TECHNICAL NOTE

Hélène Pfitzinger,[1] Ph.D.; Bertrand Ludes,[1] Ph.D.; Pascal Kintz,[1] Ph.D.; Antoine Tracqui, Ph.D.; and Patrice Mangin,[1] Ph.D.

French Caucasian Population Data for HUMTH01 and HUMFES/FPS Short Tandem Repeat (STR) Systems

REFERENCE: Pfitzinger, H., Ludes, B., Kintz, P., Tracqui, A., and Mangin, P., **"French Caucasian Population Data for HUMTH01 and HUMFES/FPS Short Tandem Repeat (STR) Systems,"** *Journal of Forensic Sciences,* JFSCA, Vol. 40, No. 2, March 1995, pp. 270–274.

ABSTRACT: The recent technology of amplification of DNA sequences by the polymerase chain reaction (PCR) has already proved to be a very useful tool for the analysis of variable number of tandem repeat (VNTR) loci. Short tandem repeat (STR) loci appear as other promising PCR-based identification systems. In fact, DNA typing based on PCR amplification of STRs is very sensitive and allows to overcome major problems encountered when using the RFLP method, such as typing of very small amounts of DNA, highly degraded DNA or mixtures of DNA from more than one individual.

Two STR systems, HUMTH01 (a tetranucleotide repeat (AATG) sequence located on chromosome 11) and HUMFES/FPS (a tetranucleotide repeat (ATTT) sequence located on chromosome 15) were investigated in order to determine allele and genotype frequencies for a French caucasian population sample.

HUMTH01 and HUMFES/FPS alleles were amplified by the use of PCR and amplified STR sequences were analyzed on 6% Hydrolink Long Ranger gels and visualized by silver staining.

The study was conducted on a sample of unrelated individuals ($N \approx 190$) randomly selected from the French caucasian population. The genotype distributions met Hardy-Weinberg expectations for both HUMTH01 and HUMFES/FPS STR systems. Furthermore, an additional allele, never reported before was observed at the HUMFES/FPS locus: it migrates as an allele containing 7 repeat units and corresponds to the smallest allele identified for this locus.

KEYWORDS: pathology and biology, PCR typing, population data, STR, HUMTH01, HUMFES/FPS

The recently developed technology of amplification of DNA sequences by the polymerase chain reaction [*1*] has already proved to be a very useful tool for the analysis of variable number of tandem repeat (VNTR or minisatellites) loci for forensic purposes [*2*]. Trimeric and tetrameric short tandem repeat (STR or microsatellites) DNA sequences are other highly polymorphic markers distributed throughout genic and extra-genic regions of the human genome [*3,4*]. STR loci appear as informative and promising PCR based identification systems. In fact, DNA typing based on amplification of STRs is very sensitive and allows to overcome major problems encountered in forensic practice such as typing of minute amounts of DNA, of highly degraded DNA or mixtures of DNA from more than one individual. The high sensitivity and the increased speed of the PCR technique are two important properties in forensic analyses.

Received for publication 18 March 1994; revised manuscript received 5 July 1994; accepted for publication 6 July 1994.

[1] Manager of the Forensic DNA Laboratory, Physician, Forensic Scientist in Charge of the Toxicology Laboratory, Physician, and Professor of Forensic Medicine/Director of the Institute, respectively, Institut de Médecine Légale, Faculté de Médecine, Strasbourg, France.

Materials and Methods

HUMTH01 Locus Description

HUMTH01 is a hypervariable microsatellite sequence of the human genome [*3–5*] located near the Human Tyrosine Hydroxylase Gene—Intron 1. It maps to chromosome 11 (11p15.5—p15) and consists of variable numbers of tetranucleotide repeat units: $(AATG)_n$.

The Population Sample—194 unrelated individuals were randomly selected from the French caucasian population. DNA was extracted from whole blood as previously described [*7*].

PCR Conditions—50 μL reaction mixture contained 10 ng DNA, 0.25 μM of each primer, 200 μM of each dNTP, 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% gelatine and 1.25 U Taq DNA polymerase (Gibco-BRL or Perkin Elmer).

