TECHNICAL NOTE

Bruce Budowle,[1] *Ph.D.; Barbara W. Koons,*[1] *M.S.; and Tamyra R. Moretti,*[1] *Ph.D.*

Subtyping of the HLA-DQA1 Locus and Independence Testing with PM and STR/VNTR Loci*

REFERENCE: Budowle B, Koons BW, Moretti TR. Subtyping of the HLA-DQA1 locus and independence testing with PM and STR/VNTR loci. J Forensic Sci 1998;43(3)657–660.

ABSTRACT: Allele and genotype frequencies for six loci (HLA-DQA1 and PM loci) were determined in African Americans, United States Caucasians, and Southwestern Hispanics. The data include allele frequencies of the HLA-DQA1 4 subtypes. The HLA-DQA1 4 allele subtyping affords greater power of discrimination in African Americans and Southwestern Hispanics than in Caucasians, due to the relatively lower 4.2/4.3 allele frequency in Caucasians. Based on the exact test, all loci, except the GYPA locus in the African American sample ($p = 0.011$), meet Hardy-Weinberg expectations. There were two examples of significant departures from expectations of independence between alleles of the HLA-DQA1 and PM loci (HBGG/Gc in African Americans, $p = 0.30$; LDLR/DQA1 in Caucasians, $p = 0.023$). The HLA-DQA1 and PM loci also were tested for associations with three STR loci and the D1S80 locus. There were four examples of significant departures from expectations of independence (TPOX/D7S8 and THO1/HBGG in African Americans, $p = 0.035$ and 0.028, respectively; THO1/LDLR in Caucasians, $p = 0.028$; and GYPA/D1S80 in Hispanics, $p = 0.046$). The HLA-DQA1 and PM allele frequency data were compared with previously reported data on other sample populations of the same population categories from our laboratory; the allele frequencies at all loci, except the D7S8 locus in Hispanics ($p = 0.028$), were statistically similar. The frequency data can be used in forensic analyses and paternity tests to estimate the frequency of a multiple locus DNA profile in various general United States populations.

KEYWORDS: validation studies, polymerase chain reaction, population databases, LDLR, GYPA, HBGG, D7S8, Gc, HLA-DQA1, D1S80, CSF1PO, TPOX, THO1, STRs, multiplex amplification

The human leukocyte antigen (HLA)-DQA1 locus and the Polymarker (PM) loci are PCR-based systems used predominately for forensic identification in the United States (1–6). Analysis of these loci is facilitated by the use of commercially available kit(s). Recently, the kit for HLA-DQA1 typing has been modified to enable subtyping of the 4 allele into the 4.1 and the 4.2/4.3 alleles. This paper presents allele frequency data (including the 4 subtyping data) for the HLA-DQA1 locus in three United States population groups. The population sample groups analyzed in the current study had been typed previously for the short tandem repeat (STR) loci CSF1PO, TPOX, and THO1 and the variable number of tandem repeat (VNTR) locus D1S80 (7), but not for the PM loci. Because on many occasions the HLA-DQA1 locus will be typed in concert with the PM loci, the five PM loci also were typed in the current study. Tests for independence between the HLA-DQA1, PM loci, three STR loci, and the D1S80 locus are presented.

Materials and Methods

Sample Preparation

Whole blood, obtained in EDTA vacutainer tubes by venipuncture from African American, Caucasians, and Southwestern Hispanic individuals, was kindly provided by Dr. Arthur Eisenberg (University of North Texas Health Science Center, Fort Worth, TX). The DNA was extracted by the phenol-chloroform method (8). The quantity of extracted DNA was estimated using the slot-blot procedure described by Waye et al. (9) and Budowle et al. (10).

