TECHNICAL NOTE

Timothy D. Kupferschmid,[1] *M.F.S.; Theresa Calicchio,*[1] *B.S.; and Bruce Budowle,*[2] *Ph.D.*

Maine Caucasian Population DNA Database Using Twelve Short Tandem Repeat Loci*

REFERENCE: Kupferschmid TD, Calicchio T, Budowle B. Maine Caucasian population DNA database using twelve short tandem repeat loci. J Forensic Sci 1999;44(2):392–395.

ABSTRACT: A population study of Caucasians residing in Maine was conducted using the AmpF1STR™ Profiler PCR Amplification Kit and the AmpF1STR™ Profiler Plus PCR Amplification Kit (Applied Biosystems Division (ABD) of Perkin Elmer, Foster City, CA). The kits contain the reagents necessary to amplify 12 different STR loci and the gender marker Amelogenin using two multiplex PCR, each containing nine STR loci. Thus, there is an overlap of six STR loci. The 12 STR loci are TH01, TPOX, CSF1PO, D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, and D7S820. These loci represent 12 of the 13 core loci selected by the CODIS STR standardization project. Dye-labeled amplification products were separated and detected using the capillary electrophoresis instrument ABI Prism™ 310 Genetic Analyzer. Allele frequencies were determined for the 12 STR loci. Statistical analysis of the data included Hardy-Weinburg equilibrium (HWE) analysis, pairwise independence testing, power of discrimination (PD), and probability of exclusion (PE).

KEYWORDS: forensic science, DNA typing, population genetics, population database, short tandem repeat loci, TH01, TPOX, CSF1PO, D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820

The typing of short tandem repeat (STR) loci is rapidly becoming routine for DNA analysis in forensic casework. STR loci consist of 2 to 7 base pair repeat units and are very common throughout the human genome (1,2). Further, coamplification of multiple STR loci is possible and has the advantage of increasing throughput (2,3). Two commercially available kits (AmpF1STR Profiler™ and AmpF1STR Profiler Plus™, ABD Perkin Elmer) allow for coamplification of 12 different STR loci and the gender marker Amelogenin. These STR loci include TH01, TPOX, CSF1PO, D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, and D7S820. These systems are highly discriminatory, very sensitive, have discrete alleles, and enable interpretation of some mixed samples.

Validation studies of STR multiplex systems are proving this technology to be reliable and reproducible (4,5). One step of the validation process recommended by the Technical Working Group on DNA Analysis Methods (TWGDAM) (6) is the establishment of population distribution data in different racial groups. Maine's population is 98.4% Caucasian (7). Therefore, a Caucasian population DNA database was established for Maine for use in estimating the statistical significance of a DNA match.

Materials and Methods

Sample Collection—One hundred fifty-one buccal swabs were collected from unrelated Caucasian volunteers. The volunteers consisted mostly of state and local law enforcement personnel and civilian employees of the State of Maine. The swabs were air dried and then stored at −70°C until DNA extraction.

DNA Typing—DNA was extracted from the samples using Chelex 100 resin as described by Walsh et al. (8). The quantity of the extracted DNA was estimated using a slot-blot assay (9). Approximately 1 ng of DNA was amplified. Amplification and typing was performed according to the AmpF1STR™ Profiler PCR Amplification Kit's recommended protocol. Amplification reactions were conducted in 25 μL reaction volumes containing 10.5 μL AmpF1STR PCR Reaction Mix, 5.5 μL AmpF1STR Primer Set, and 0.5 μL AmpliTaq Gold (2.5 Units). A GeneAmp PCR System 9600 (Perkin Elmer) thermal cycler was used per manufacturer's protocol. Samples were prepared for electrophoresis by combining 1.5 μL of PCR product with 25 μL of a master mix (1.2 mL deionized formamide and 25 μL GeneScan-350 [ROX] internal size standard). The PCR products were heat denatured at 95°C and snap cooled in a ice water bath. Amplified products were separated by electrophoresis on the ABI Prism™ 310 Genetic Analyzer. PCR products were injected for 5 s into a 36 (L_d) cm, 50 μm inside diameter capillary. Separations were run at 15 kV at 60°C containing Performance Optimized Polymer-4 (POP4™) (ABD Perkin Elmer). Data were collected using ABI Prism 310 Collection software, version 1.0.2, and analyzed using the GeneScan Analysis software, version 2.1 and Genotyper, version 2.0.