Primer Sequences [*8*]—

1-5′-GTGGGCTGAAAAGCTCCCGATTAT-3′
2-5′-GTGATTCCCATTGGCCTGTTCCTC-3′

Locus Specific PCR Thermal Cycles—the DNA amplification was carried out in a Perkin-Elmer Gene Amp™ PCR System 9600 as follows: 94°C—45 s; 60°C—30 s; 72°C—30 s for 30 cycles.

HUMFES/FPS Locus Description

HUMFES/FPS is a hypervariable microsatellite sequence of the human genome [*6*] located near the Human C-FES/FPS proto oncogene and it maps to chromosome 15 (15q25-qter). The

HUMFES/FPS sequences consist of variable numbers of tetranucleotide repeat units: $(ATTT)_n$.

The population sample numbered 190 unrelated, randomly selected French caucasian individuals and *the PCR reaction conditions* were identical to that described for the HUMTH01 locus.

Primer Sequences [8]—

1-5′-GTGGGCTGAAAAGCTCCCGATTAT-3′
2-5′-GTGATTCCCATTGGCCTGTTCCTC-3′

The locus-specific PCR thermal cycles were as follows: 94°C—45 s; 54°C—30 s; 72°C—30 s for 30 cycles.

Identification of HUMTH01 and HUMFES/FPS Alleles

6 μL aliquots of the reaction mixtures were loaded on 6% Hydrolink Long Ranger gels (AT Biochem) and the alleles were visualized by silver staining.

Assignment of alleles—the allele designation used was based on the number of repeat units present in the amplified allelic fragments as already proposed (3–4 and 9; DNA recommendations 1992). The identification of alleles was achieved against an allelic ladder consisting of human alleles prepared from DNA of known genotypes. Both HUMTH01 and HUMFES/FPS allelic ladders were provided by the Forensic Science Service laboratory (Aldermaston, UK) as part of an EDNAP (European DNA Profiling Group) inter-laboratory exercise [10]. The HUMTH01 allelic ladder originally constructed by the FBI (unpublished data) and donated by Dr. B. Budowle (FBI, Quantico) consisted of 7 alleles (alleles 5, 6, 7, 8, 9, 10-1 and 11) among which the variant allele 10-1 (10) was present whereas allele 10 was absent.

The allelic ladders were obtained by reamplification of an aliquot of a 10^5 dilution of the PCR reaction mixture containing the already amplified allelic ladder. The HUMFES/FPS allelic ladder also consisted of 7 alleles: alleles 8, 9, 10, 11, 12, 13 and 14 [8].

Results and Discussion

The STR systems, HUMTH01 and HUMFES/FPS, were selected and investigated for their polymorphism. The French caucasian population data of both systems were compared to already published caucasian data.

HUMTH01 Population Studies

HUMTH01 Alleles—Six different HUMTH01 alleles (5, 6, 7, 8, 9 and 10-1) were identified in our population sample ($N = 194$) (Table 1 and Figs. 1 and 2). The PCR product sizes ranged from 154 to 178 bp. The smallest allele observed migrated as allele 5 containing 5 repeat units and the largest allele observed migrated as allele 10-1, the variant allele of the HUMTH01 system which differs from allele 10 by the absence of a single base-pair only [9]. Allele 11 did not appear in our population sample. The two most frequent alleles were allele 6 (25.0%) and allele 10-1 (30.67%) as already observed in other population samples (3, 4, 8, 9 and 11). However, allele 10 never identified in our population sample and rare (≈0.5%) in other population samples (4, 9 and 11) may have been mis-typed as allele 10-1 as a single base-pair resolution appears difficult in our gel system. Population data comparisons were therefore established by grouping alleles 10-1 and 10 for frequency purposes.

TABLE 1—*Population sample.*

Allele distribution: HUMTH01 locus		
Alleles	*n*	%
5	2	0.52
6	97	25.00
7	73	18.81
8	43	11.08
9	54	13.92
(10-1)	119	30.67
11	0	0
Total	$N = 388$	100%

