIS loci and validate their possible application to forensic tests using those kits. One such system composed of 13 STR loci showed an average match probability rarer than one in a trillion (1). This system showed an average exclusion probability of 99.99% with the complete data available for mother, child, and the alleged father (trio) in parentage analysis (1). However, these STR loci selected for CODIS are not as informative in all ethnic populations; for example, in the Japanese population, some of those loci estimated the observed heterozygosities as less than 0.7 (2). Consequently, this system shows exclusion probabilities of as low as 99% with some trios, and sometimes much lower values with the cases lacking mother and shibship tests. Thus, it would be desirable to search for more informative tetranucleotide STR loci for each ethnic group to obtain higher exclusion probabilities in parentage analyses. Distribution of STRs in the human genome has been reported to occur approximately every two Kb with the majority of repeat sequences dinucleotides (3). Tetranucleotide STRs are also abundant (3) and more amenable to amplification with fewer artifacts than the dinucleotides.

Armour et al. reported a rapid method to isolate microsatellite arrays and discovered many polymorphic microsatellite loci in the human genome, some of which were published (4). In previous papers, we evaluated the forensic utility of those four loci, 9q2h2 (D2S3020), D15S233, D14S299, and D7S809, in the Japanese population and demonstrated a high power of discrimination in personal identification (5–8).

The apparent possibility of germ line mutations in such STR systems is not negligible (9) in parentage analyses. When a single exclusion case in paternity testing is encountered, the exclusion probability decreases drastically. Therefore, it is desirable to type other

[1] Research student, associate professor, postgraduate student, assistant professor, research student, and professor, respectively, Department of Legal Medicine and Bioethics, Graduate School of Medicine, Nagoya University, 65 Tsurumai-cho, Showa-ku, Nagoya, Japan.
[2] Professor, Department of Forensic Medicine, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand.
[3] Professor, Department of Immunology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Received 11 Jan. 2002; and in revised form 8 June 2002 and 2 Sept. 2002; accepted 2 Sept. 2002; published 11 Dec. 2002.

STR loci to raise the possibility and to ascertain whether the mutation is significant. Therefore, other efficient STR typing systems would be necessary to resolve those problems and to also apply to more complicated kinships.

In the present study, we present a newly devised quadruplex typing system for all tetranucleotide repeat STR loci using the primer sets labeled with two-colored fluorescent dyes. The usefulness of this system was investigated in the Thai population the same as the Japanese population, by comparing the allele frequency distributions with those of Japanese.

Materials and Methods

Sample Preparations—Fresh blood samples were collected from 183 unrelated healthy Thai individuals (Bangkok City) with informed consent. DNA was extracted from blood buffy coats using a usual organic extraction method (10). The DNA concentration of each sample was quantified fluorometrically with bisbenziamide Hoechest No. 33258 trihydrochloride (SIGMA Chemical Co., St. Louis, MO) using the DyNA Quant™ 200 Fluorometer (Hoefer Pharmacia Biotech Inc. San Francisco, CA).

PCR Amplification and Genotyping—PCR amplification was carried out in a volume of 10 μL containing 4.5 m*M* $MgCl_2$, 0.2 μM of each dNTPs, 15 μM of each primer at D7S809, 10 μM of each primer at D2S3020 and D14S299, 5 μM of primer at D15S233, 0.5 U AmpliTaq DNA Polymerase (Perkin-Elmer, Norwalk, CT), and 1 ng template DNA. Cycling conditions were one cycle at 95°C for 3 min, 30 cycles at 94°C for 1 min and at 60°C for 1 min and at 72°C for 1 min, and one cycle at 72°C for 10 min using the GeneAmp PCR system 9700 (PE Applied Biosystems, Foster City, CA). The sequences of the primers and the fluorescent dyes (FAM and HEX) labeled were as follows: wg1g9a: 5′-GATCGTGCCACTGTACC-3′; HEX-wg1g9b: HEX-5′-CGCGTGTCTAACCTCATGGC-3′; wg1c5a1:′-GATCTCAATAAACATTGATACTGG-3′; FAM-wg1c5b1: FAM-5′-CTGCATGAGCTAAAGCATACTG-3′; FAM-wg1d1a: FAM-5′-CACCACGTCCCGCCTCTAAT-3′; wg1d1b: 5′-CACCACGTCCCGCCTCTAAT-3′; 9q2h2a:5′-ACAGAGCAAGACTCTGTCAC-3′; HEX-9q2h2b: HEX-5′-GTATACACTTGAGCTAGTAGC-3′. The sequences of these primers were described previously (5–8). PCR products were analyzed using the ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems) with GeneScan analysis software 3.1. Each 1.5 μL of PCR product or of the allelic ladder marker mentioned below was loaded in 24.5 μL of deionized formamide and 0.5 μL of GeneScan™-500 [ROX] Size Standard (PE Applied Biosystems). The analyzed data were automatically genotyped using Genotyper 2.5 software (PE Applied Biosystems) with an appropriate template constructed from size data of an allelic ladder marker made by ourselves.