Statistical Analysis—The frequency of each allele for each locus was calculated from the numbers of each genotype in the sample

[1] Senior forensic DNA analyst and forensic DNA chemist, respectively, Maine State Police Crime Laboratory, Augusta, Maine.

[2] Program manager, Forensic Science Research and Training Center, FBI Laboratory, FBI Academy, Quantico, VA.

* Names of commercial manufacturers are provided for identification only and inclusion does not imply endorsement by the Federal Bureau of Identification or the Maine Department of Public Safety.

Received 29 May 1998; accepted 30 July 1998.

Copyright © 1999 by ASTM International

TABLE 1—*Observed allele frequency distributions of the STR loci TH01, TPOX, CSF1PO, D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, and D7S820.*

TH01			TPOX			CSF1PO		
Allele	No. Observ.	Frequency	Allele	No. Observ.	Frequency	Allele	No. Observ.	Frequency
6	79	0.265	8	148	0.497	7	1	0.003
7	49	0.164	9	33	0.111	8	1	0.003
8	32	0.107	10	12	0.040	9	6	0.021
9	44	0.148	11	94	0.315	10	92	0.317
9.3	91	0.305	12	10	0.034	11	79	0.272
10	3	0.010	13	1	0.003	12	88	0.303
N = 149*	...	...	N = 149*	...	...	13	18	0.062
...	...	...	...	...	...	14	5	0.017
...	...	...	...	...	...	...	N = 145*	...
D3S1358			vWA			FGA		
11	1	0.003	14	40	0.132	18	3	0.010
13	2	0.007	15	24	0.079	19	17	0.056
14	56	0.187	16	69	0.228	20	27	0.089
15	64	0.213	17	70	0.232	21	58	0.192
16	67	0.223	18	68	0.225	21.2	1	0.003
17	67	0.223	19	27	0.089	22	56	0.185
18	38	0.127	20	4	0.013	22.2	4	0.013
19	5	0.017	...	...	...	23	52	0.172
N = 150*	...	...	N = 151*	...	...	23.2	3	0.010
...	...	...	...	...	...	24	37	0.123
...	...	...	...	...	...	25	33	0.109
...	...	...	...	...	...	26	10	0.033
...	...	...	...	...	...	27	1	0.003
...	...	...	...	...	...	N = 151*	...	...
D8S1179			D21S11			D18S51		
8	10	0.033	27	10	0.033	10	3	0.010
9	4	0.013	28	49	0.163	11	4	0.013
10	33	0.109	29	75	0.250	12	43	0.144
11	25	0.083	29.2	1	0.033	13	35	0.117
12	33	0.109	30	76	0.253	14	50	0.168
13	90	0.298	30.2	13	0.043	15	44	0.148
14	64	0.212	31	18	0.060	16	34	0.114
15	33	0.109	31.2	23	0.077	17	30	0.100
16	7	0.023	32	2	0.007	18	25	0.084
17	3	0.010	32.2	23	0.077	19	16	0.054
...	...	...	33.2	7	0.023	20	7	0.024
...	...	...	34.2	3	0.010	21	4	0.013
...	...	...	...	...	...	22	1	0.003
...	...	...	...	...	...	23	1	0.003
...	...	...	...	...	...	26	1	0.003
N = 151*	...	...	N = 150*	...	...	N = 149*	...	...
D5S818			D13S317			D7S820		
9	9	0.030	8	31	0.103	7	2	0.007
10	8	0.026	9	23	0.076	8	37	0.123
11	134	0.444	10	20	0.066	9	52	0.172
12	100	0.331	11	102	0.338	10	76	0.252
13	47	0.156	12	81	0.268	11	67	0.222
14	3	0.010	13	33	0.109	12	54	0.179
15	1	0.003	14	11	0.036	13	11	0.036
N = 151*	...	...	15	1	0.003	14	3	0.010
...	...	...	N = 151*	...	...	N = 151*	...	...