Typing

The HLA-DQA1 and PM loci were typed using the AmpliType® PM PCR Amplification and Typing Kit (Perkin Elmer Corporation, Norwalk, CT). The amplification conditions were those recommended by the manufacturer, except that 0.16 μg/μL of bovine serum albumin (Sigma, catalog #3350) were added to the PCR (11,12). Amplification was carried out in a Perkin-Elmer GeneAmp PCR System 9600 thermal cycler. The HLA-DQA1 typing strips were kindly provided by Dr. Rebecca Reynolds (Roche Molecular Systems, Alameda, CA).

Statistical Analysis

The frequency of each allele for each locus was calculated from the numbers of each genotype in the sample set (i.e., the gene count method). Unbiased estimates of expected heterozygosity were computed as described by Edwards et al. (13). Possible divergence from Hardy-Weinberg expectations (HWE) was tested by

[1]Forensic Science Research and Training Center, FBI Laboratory, FBI Academy, Quantico, VA.

*This is publication number 97-07 of the Laboratory Division of the Federal Bureau of Investigation. Names of commercial manufacturers are provided for identification only, and inclusion does not imply endorsement by the Federal Bureau of Investigation.

Received 21 April 1997; and in revised form 16 July 1997; accepted 27 Aug. 1997.

calculating the unbiased estimate of the expected homozygote/heterozygote frequencies (14–17), the likelihood ratio test (13,15,18), and the exact test (19), based on 2000 shuffling experiments. An interclass correlation criterion (20) for two-locus associations was used for detecting disequilibrium between the loci.

A 2 × N contingency table exact test was used to generate a G-statistic (2000 shuffling experiments) (21,22) to test for homogeneity between sample populations. The program for this analysis was kindly provided by R. Chakraborty (University of Texas School of Biomedical Sciences, Houston, TX).

Results and Discussion

The distribution of the observed HLA-DQA1 allele frequencies in African Americans, U.S. Caucasians, and Southwestern Hispanics are shown in Table 1. To the best of our knowledge this is the first published report on HLA-DQA1 subtyping of the 4 allele in U.S. populations. The HLA-DQA1 4.1 allele is more frequent than the 4.2/4.3 allele in all three population samples. Of the three U.S. populations, the Caucasian sample had the lowest 4.2/4.3 allele frequency ($f = 0.025$). Thus, the subtyping of the HLA-DQA1 4 allele affords greater power of discrimination in African Americans and Southwestern Hispanics than in Caucasians. The HLA-DQA1 locus meets HWE in all three sample groups.

The three U.S. sample population groups also were typed for the PM loci (Tables 2–4). Based on the exact test, all loci, except the GYPA locus in the African American sample ($p = 0.011$), meet Hardy-Weinberg expectations (HWE).

The data in the current study were compared with previously reported (1) HLA-DQA1 and PM population data on different samples of African Americans, U.S. Caucasians, and Southwestern Hispanics. First, the allele frequency distribution of the GYPA locus in African Americans is statistically similar to allele frequency data previously typed by our laboratory ($p = 0.935$) and in which no deviations were detected (1). Because the GYPA locus was the only example of a departure from HWE in the three populations and because the allele frequencies are similar to other African American data, the GYPA allele frequency data in the current African American sample can still be used reliably to estimate the frequency of a GYPA genotype. Second, within each major population group comparison, the allele frequencies at all loci are statistically similar, except the D7S8 locus in Southwestern Hispanics ($p = 0.028$). The D7S8 A and B allele frequencies for Southwestern Hispanics in this study are 0.587 and 0.413, respectively (Table 3), while in the previous study, Budowle et al. (1) reported the frequencies as 0.682 and 0.318, respectively. Thus, out of 18 comparisons (i.e., six loci comparisons in each of three population groups), only one comparison was significantly different. This observation is no more than would be expected by chance.

These results suggest that, within a major population group, the databases described here and by Budowle et al. (1) can be pooled. The D7S8 locus in the pooled southwestern Hispanic sample population was tested for departures from HWE, and no departures from expectations were observed ($N = 304$; Observed Homozygosity = 0.536; Expected Homozygosity (unbiased) = 0.527; HWE-Homozygosity Test ($p = 0.735$), Likelihood Ratio Test ($p = 0.805$), Exact Test ($p = 0.805$)).