Sequence Analysis and Preparation of Allelic Ladder Markers—Thirteen "nominal allele motif" alleles at D7S809 observed in the Japanese population (Alleles 18 to 31, excluding 22, ranged from 245 to 297 bp) were separated by 6% polyacrylamide gel electrophoresis and silver staining and were re-amplified and sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction kit (PE Applied Biosystems) with the ABI PRISM 377 DNA sequencer (PE Applied Biosystems). Similarly, the alleles newly observed in the Thai population were also sequenced. The 13 "nominal allele motif" alleles were inserted into plasmids (pCR®2.1-TOPO®) included in a TOPO TA Cloning® Kit (Invitrogen, Carlsbad, CA). The ten alleles ranged from 131 to 167 bp (Alleles 12 to 21) at the D2S3020, nine alleles ranged from 195 to 227 bp (Alleles 15 to 23) at the D15S233, and nine alleles ranged from 299 to 331 bp (Alleles 28 to 36) at the D14S299 described previously (11) and were cloned similarly. Each cloned allele at D2S3020, D7S809, D15S233, and D14S299 was amplified with each primer set labeled by FAM, FAM, HEX, and HEX, respectively. Each PCR product was mixed together in a proportion to show their peak heights evenly balanced between 500 and 1300 RFU (relative fluorescent units) on an electropherogram by the Genentic Analyzer 310 as recommended by Griffiths et al. (12). The nomenclature used in the present study followed the recommendations of the DNA commission of IFSH (1994) based on the number of repeats of each allele.

Statistical Analysis—Tests for Hardy-Weinberg equilibrium (HWE) were carried out using a homozygosity test (13), a likelihood ratio test (14), and an exact test (15). The probabilities of the likelihood ratio test and the exact test were estimated based on 10 000 shuffling experiments. Each allele frequency at the four STR loci in Thais was compared with that in Japanese by the Genepop 3.1c software.

Sensitivity Study—Various amounts of high molecular DNA samples (1.0, 0.5, 0.2, 0.1, and 0.05 ng) were prepared as template DNA for sensitivity examination in the present quadruplex system. PCR was performed at 30 and 35 cycles. The results were compared to those from the AmpFℓSTR profiler plus kit (PE Applied Biosystems).

Results and Discussion

We constructed a novel multiplex amplification and typing system for four STR loci (D2S3020, D7S809, D15S233, and D14S299) and genotyped 183 Thai individuals with this system automatically using the Genotyper by running its macro revised from a commercially released kit with an appropriate template. The Genotyper plots of the allelic ladder markers made in the present study and one of a DNA sample are shown in Fig. 1. We also calculated the allele frequencies at four STR loci in Table 1. Deviations from HWE ($p < 0.05$) were observed only in the two tests at the locus D7S809. In three other loci, however, no significant deviation ($p < 0.05$) from HWE was observed using the three tests. The departures from HWE are as would be expected ($p = 0.020$ and 0.033) for the D7S809 locus as they may be due to chance or sampling (16,17).