* N refers to the number of individuals in the database whose sample yielded genotype information at the given locus.

set (i.e., the gene count method). Unbiased estimates of expected heterozygosity were computed as described by Edwards et al. (1). Possible divergence from HWE was tested by calculating the unbiased estimate of the expected homozygote/heterozygote frequencies (10–13), the likelihood ratio test (10,12,14), and the exact test (15), based on 2000 shuffling experiments. An interclass correlation criterion (16) for two-locus associations was used for detecting disequilibrium between the STR loci.

Results and Discussion

Genotype and allele frequency data were obtained from 151 Maine Caucasians at the STR loci TH01, TPOX, CSF1PO, D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, and D7S820 (although a few samples did not yield interpretable genotype data at all loci, see Table 1). There is little evidence of departure from HWE for these loci based on the

TABLE 2—*Tests for independence on the STR loci TH01, TPOX, CSF1PO, D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, and D7S820.*

Locus/Test	Result
TH01	
Obs. homozygosity	23.5%
Exp. homozygosity	22.1%
Homozygosity test*	0.691
Likelihood test*	0.966
Exact test*	0.946
TPOX	
Obs. homozygosity	39.6%
Exp. homozygosity	35.9%
Homozygosity test*	0.347
Likelihood test*	0.693
Exact test*	0.827
CSF1PO	
Obs. homozygosity	28.3%
Exp. homozygosity	26.9%
Homozygosity test*	0.709
Likelihood test*	0.247
Exact test*	0.252
D3S1358	
Obs. homozygosity	22.0%
Exp. homozygosity	19.4%
Homozygosity test*	0.417
Likelihood test*	0.738
Exact test*	0.833
vWA	
Obs. homozygosity	16.6%
Exp. homozygosity	18.6%
Homozygosity test*	0.519
Likelihood test*	0.786
Exact test*	0.840
FGA	
Obs. homozygosity	8.6%
Exp. homozygosity	13.8%
Homozygosity test*	0.066
Likelihood test*	0.868
Exact test*	0.952
D8S1179	
Obs. homozygosity	21.2%
Exp. homozygosity	17.6%
Homozygosity test*	0.240
Likelihood test*	0.505
Exact test*	0.509
D21S11	
Obs. homozygosity	19.3%
Exp. homozygosity	17.0%
Homozygosity test*	0.439
Likelihood test*	0.037
Exact test*	0.130
D18S51	
Obs. homozygosity	14.1%
Exp. homozygosity	11.6%
Homozygosity test*	0.336
Likelihood test*	0.445
Exact test*	0.287
D5S818	
Obs. homozygosity	29.8%
Exp. homozygosity	33.0%
Homozygosity test*	0.400
Likelihood test*	0.438
Exact test*	0.469
D13S317	
Obs. homozygosity	23.2%
Exp. homozygosity	21.7%
Homozygosity test*	0.669
Likelihood test*	0.219
Exact test*	0.220
D7S820	
Obs. homozygosity	21.2%
Exp. homozygosity	18.8%
Homozygosity test*	0.451
Likelihood test*	0.618
Exact test*	0.457

* These values are probability values.

