TECHNICAL NOTE

Kwang Man Woo,[1] *M.S. and Bruce Budowle,*[2] *Ph.D.,*

Korean Population Data on the PCR-Based Loci LDLR, GYPA, HBGG, D7S8, Gc, HLA-DQA1, and D1S80

REFERENCE: Woo, K. M. and Budowle, B., "**Korean Population Data on the PCR-Based Loci LDLR, GYPA, HBGG, D7S8, Gc, HLA-DQA1, and D1S80,**" *Journal of Forensic Sciences,* JFSCA, Vol. 40, No. 4, July 1995, pp. 645–648.

ABSTRACT: Korean population data was generated for the loci LDLR, GYPA, HBGG, D7S8, Gc, HLA-DQA1, and D1S80. The genotype frequency distributions for the loci do not deviate from Hardy Weinberg expectations. Furthermore, there was no evidence for departures from expectations of independence between the loci. Using a test for homogeneity all the loci, except for D1S80, were similar between the Korean population sample and a Chinese population sample.

KEYWORDS: Korean, population databases, PCR, Hardy-Weinberg expectations, HLA-DQA1, LDLR, GYPA, HBGG, D7S8, Gc, and D1S80

There now are several genetic loci that can be amplified by the polymerase chain reaction (PCR) [*1*] and subsequently typed using commercially available kits. The most characterized of these PCR-based loci for forensic use are low density lipoprotein receptor (LDLR) [2], glycophorin A (GYPA) [*3*], hemoglobin G gammaglobin (HBGG) [*4*], D7S8 [*5*], group-specific component (Gc) [*6*] (PM loci), HLA-DQA1 [*7,8*], and D1S80 [*9,10*]. In order to use genetic loci in identity testing, some population data are needed. This paper presents allele and genotype frequency data in a Korean population sample. In addition, the Korean population data are compared with data from a Chinese population sample [*11,12*].

Materials and Methods

Sample Preparation

Whole blood was obtained in EDTA Vacutainer tubes by venipuncture from 116 unrelated Korean individuals collected by the Gene Analysis Section Laboratory of the General Affairs Department in Seoul, Korea.

Received for publication 31 Oct. 1994; revised manuscript received 6 Jan. 1995; accepted for publication 7 Jan. 1995.

[1]Assistant Chief of Gene Analysis Section, General Affairs Department, The Supreme Public Prosecutor's Office, Seoul, Korea.

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

This is publication number 95-03 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.

The DNA was extracted by the phenol-chloroform method according to the method of Budowle and Baechtel [*13*]. The quantity of extracted DNA was estimated using the slot-blot procedure described by Waye, et al. [*14*]. One-to-five ng of DNA were used for PCR.

Typing

The DNA samples were amplified and typed for the PM loci by using the AmpliType® PM PCR Amplification and Typing Kit (Perkin Elmer Corporation, Norwalk, CT) according to the manufacturer's protocol. The HLA-DQA1 locus was amplified and typed by using the Amplitype HLA-DQα Forensic DNA Amplification and Typing Kit (Perkin Elmer Corporation, Norwalk, CT). Amplification was carried out in a Perkin-Elmer DNA thermal cycler 480. The D1S80 locus was typed according to the method described by Budowle, et al. [*15*].

Statistical Analysis

The frequency of each allele for each locus was calculated from the numbers of each genotype in the sample set. Unbiased estimates of expected heterozygosity were computed as described by Edwards, et al. [*16*]. Possible divergence from Hardy-Weinberg expectations (HWE) was determined by calculating the unbiased estimate of the expected homozygote/heterozygote frequencies [*17–19*], the likelihood ratio test [*16,20,21*], and the exact test [*22*]. An inter-class correlation criterion [*23*] was used for detecting disequilibrium between loci pairs. Independence among the seven PCR-based loci was determined by examining whether or not the observed variance of the number of heterozygous loci in the population sample is within its confidence interval under the assumption of independence [*24,25*].

A 2 × C contingency table exact test was used to generate a G-statistic (1000 shuffling experiments) [*26,27*] to test for homogeneity between Korean and Chinese population samples. The program was kindly provided by R. Chakraborty (University of Texas School of Biomedical Sciences, Houston, Texas).

Results and Discussion

The distributions of observed allele and genotype frequencies for the PM, HLA-DQA1, and D1S80 loci in the Korean population

sample are shown in Tables 1, 2, 4, 5, and 6. The genotype frequency distributions for the loci do not deviate from HWE based on the homozygosity test, likelihood ratio test, and the exact test (Tables 3, 4, and 6).

An analysis was performed to determine whether or not there were any detectable associations between any of the seven PCR-based loci. An inter-class correlation test analysis demonstrated that there is no evidence for correlation between the alleles at any of the pairs of loci (Table 7). As an additional test for association, independence among the seven loci was evaluated by examining whether or not the observed variance (s_k^2) of the number of heterozygous loci in the population sample is within its confidence interval under the assumption of independence using the procedure described by Brown, et al. [*24*]. There was no evidence of association for the seven loci described in our Korean sample population using the s_k^2 criterion (s_k^2 = 1.492; 95% confidence interval of variance is 1.070–1.766).