Only one off-ladder allele, allele 11, 14, and 27, was observed at each D2S3020, D15S233, and D14S299 locus, respectively. Alternatively, at D7S809, three off-ladder alleles (allele 12, 22, and 32) and one "microvariant" (shown as .1, .2 and .3) allele (allele 20.2) were additionally observed in the Thai population, and 17 alleles were totally observed in the Japanese and Thai populations. A total of 33 "nominal allele motif" alleles and one "microvariant" allele observed in both populations were sequenced. Consequently, the sequence analysis constructed a model of allelic sequence structure at D7S809 shown in Fig. 2, and four variable regions (x, y, z, and m) were supposed. "Microvariants," different sequence structures with the same allele size, were observed in 3 of 17 alleles from both populations. Sequence analysis of an off-ladder allele 12 observed in the Thai population also revealed a deletion of 24 nucleotides, the $(CAGG)_6$ sequence, downstream of the variable region "m". Alternatively, according to the sequence structure of a "microvariant" allele (20.2), two nucleotides (GA) were inserted next to the downstream of the variable region "x". Although it was supposed

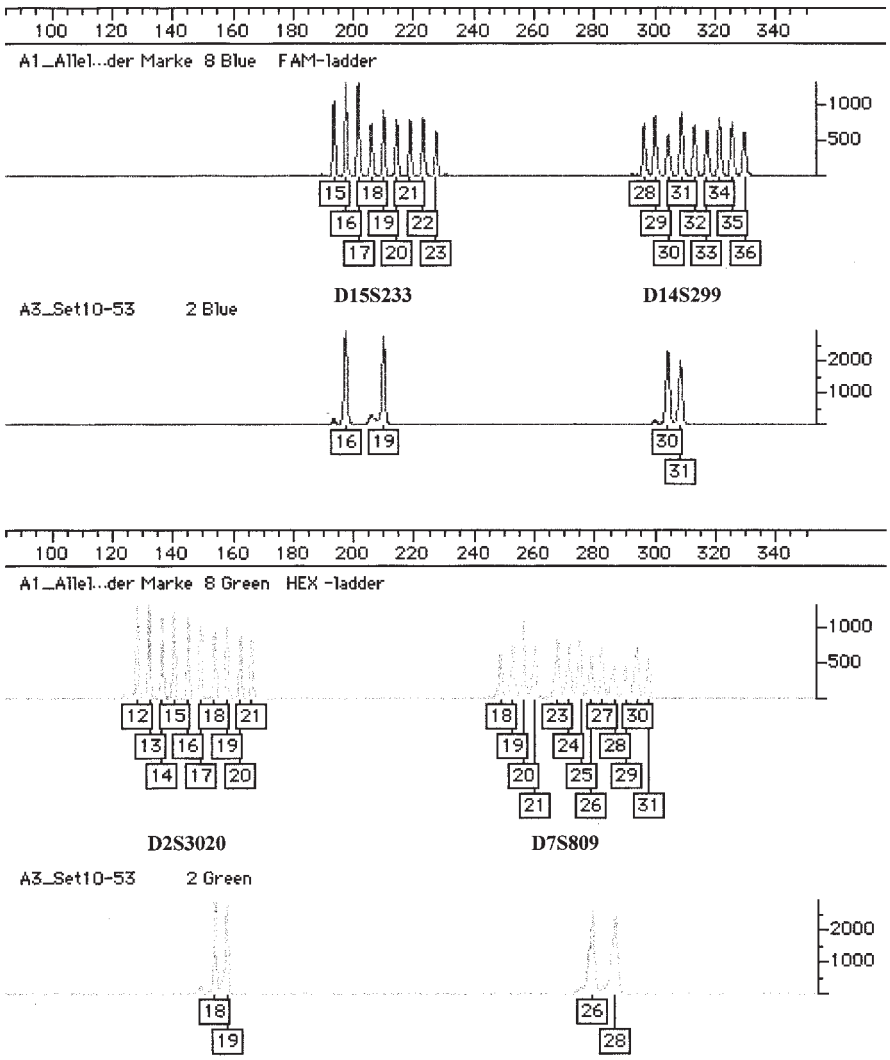

FIG. 1—*Genotyper plots of the allelic ladder marker and an example of sample for four STR loci in the present system.*

that this locus would imply more diverse and complicated sequence structures, all types were consistent with its number of repeats at least among the alleles analyzed in the present study.