TABLE 3—*Power of discrimination (PD) and probability of exclusion (PE).*

Locus	PD (Obs)	PD (Exp)	PE
TH01	0.91473357	0.91462289	0.56274017
TPOX	0.82257556	0.80999523	0.38213948
CSF1PO	0.87400713	0.87666854	0.48165096
D3S1358	0.93128889	0.93130272	0.60582646
vWA	0.93302925	0.93741281	0.62453996
FGA	0.95995790	0.96437297	0.71734034
D8S1179	0.94653743	0.94679635	0.65427542
D21S11	0.94897778	0.94938035	0.66385719
D18S51	0.96833476	0.97404815	0.75940184
D5S818	0.81645542	0.83033570	0.40762582
D13S317	0.9241568	0.92168531	0.58396014
D7S820	0.93145037	0.93586711	0.61960188
Total	>0.99999999	>0.99999999	>0.99999999

homozygosity test, likelihood ratio test, and the exact test (Table 2). There is only one example of departure from HWE and that is the D21S11 locus with the likelihood ratio test ($p = 0.037$). The power of discrimination (PD) and probability of exclusion (PE) data are shown in Table 3. In the pairwise independence tests, there were only two departures (FGA/D7S820 and D21S11/D5S818, $p = 0.034$ and $p = 0.049$, respectively) out of 66 comparisons. This meets statistical expectations.

Conclusion

The loci are highly polymorphic, the distribution of the allele frequencies meet HWE, and there is little evidence for association between the 12 STR loci in the Maine Caucasian population database. Thus, these data indicate that allele frequency estimations can be reliably calculated from forensic DNA casework.

References

1. Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 1992;12:241–53.
2. Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R. Evaluation of 13 short tandem repeat loci for use in personal identification applications. Am J Hum Genet 1994;55:175–89.
3. Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M. Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Meth Applic 1993;3:13–22.
4. Budowle B, Moretti TR, Keys KM, Koons BW, Smerick JB. Validation studies of the CTT STR multiplex system. J Forensic Sci 1997;42(2):701–7.
5. Wallin JM, Buoncristiani MR, Lazaruk KD, Fildes N, Holt CL, Walsh PS. TWGDAM validation of the AmpF1STR™ blue PCR amplification kit for forensic casework analysis. J Forensic Sci 1998;43(4):117.
6. Technical Working Group on DNA Analysis Methods. Guidelines for a quality assurance program for DNA analysis. Crime Lab Dig 1995;22(2):21–43.
7. 1990 general population characteristics: Maine. US Bureau of Census, Washington, DC: U.S. Gov. Printing Office, 1992.
8. Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Bio Techniques 1991;10:506–18.
9. Walsh PS, Valaro J, Reynolds R. A rapid chemiluminescent method for quantitation of human DNA. Nucleic Acids Res 1992;20: 5061–5.
10. Chakraborty R, Smouse PE, Neel JV. Population amalgamation and genetic variation: observations on artificially agglomerated tribal populations of Central and South America. Am J Hum Genet 1988; 43:709–25.

11. Chakraborty R, Fornage M, Guegue R, Boerwinkle E. Population genetics of hypervariable loci: analysis of PCR based VNTR polymorphism within a population. In: T Burke, G Dolf, AJ Jeffreys, and R Wolff, editors. DNA fingerprinting: approaches and applications, Berlin: Birkhauser Verlag, 1991;127–43.
12. Nei M, Roychoudhury AK. Sampling variances of heterozygosity and genetic distance. Genetics 1974;76:379–90.
13. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 1978;89:583–90.
14. Weir BS. Independence of VNTR alleles defined by fixed bins. Genetics 1992;130:873–87.
15. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 1992;48:361–72.
16. Karlin S, Cameron EC, Williams PT. Sibling and parent-offspring correlation estimation with variable family size. Proc Natl Acad Sci USA 1981;78:2664–8.

Additional information and reprint requests:
Timothy D. Kupferschmid
Maine State Police Crime Laboratory
30 Hospital Street
Augusta, Maine 04333-0133