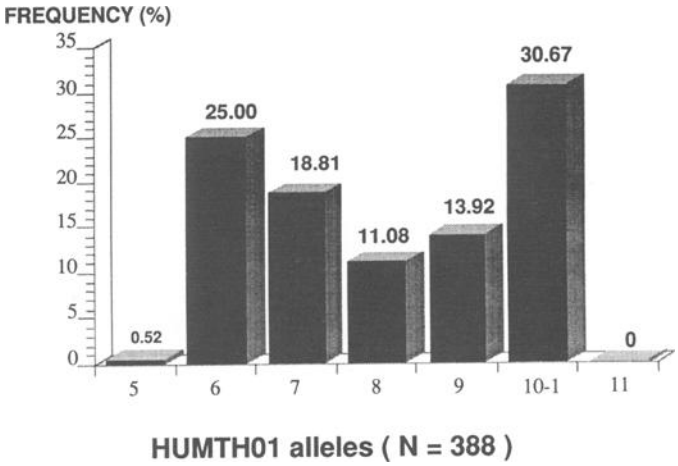

FIG. 1—*Distribution of HUMTH01 alleles in the French caucasian population.*

Our allele frequencies distribution was shown, using a R × C contingency table, to be similar to that observed by Wiegand et al. [11] for a German caucasian sample ($\chi^2 = 4.68$; df = 5; $0.3 < P < 0.5$) and also similar to that observed by Edwards et al. 1992 [4] for an American caucasian sample ($\chi^2 = 3.35$; df = 5; $0.5 < P < 0.9$). All classes with less than 5 observations were pooled.

HUMTH01 Genotypes

We observed 17 distinct genotypes out of 28 possible genotypes for the population sample studied. Among the seven possible homozygote genotypes only 5 were observed (Table 2): observed heterozygosity = 77.32%; allelic diversity: h = 0.77 and discrimination power: Pd = 0.91.

Because of relatively small population samples, a reliable estimation of deviations from the Hardy-Weinberg (HW) equilibrium was not possible using each separate allele. Alleles were therefore categorized into allele groups (12) and calculations have been carried out using a 5-allele model: no significant deviations were found between observed and expected values: $\chi^2 = 4.54$; df = 13; $P > 0.9$. This chi² test was repeated after regrouping of alleles in new subdivisions and again no significant deviations were observed.

The homogeneity of HUMTH01 genotype frequencies for our French caucasian population sample was assessed, using a R × C contingency table, with a German caucasian sample [11] and no significant differences appeared between the two samples: $\chi^2 = 11.38$; df = 12; $0.3 < P < 0.5$. A second homogeneity test