The statistical properties, such as the observed heterozygosity (Obs.Hz), expected heterozygosity (Exp. Hz), power of discrimination (PD), polymorphism information content (PIC), and paternity exclusion rate (PE), were calculated as shown in Table 2. The Obs.Hz and PD at all four loci in the Thai population were higher than 0.8 and 0.9, respectively, and all of these statistical values in the Thai population were slightly higher than those in the Japanese population. The combined PD and combined PE of these four loci in the Thai population were 0.99998 and 0.978, respectively, almost the same as in the Japanese population (0.99996 and 0.971). Allele frequencies at all four loci in the Thai population were compared to the allele frequencies in the Japanese population described previously (8,11). The most frequent allele at D14S299 in the Thai population was allele 30 (32% compared to 45% in the Japanese population), while the next most frequent allele 31 was 23 and 30% in the Thai population and the Japanese population, respectively. The most frequent allele at D2S3020 in the Thai population was allele 15 (38%), which was 32% in the Japanese population, followed by allele 16 at 23 and 29% in the Thai population and the Japanese population, respectively. Probably due to these major differences, the allele distributions at D2S3020 and D14S299 in the Thai population were significantly different ($p < 0.05$) from those in the Japanese population when both populations were statistically compared using the Genepop software ($p = 0.0166$ and 0.0000, respectively). The allele frequency distributions at D15S233 and D7S809 were not significantly different between the Japanese population and the Thai population (*p*-values: 0.0667 and 0.0972, respectively). Interestingly, the allele distribution at D2S3020 in the Japanese population was not significantly different from Caucasians, as described previously (11).

The sensitivity of the present system was 0.1 to 0.2 ng of template DNA at 30 cycles (Fig. 3) and almost all samples containing 0.1 ng of DNA were genotyped accurately at 36 cycles. This sensitivity was almost the same as that of a commercially available kit. Furthermore, all PD at these four loci (0.908, 0.904, 0.956, and 0.882 at D2S3020, D15S233, D7S809, and D14S299, respectively) were higher than those at the "worst" four loci in CODIS core 13 loci in the Japanese population (0.788, 0.844, 0.861, and 0.861 at TPOX, TH01, CSF1PO, and D3S1358, respectively) (2). The highest PD, 0.956 at D7S809, was almost the same as the PD at the

TABLE 1—*The allele frequency distributions at the four STR loci of the Thai population compared with the Japanese population.*

	D2S3020			D15S233			D7S809			D14S299	
Allele	Thais (n = 366)	Japanese* (n = 548)	Allele	Thais (n = 366)	Japanese* (n = 712)	Allele	Thais (n = 366)	Japanese* (n = 256)	Allele	Thais (n = 366)	Japanese* (n = 656)
11	0.003		13	0.003		12	0.005		27	0.005	
12	0.008	0.002	14			18	0.014	0.004	28	0.003	0.005
13	0.057	0.055	15	0.033	0.028	19	0.060	0.055	29	0.036	0.084
14	0.148	0.184	16	0.178	0.173	20	0.022	0.027	30	0.317	0.453
15	0.377	0.323	17	0.295	0.317	20.2	0.003		31	0.298	0.229
16	0.232	0.292	18	0.148	0.118	21	0.027	0.012	32	0.197	0.130
17	0.090	0.091	19	0.293	0.291	22	0.005		33	0.049	0.038
18	0.063	0.029	20	0.096	0.053	23		0.004	34	0.041	0.043
19	0.022	0.015	21	0.005	0.014	24	0.016	0.012	35	0.044	0.005
20		0.005	22		0.003	25	0.068	0.145	36	0.011	0.011
21		0.004	23		0.003	26	0.227	0.191	37		
						27	0.191	0.164	38		0.003
						28	0.142	0.180	39		0.002
						29	0.128	0.152			
						30	0.049	0.039			
						31	0.036	0.016			
						32	0.005				
Tests for Hardy-Weinberg equilibrium (p-values)											
H. test†	0.192			0.114			0.332			0.093	
L. test‡	0.363			0.422			0.020			0.419	
E. test§	0.371			0.564			0.033			0.725	

n = Number of the allele examined.
* Allele frequencies in the Japanese population were published previously (Tamaki et al. 1996 and Yoshimoto et al. 2001).
† Homozygosity test.
‡ Likelihood ratio test.
§ Exact test.