TABLE 1—*HLA-DQA1 observed allele frequencies in several United States general population groups.*

Allele	African Americans* ($N = 206$)§	Caucasians† ($N = 199$)§	Southwestern Hispanics‡ ($N = 208$)§
1.1	0.112	0.153	0.103
1.2	0.308	0.191	0.125
1.3	0.051	0.060	0.031
2	0.078	0.151	0.103
3	0.148	0.196	0.240
4.1	0.189	0.224	0.269
4.2/4.3	0.114	0.025	0.127

African Americans—Observed Homozygosity = 0.121; Expected Homozygosity (unbiased) = 0.185; HWE-Homozygosity Test ($p = 0.019$), Likelihood Ratio Test ($p = 0.173$), Exact Test ($p = 0.232$).

†Caucasians—Observed Homozygosity = 0.141; Expected Homozygosity (unbiased = 0.173; HWE-Homozygosity Test ($p = 0.225$), Likelihood Ratio Test ($p = 0.371$), Exact Test ($p = 0.559$).

‡Southwestern Hispanics—Observed Homozygosity = 0.173; Expected Homozygosity (unbiased) = 0.183; HWE-Homozygosity Test ($p = 0.725$), Likelihood Ratio Test ($p = 0.114$), Exact Test ($p = 0.226$).

§N = Number of individuals in database.

TABLE 2—*Observed frequency distributions of PM loci genotypes.*

Genotype	African American ($N = 206$)*	Caucasian ($N = 199$)*	Southwestern Hispanic ($N = 208$)*
LDLR AA	0.029	0.171	0.255
LDLR AB	0.330	0.528	0.534
LDLR BB	0.641	0.302	0.212
GYPA AA	0.180	0.276	0.399
GYPA AB	0.592	0.513	0.472
GYPA BB	0.228	0.211	0.130
HBGG AA	0.189	0.271	0.115
HBGG AB	0.160	0.427	0.418
HBGG BB	0.078	0.302	0.428
HBGG AC	0.320	0.000	0.019
HBGG BC	0.160	0.000	0.019
HBGG CC	0.092	0.000	0.000
D7S8 AA	0.447	0.317	0.351
D7S8 AB	0.432	0.508	0.472
D7S8 BB	0.121	0.176	0.178
Gc AA	0.005	0.095	0.043
Gc AB	0.136	0.121	0.096
Gc BB	0.568	0.020	0.072
Gc AC	0.029	0.302	0.231
Gc BC	0.228	0.146	0.298
Gc CC	0.034	0.317	0.260

**N* refers to number of individuals in database.

TABLE 3—*Observed allele frequency distributions for PM loci.*

Allele	African American ($N = 206$)*	Caucasian ($N = 199$)*	Southwestern Hispanic ($N = 208$)*
LDLR A	0.194	0.435	0.467
LDLR B	0.806	0.565	0.534
GYPA A	0.476	0.533	0.635
GYPA B	0.524	0.467	0.365
HBGG A	0.430	0.485	0.334
HBGG B	0.238	0.515	0.647
HBGG C	0.333	0.000	0.019
D7S8 A	0.663	0.570	0.587
D7S9 B	0.337	0.430	0.413
Gc A	0.087	0.307	0.207
Gc B	0.750	0.153	0.269
Gc C	0.163	0.540	0.524

**N* refers to number of individuals in database.

An inter-class correlation test analysis detected only two examples of significant departures from expectations of independence between allele of the HLA-DQA1 and PM loci (HBGG/Gc in African Americans, $p = 0.030$; LDLR/DQA1 in Caucasians, $p = 0.023$). The African American, U.S. Caucasians, and Southwestern Hispanic samples had been typed previously for the STR loci CSF1PO, TPOX, and THO1 and the D1S80 locus (7). Thus, independence testing between the HLA-DQA1 and PM loci and the STR and D1S80 loci was performed. There were four examples of significant departures (TPOX/D7S8 and THO1/HBGG in African Americans, $p = 0.035$ and 0.028, respectively; THO1/LDLR in Caucasians, $p = 0.028$, and GYPA/D1S80 in Southwestern Hispanics, $p = 0.046$). This number of pair-wise departures is not substantially more than would be expected. Thus, the data support that the ten PCR-based loci meet expectations of independence for our African American, U.S. Caucasian, and Southwestern Hispanic populations.