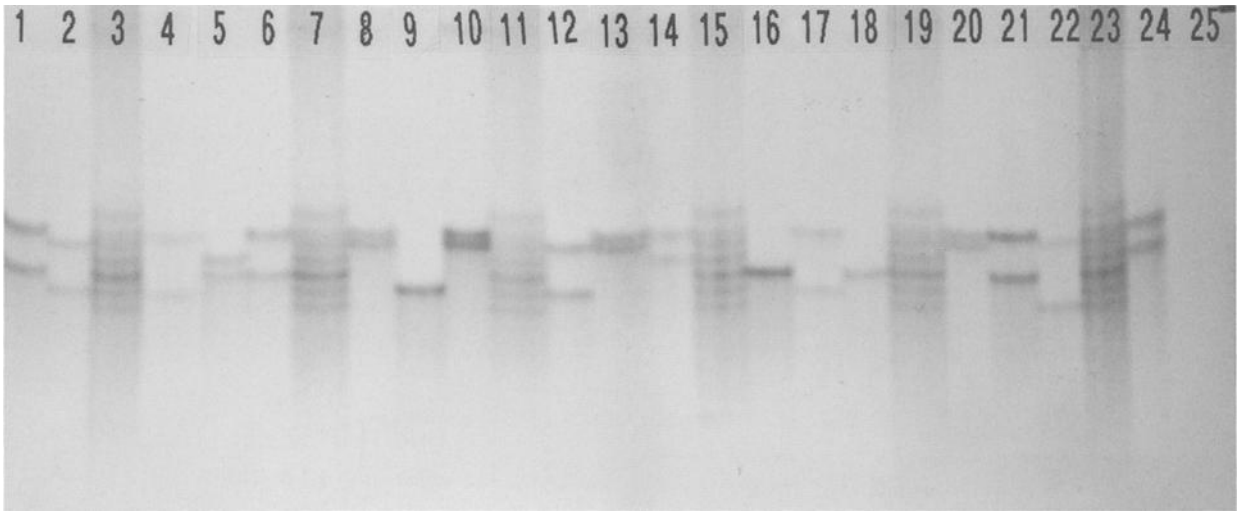

FIG. 2—*HUMTH01 data base samples visualized on a 6% Hydrolink Long Ranger gel using silver staining. The HUMTH01 genotypes are: lanes 1, 6 and 21: 7-(10-1); lanes 2, 4 and 12: 6-9; lane 5: 7-8; lanes 8, 10, 13 and 20: 9-(10-1); lane 9: 6-6; lanes 14 and 24 (control): 8-(10-1); lanes 16 and 18: 7-7; lane 17: 6-(10-1); lane 22: 5-9; lane 25: T_0; lanes 3, 7, 11, 15, 19 and 23: allelic ladder consisting of alleles 5, 6, 7, 8, 9, 10-1 and 11.*

TABLE 2—*Genotype distribution: HUMTH01 locus.*

Genotypes	*n* observed	*n* expected	% observed
5-5	/	/	/
5-6	1	0.5	0.52
5-7	/	/	/
5-8	/	/	/
5-9	1	0.27	0.52
5-(10-1)	/	/	/
5-11	/	/	/
6-6	13	12.13	6.7
6-7	14	18.26	7.22
6-8	7	10.75	3.61
6-9	14	13.5	7.22
6-(10-1)	35	29.76	18.04
6-11	/	/	/
7-7	8	6.87	4.12
7-8	8	8.09	4.12
7-9	12	10.17	6.18
7-(10-1)	23	22.39	11.86
7-11	/	/	/
8-8	5	2.39	2.58
8-9	6	5.98	3.09
8-(10-1)	12	13.19	6.18
8-11	/	/	/
9-9	2	3.76	1.03
9-(10-1)	17	16.57	8.76
9-11	/	/	/
(10-1)-(10-1)	16	18.26	8.25
(10-1)-11	/	/	/
11-11	/	/	/
Total	$N = 194$	/	100%

was conducted with an Italian caucasian sample (Genestandards, Sistema HUMTH01, Societa Italiana Chimici, Divisione Scientifica r. l. 00162 Roma, via L. Pulci, 27) and no significant differences were observed: $\chi^2 = 20.49$; df = 12; $P > 0.05$ (all classes with less than 3 observations were pooled).

HUMFES/FPS Population Studies

HUMFES/FPS Alleles—Seven different HUMFES/FPS alleles (7, 8, 9, 10, 11, 12 and 13) were identified in our population sample ($N = 190$) (Table 3 and Figs. 3 and 4). PCR product sizes

TABLE 3—*Allele distribution: HUMFES/FPS locus.*

Alleles	*n*	%
7	1	0.26
8	5	1.32
9	1	0.26
10	136	35.79
11	145	38.16
12	68	17.89
13	24	6.32
14	0	0
Total	380	100%