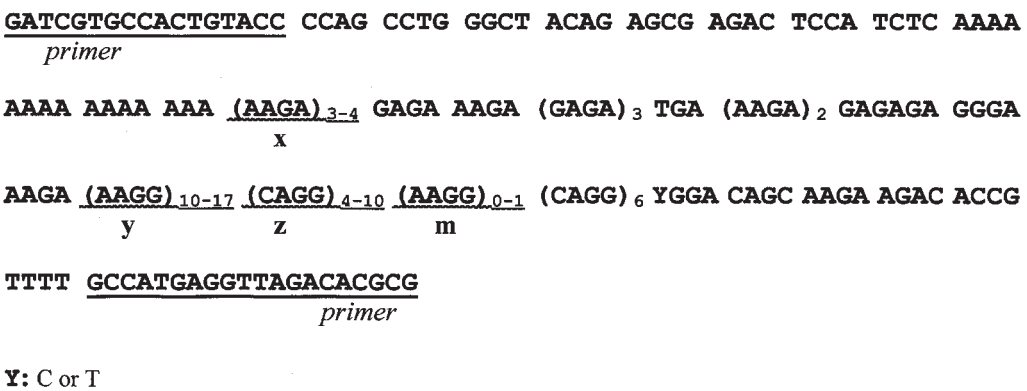

FIG. 2—*The allele sequence structure model at D7S809: Variable regions (x, y, z, and m) are shown as waved underlines.*

TABLE 2—*Statistical properties at the four loci of the Thai population compared with the Japanese population.*

	D2S3020		D15S233		D7S809		D14S299	
Locus	Thais	Japanese*	Thais	Japanese*	Thais	Japanese*	Thais	Japanese*
Obs. Hz†	0.803	0.770	0.836	0.778	0.885	0.859	0.814	0.729
Exp. Hz‡	0.768	0.765	0.792	0.768	0.864	0.858	0.767	0.717
PD§	0.914	0.908	0.925	0.904	0.966	0.956	0.909	0.882
PIC‖	0.735	0.728	0.759	0.731	0.847	0.837	0.729	0.680
PE¶	0.565	0.550	0.589	0.554	0.726	0.707	0.553	0.500

* Values in the Japanese population were published previously (Tamaki et al. 1996 and Yoshimoto et al. 2001).
† Observed heterozygosity.
‡ Expected heterozygosity.
§ Power of discrimination
‖ Polymorphism information content.
¶ Paternity exclusion rates (the combined PE of the four loci were 0.978 and 0.971 in the Thai population and in the Japanese population, respectively).