In conclusion, African American, U.S. Caucasian, and Southwestern Hispanic databases are described for the loci HLA-DQA1, LDLR, GYPA, HBGG, D7S8, and Gc. Data are presented on the HLA-DQA1 4 allele subtypes. There is little evidence for departures from independence for these PCR-based loci. By applying the recommendations of the NRC II Report (23), these allele frequency data can be used to estimate the rarity of a multiple locus profile.

TABLE 4—*Tests for independence on PM loci.*

	African American ($N = 206$)*	Caucasian ($N = 199$)*	Southwestern Hispanic ($N = 208$)*
LDLR			
Observed Homozygosity	67.0%	47.2%	46.6%
Expected Homozygosity†	68.6%	50.7%	50.0%
Homozygosity Test‡	0.612	0.324	0.336
Likelihood Ratio Test‡	0.512	0.311	0.335
Exact Test‡	0.512	0.377	0.400
GYPA			
Observed Homozygosity	40.8%	48.7%	52.9%
Expected Homozygosity†	50.0%	50.1%	53.5%
Homozygosity Test‡	0.008	0.704	0.856
Likelihood Ratio Test‡	0.006	0.763	0.881
Exact Test‡	0.011	0.763	0.881
HBGG			
Observed Homozygosity	35.9%	57.3%	54.3%
Expected Homozygosity	35.0%		
Test†		49.9%	52.9%
Homozygosity‡	0.785	0.038	0.681
Likelihood Ratio Test‡	0.169	0.052	0.599
Exact Test‡	0.153	0.052	0.599
D7S8			
Observed Homozygosity	56.8%	49.3%	52.9%
Expected Homozygosity†	55.2%	50.9%	51.4%
Homozygosity Test‡	0.641	0.648	0.664
Likelihood Ratio Test‡	0.638	0.656	0.778
Exact Test‡	0.638	0.656	0.675
Gc			
Observed Homozygosity	60.7%	43.2%	37.5%
Expected Homozygosity†	59.6%	40.8%	38.8%
Homozygosity Test‡	0.743	0.484	0.692
Likelihood Ratio Test‡	0.828	0.402	0.804
Exact Test‡	0.824	0.381	0.815

*N = Number of individuals in database.
†Expected homozygosity is an unbiased estimate.
‡These values are probability values.