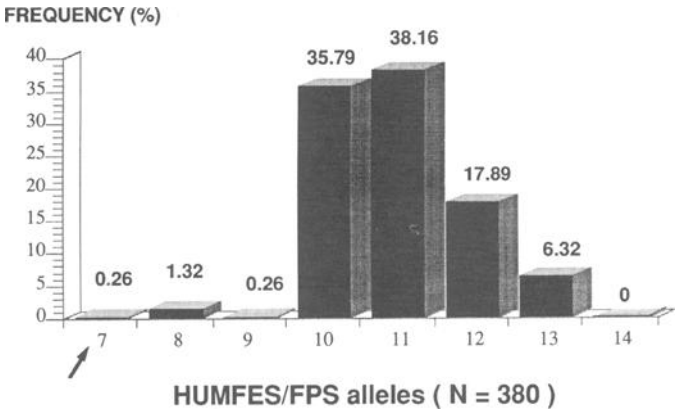

FIG. 3—*Distribution of HUMFES/FPS alleles in the French caucasian population.*

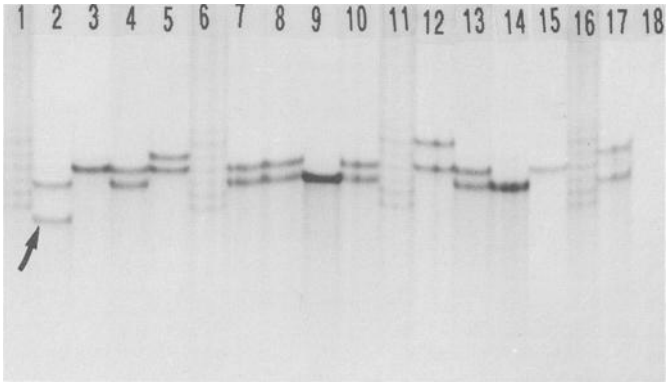

FIG. 4—*HUMFES/FPS data base samples visualized on a 6% Hydrolink Long Ranger gel using silver staining. The HUMFES/FPS genotypes are: lanes 4, 7, 8, 10 and 13: 10-11; lanes 3 and 15: 11-11; lanes 9 and 14: 10-10; lane 5: 11-12; lane 12: 11-13; lane 2: 7-10 (arrow = allele 7); lane 17: control = 10-12; lane 18: T_0; lanes 1, 6, 11 and 16: allelic ladder consisting of alleles 8, 9, 10, 11, 12, 13 and 14.*

ranged from 209 to 238 bp. The smallest allele observed migrated as allele 7, the largest allele observed migrated as allele 13 and allele 14 did not appear in our population sample. The two most frequent alleles were allele 10 (35.8%) and allele 11 (38.2%).

The allele frequencies distribution appeared similar to that observed by Kimpton et al. for an English caucasian population sample [*8*]: alleles 10, 11 and 12 were highly represented whereas allele 14 was not observed. Moreover, allele 7 identified in our population sample (Fig. 4, lane 2-arrow) is an additional allele not reported before and appears to migrate as a 7-repeat allele that has not been sequenced now.

HUMFES/FPS Genotypes—We observed 15 distinct genotypes out of 36 possible genotypes for the population sample studied (Table 4). Among the 8 possible homozygote genotypes only 4 were observed: observed heterozygosity = 72.7%; allelic diversity: h = 0.7 and discrimination power: Pd = 0.8.

The HW equilibrium was tested for HUMFES/FPS genotypes and calculations have been carried out using a 5-allele model. No significant deviations were found between observed and expected genotypes: $\chi^2 = 5.46$; df = 9; $0.5 < P < 0.9$.

Conclusion

Both HUMFES/FPS and HUMTH01 tetranucleotide STR loci were easily amplified using the PCR method and allowed us to gather population data for a sample of 190 and 194 French caucasian individuals respectively. The HUMTH01 and the HUMFES/FPS genotype distributions were in Hardy-Weinberg equilibrium. Genotype and allele frequencies were similar to that already observed for other caucasian population samples (3, 4, 8, 9, 10 and 11). In our HUMFES/FPS population sample we observed an additional allele not reported by Kimpton et al. [*8*] and migrating as a 7-repeat allele.

However, HUMTH01 and HUMFES/FPS PCR-based STR systems were studied for a French caucasian population sample for forensic purposes and as they require a greater number of systems to be run in order to reach the equivalent DNA discrimination obtainable with conventional RFLP VNTR systems [*8*], population data for other STR systems (HUMVWA31 and HUMF13A1) [*8*] are developed. Furthermore, multiplex reactions testing simultaneously several systems are investigated by the use of different fluorescent dyes in conjunction with the Applied Biosystems automated DNA sequencer (model 373A). The use of denaturing gels will also provide us an improved gel system allowing single base resolution and the detection of rare alleles differing by only one base-pair.

TABLE 4—*Genotype distribution: HUMFES/FPS locus.*

Genotypes	*n* observed	*n* expected	% observed
7-10	1	0.36	0.53
8-8	/	/	/
8-9	/	/	/
8-10	2	1.78	1.05
8-11	1	1.91	0.53
8-12	/	/	/
8-13	2	0.32	1.05
8-14	/	/	/
9-9	/	/	/
9-10	1	0.36	0.53
9-11	/	/	/
9-12	/	/	/
9-13	/	/	/
9-14	/	/	/
10-10	22	24.33	11.58
10-11	54	51.89	28.42
10-12	24	24.34	12.63
10-13	10	8.59	5.26
10-14	/	/	/
11-11	25	27.66	13.16
11-12	33	25.96	17.37
11-13	7	9.16	3.68
11-14	/	/	/
12-12	4	6.08	2.1
12-13	3	4.29	1.58
12-14	/	/	/
13-13	1	0.76	0.53
13-14	/	/	/
14-14	/	/	/
Total	$N = 190$	/	100%

References

[*1*] Saiki, R., Bugawan, T., Horn, G., Mullis, K., and Erlich, H., "Analysis of Enzymatically Amplified B-Globin and HLA-DQα DNA with Allele Specific Oliginucleotide Probe," *Nature,* Vol. 324, 1986, pp. 163–166.