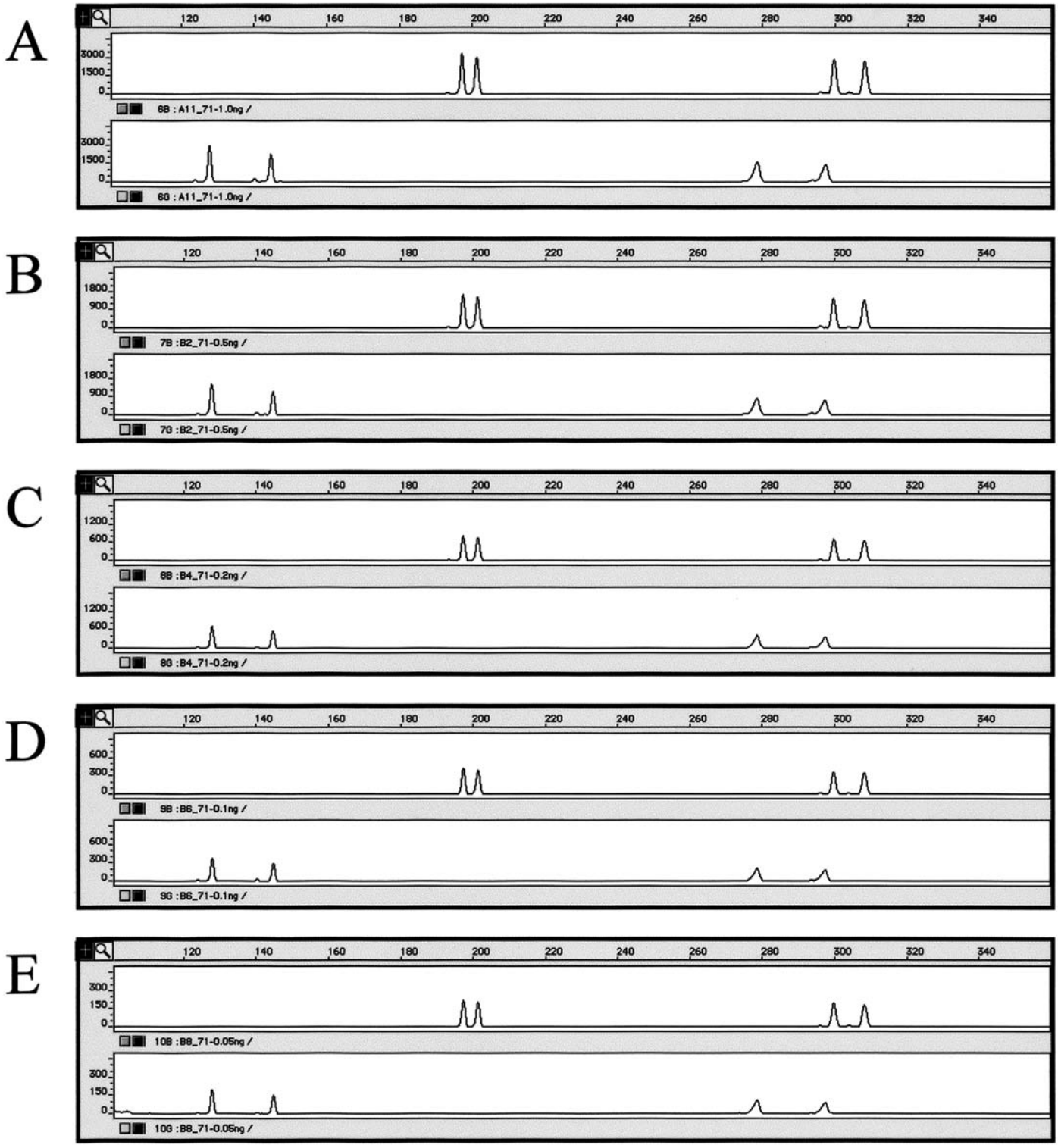

FIG. 3—*Effect of amplification on various amounts of high molecular template DNA ranging from 1.0 to 0.05 ng (A=1.0 ng, B=0.5 ng, C=0.2 ng, D=0.1 ng, E=0.05 ng). Each upper and lower panel shows the electropherogram of the blue and green lane, respectively.*

"best" three loci in CODIS core 13 loci in the Japanese population (0.964, 0.957, and 0.952 at FGA, D18S51, and D8S1179, respectively) (2, unpublished data).

Consequently, this system consisted of four informative STR loci with "nominal allele motif" tetranucleotide repetitive alleles except having a rare "microvariant" allele at D7S809. The allele detection system by Genetic Analyzer 310 is a convenient method to genotype all these four STR loci described using the two-colored fluorescent dyes and genotyping software. Therefore, this novel PCR amplification and typing system for four STR loci would be available as a DNA profiling system in the forensic field in Thais as well as Japanese.