References

1. Budowle B, Lindsey JA, DeCou JA, Koons BW, Giusti AM, Comey CT. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8, and Gc (PM loci), and HLA-DQα using a multiplex amplification and typing procedure. J Forensic Sci 1995;40(1): 45–54.
2. Comey CT, Budowle B. Validation studies on the analysis of the HLA-DQ alpha locus using the polymerase chain reaction. J Forensic Sci 1991;36:1633–48.
3. Comey CT, Budowle B, Adams DE, Baumstark AL, Lindsey JA, Presley LA. PCR amplification and typing of the HLA DQα gene in forensic samples. J Forensic Sci 1993;38:239–49.
4. Gyllensten UB, Erlich HA. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQ alpha locus. Proc Natl Acad Sci USA 1988; 85:7652–6.
5. Walsh S, Fildes N, Louie AS, Higuchi R. Report of the blind trial of the Cetus AmpliType HLA DQα forensic deoxyribonucleic acid (DNA) amplification and typing kit. J Forensic Sci 1991;36: 1551–6.
6. Word CJ, Sawosik TM, Bing DH. Summary of validation studies from twenty-six forensic laboratories in the United States and Canada on the use of the AmpliType PM PCR amplification and typing kit. J Forensic Sci 1997;42:39–48.
7. Budowle B, Smerick JB, Keys KM, Moretti TR. United States population data on the multiplex short tandem repeat loci—HUMTHO1, TPOX, and CSF1PO and the variable number tandem repeat locus D1S80. J Forensic Sci 1997;42(5):846–9.
8. Comey CT, Koons BW, Presley KW, Smerick JB, Sobieralski CA, Stanley DM et al. DNA extraction strategies for amplified fragment length polymorphism analysis. J Forensic Sci 1994;39:1254–69.
9. Waye JS, Presley L, Budowle B, Shutler GG, Fourney RM. A simple method for quantifying human genomic DNA in forensic specimen extracts. Biotechniques 1989;7:852–5.
10. Budowle B, Baechtel FS, Comey CT, Giusti AM, Klevan L. Simple protocols for typing forensic biological evidence: chemiluminescent detection for human DNA quantitation and RFLP analyses and manual typing of PCR amplified polymorphisms. Electrophoresis 1995;16:1559–67.
11. Hochmeister MN, Budowle B, Borer UV, Eggmann UT, Comey CT, Dirnhofer R. Typing of DNA extracted from compact bone tissue from human remains. J Forensic Sci 1991;36:1649–61.
12. Paabo S, Gifford JA, Wilson AC. Mitochondrial DNA sequences from a 7000-year-old brain. Nuc Acids Res 1988;16:9775–8.
13. Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R. Genetic variation at five trimeric and tetrameric repeat loci in four human population groups. Genomics 1992;12:241–53.
14. Chakraborty R, Smouse PE, Neel JV. Population amalgamation and genetic variation: observations on artificially agglomerated tribal populations of Central and South America. Amer J Hum Genet 1988;43:709–25.
15. 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,

and Wolff R, editors, DNA Fingerprinting: Approaches and Applications, Birkhauser Verlag, Berlin, 1991;127–43.
16. Nei M, Roychoudhury AK. Sampling variances of heterozygosity and genetic distance. Genetics 1974;76:379–90.
17. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 1978;89:583–90.
18. Weir BS. Independence of VNTR alleles defined by fixed bins. Genetics 1992;130:873–87.
19. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 1992;48:361–72.
20. 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.
21. Roff DA, Bentzen P. The statistical analysis of mitochondrial DNA polymorphisms: χ^2 and the problem of small samples. Mol Biol Evol 1989;6:539–45.
22. Lewontin RC, Felsenstein J. The robustness of homogeneity tests in $2 \times N$ tables. Biometrics 1965;21:19–33.
23. National Research Council. The Evaluation of Forensic DNA Evidence, Washington, DC, National Academy Press, 1996.

Additional information and reprint request:
Bruce Budowle, Ph.D.
Forensic Science and Training Center
FBI Laboratory, FBI Academy
Quantico, VA 22135

ERRATUM

Erratum/Correction of Budowle B, Koons BW, Moretti TR. Subtyping of the HLA-DQA1 locus and independence testing with PM and STR/VNTR loci. J Forensic Sci 1998;43(3):657–660.

Sir:

An error in the values in one of the columns of Table 3 of the above referenced previously published paper requires correction. The allele frequencies at the LDLR locus for southwestern Hispanics are incorrect. My records show that the correct LDLR frequencies in this population are 0.521 for A and 0.478 for B. This transcriptional error was missed during page proof review. I hope that this error has not caused any inconvenience.

Bruce Budowle, Ph.D.
Forensic Science Research and Training Center
F.B.I. Academy
Quantico, VA 22135

Editor's Note: Any and all future citations of the above-referenced paper should read: Budowle B, Koons BW, Moretti TR. Subtyping of the HLA-DQA1 locus and independence testing with PM and STR/VNTR loci. [published erratum appears in J Forensic Sci 1998 Sep;43(5)] J Forensic Sci 1998;43(3):657–660.