[*2*] Budowle, B., Chakraborty, R., Giusti, A., Eisenberg, A., and Allen, R., "Analysis of the VNTR Locus D1S80 by the PCR Followed by High-Resolution PAGE," *American Journal of Human Genetics,* Vol. 48, 1991, pp. 137–144.

[*3*] Edwards, A., Civitello, A., Hammond, H. A., and Caskey, C. T., "DNA Typing and Genetic Mapping with Trimeric and Tetrameric Tandem Repeats," *American Journal of Human Genetics,* Vol. 49, 1991, pp. 746–756.

[*4*] Edwards, A., Hammond, H. A., Jin, L., Caskey, C. T., and Chakraborty, R., "Genetic Variation at Five Trimeric and Tetrameric Tandem Repeat Loci in Four Human Population Groups," *Genomics,* Vol. 12, 1992, pp. 241–253.

[*5*] Polymeropoulos, M. H., Xiao, H., Rath, D. S., and Merril, C. R., "Tetranucleotide Repeat Polymorphism at the Human Tyrosine Hydroxylase Gene (TH)," *Nucleic Acids Research,* 1991, Vol. 19, No. 13, p. 3753.

[*6*] Polymeropoulos, M. H., Rath, D. S., Xiao, H., and Merril, C. R., "Tetranucleotide Repeat Polymorphism at the Human c-fes/fps Protooncogene (FES)," *Nucleic Acids Research,* 1991, Vol. 19, No. 14, p. 4018.

[7] Ludes, B., Mangin, P., and Chaumont, A. J., "Stability of DNA in Brain Cortex After Long Post Mortem Periods, In G. Berghaus, B. Brinkmann, C. Rittner, and M. Staak, Eds., *DNA-Technology and its Forensic Application,* Springer, Berlin Heidelberg New York, 1991, pp. 187–191.

[8] Kimpton, C. P., Gill, P., Walton, A., Urquhart, A., Millican, E. S., and Adams, M., "Automated DNA Profiling Employing Multiplex Amplification of Short Tandem Repeat Loci," *PCR Methods and Applications,* 1993, Vol. 3, pp. 13–22.

[9] Puers, C., Hammond, H. A., Jin, L., Caskey, T. C., and Schumm, J., "Identification of Repeat Sequence Heterogeneity at the Polymorphic Short Tandem Repeat Locus HUMTH01 (AATG)n and Reassignment of Alleles in Population Analysis by Using a Locus-Specific Allelic Ladder," *American Journal of Human Genetics,* 1993, Vol. 53, pp. 953–958.

[10] Gill, P., Kimpton, C., D'Aloja, E., Andersen, J. F., Bar, W., Brinkmann, B., Holgersson, S., Johnsson, V., Kloosterman, A. D., Lareu, M. V., Nellemann, L., Pfitzinger, H., Phillips, C. P., Schmitter, H., Schneider, P. M., and Stenersen, M., "Report of the European DNA Profiling Group Towards Standardisation of Short Tandem Repeat (STR) Loci," *Forensic Science International,* 1994, Vol. 65, pp. 51–59.

[11] Wiegand, P., Budowle, B., Rand, S., and Brinkmann, B., "Forensic Validation of the STR Systems SE 33 and TC 11," *International Journal of Legal Medicine,* 1993, Vol. 105, pp. 315–320.

[12] Rand, S., Puers, C., Skowasch, K., Wiegand, P., Budowle, B., and Brinkmann, B., "Population Genetics and Forensic Efficiency Data of 4 AMPFLP's," *International Journal of Legal Medicine,* 1992, Vol. 104, pp. 329–333.

Address requests for reprints or additional information to
Hélène Pfitzinger, Ph.D.
Institut de Médecine Légale
Faculté de Médecine
11, rue Humann
67085 Strasbourg France