Acknowledgments

This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1. Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B. The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. Electrophoresis 1999;20: 1682–96.
2. Yamamoto T, Uchihi R, Nozawa H, Huang X-L, Leong Y-K, et al. Allele distribution at nine STR Loci-D3S1358, vWA, FGA, TH01, TPOX, CSF1PO, D5S818, D13S317 and D7S820-in the Japanese population by multiplex PCR and capillary electrophoresis. J Forensic Sci 1999;44: 167–70.
3. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921.
4. Armour JAL, Neumamm R, Gobert S, Jeffreys AJ. Isolation of human simple repeat loci by hybridization selection. Hum Mol Genet 1994;3: 599–605.
5. Mizutani M, Huang X-L, Tamaki K, Yoshimoto T, Uchihi R, Yamomoto T, et al. Evaluation of two new STR loci 9q2h2 and wg3f12 in a Japanese population. Legal Med 1999;1:25–8.
6. Kojima T, Yamamoto T, Yoshimoto T, Tamaki K, Huang XL, Ohtaki H, et al. Tetrameric short tandem repeat (STR) system D15S233 (wg1d1): sequencing and frequency data in the Japanese and Chinese population. Legal Med 1999;1:119–26.
7. Yoshimoto T, Tamaki K, Katsumata S, Huang XL, Uchihi R, Tanaki M, et al. Sequence analysis of alleles at a microsatellite locus D14S299 (wg1c5) and population genetic comparisons. Int J Legal Med 1999;113: 15–18.
8. Tamaki K, Huang XL, Nozawa H, Yamamoto T, Uchihi R, Katsumata Y, et al. Evalution of tetranucleotide repeat locus D7S809 (wg1g9) in the Japanese population. Forensic Sci Int 1996;81:133–40.
9. Thomson JA, Pilotti V, Stevens P, Ayres KL, Debenham PG. Validation of short tandem repeat analysis for the investigation of cases of disputed paternity. Forensic Sci Int 1999;100:1–6.
10. Tamaki K, Yamamoto T, Uchihi R, Katsumata Y, Kondo K, Mizuno S, et al. Frequency of HLA-DQA1 alleles in the Japanese. Hum Hered 1991;41:209–14.
11. Yoshimoto T, Yamamoto T, Uchihi R, Tamaki K, Huang XL, Mizutani M, et al. A new triplex STR system without irregular alleles by silver staining and its potential application to forensic analysis. J Forensic Sci 2001;46:448–52.
12. Griffiths RAL, Barber MD, Johnson PE, Gillbard SM, Haywood MD, Smith CD, et al. New reference allelic ladders to improve allelic designation in a multiplex STR system. Int J Med 1998;111:267–72.
13. Weir BS. Independence of VNTR alleles defined by fixed bins. Genetics 1992;130:873–87.
14. Chakraborty R, Fornage M, Guegue R, Boerwinkle E. Population genetics of hypervariable loci: analysis of PCR based VNTR polymorphism within a population. In: Burke T, Dolf G, Jeffreys AJ, Wolff R, editors, DNA fingerprinting: approaches and applications. Berlin: Birkhauser, 1991;127–43.
15. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 1992; 48:361–72.
16. Budowle B, Smerick JB, Keys KM, Moretti TR. United States population data on the multiplex short tandem repeat loci—HUMTH01, TPOX, and CSF1PO—and variable number tandem locus D1S80. J Forensic Sci 1997;42:846–9.
17. Bell B, Budowle B, Martinez-Jarreta B, Casalod Y, Abecia E, Castellano M. Distribution of types of six PCR-based loci: LDLR, GYPA, HBGG, D7S8, GC, and HLA-DQA1 in Central Pyrenees and Teruel (Spain). J Forensic Sci 1997;42:510–3.

Additional information and reprint requests:
Toshimichi Yamamoto, Ph.D.
Department of Legal Medicine and Bioethics
Graduate School of Medicine
Nagoya University
65 Tsurumai-cho, Showa-ku
Nagoya 466-8550, Japan
E-mail: yamachan@med.nagoya-u.ac.jp
Tel.: +81-52-744-2119
Fax: +81-52-744-2121