Extant data demonstrate that general population groups (i.e., U.S. Caucasians, African Americans, etc.) are appropriate as reference groups for forensic identity testing [*28–30*]. Therefore, it was

TABLE 1—*Observed frequency distributions of PM loci genotypes in 116 unrelated Koreans.*

GENOTYPE	KOREAN
LDLR AA	0.043
LDLR AB	0.241
LDLR BB	0.716
GYPA AA	0.259
GYPA AB	0.552
GYPA BB	0.190
HBGG AA	0.112
HBGG AB	0.353
HBGG BB	0.534
HBGG AC	0.000
HBGG BC	0.000
HBGG CC	0.000
D7S8 AA	0.302
D7S8 AB	0.422
D7S8 BB	0.276
Gc AA	0.060
Gc AB	0.293
Gc BB	0.207
Gc AC	0.155
Gc BC	0.241
Gc CC	0.043

TABLE 2—*Observed allele frequency distributions for PM loci in 116 unrelated Koreans.*

ALLELE	KOREAN
LDLR A	0.164
LDLR B	0.836
GYPA A	0.534
GYPA B	0.466
HBGG A	0.289
HBGG B	0.711
HBGG C	0.000
D7S8 A	0.513
D7S8 B	0.487
Gc A	0.284
Gc B	0.474
Gc C	0.241

TABLE 3—*Tests for independence on PM loci.*

LDLR	
Obs. Homozygosity	75.9%
Exp. Homozygosity[a]	72.5%
Homozygosity Test[b]	0.416
Likelihood Ratio Test[b]	0.264
Exact Test[b]	0.153
GYPA	
Obs. Homozygosity	44.8%
Exp. Homozygosity[a]	50.0%
Homozygosity Test[b]	0.263
Likelihood Ratio Test[b]	0.264
Exact Test[b]	0.264
HBGG	
Obs. Homozygosity	64.7%
Exp. Homozygosity[a]	58.7%
Homozygosity Test[b]	0.196
Likelihood Ratio Test[b]	0.178
Exact Test[b]	0.178
D7S8	
Obs. Homozygosity	57.8%
Exp. Homozygosity[a]	49.8%
Homozygosity Test[b]	0.087
Likelihood Ratio Test[b]	0.152
Exact Test[b]	0.100
Gc	
Obs. Homozygosity	31.0%
Exp. Homozygosity[a]	36.1%
Homozygosity Test[b]	0.254
Likelihood Ratio Test[b]	0.583
Exact Test[a]	0.673

[a]Expected homozygosity is an unbiased estimate.
[b]These values are probability values.
[c]Test considered significant if $P < 0.05$.

TABLE 4—*Distribution of observed HLA-DQA1 genotype frequencies in 116 unrelated Koreans.*

GENOTYPE	KOREAN[a]
1.1–1.1	0.009
1.1–1.2	0.017
1.1–1.3	0.078
1.1–2	0.000
1.1–3	0.129
1.1–4	0.069
1.2–1.2	0.009
1.2–1.3	0.026
1.2–2	0.017
1.2–3	0.103
1.2–4	0.060
1.3–1.3	0.009
1.3–2	0.017
1.3–3	0.069
1.3–4	0.026
2–2	0.000
2–3	0.026
2–4	0.017
3–3	0.129
3–4	0.138
4–4	0.052

[a]Observed Homozygosity = 20.7%; Expected Homozygosity (unbiased) = 22.4%; HWE-Homozygosity Test ($P = 0.654$), Likelihood Ratio Test ($P = 0.549$), Exact Test ($P = 0.556$); Test considered significant if $P < 0.05$.

TABLE 5—*HLA-DQA1 observed allele frequencies in 116 unrelated Koreans.*

allele 1.1	0.155
allele 1.2	0.121
allele 1.3	0.116
allele 2	0.039
allele 3	0.362
allele 4	0.207

TABLE 6—*D1S80 allele frequencies in a sample of 116 unrelated Koreans.*

ALLELE	FREQUENCY
14	0.004
15	0.004
16	0.047
17	0.004
18	0.190
19	0.030
20	0.004
21	0.022
22	0.009
23	0.017
24	0.172
25	0.013
26	0.004
27	0.052
28	0.125
29	0.052
30	0.138
31	0.047
32	0.013
33	0.000
34	0.004
35	0.000
36	0.026
37	0.004
38	0.004
39	0.000
40	0.000
41	0.000
>41[a]	0.013

[a]All alleles migrating slower than the largest allele in the ladder (i.e., allele #41) are placed in the >41 allele class.
[b]Observed Homozygosity = 16.4%.
[c]Expected Homozygosity (unbiased) = 11.0%.
[d]HWE—Homozygosity Test ($P = 0.061$), Likelihood Ratio Test ($P = 0.114$), Exact Test ($P = 0.090$).
[e]Test considered significant if $P < 0.05$.

anticipated that subgroup population data on the PM, HLA-DQA1, and D1S80 loci would not be necessary for estimating DNA profile frequencies. However, it is desirable to generate some subgroup population data to demonstrate that subgroups within a major population category are more similar than major population groups. While significant differences for the seven PCR-based loci between the Korean and U.S. Caucasian or African Americans [*31,32*] were observed (data not shown), the degree of similarity between two Asian groups was not known. Huang, et al. [*11,12*] recently described population data for the PM, HLA-DQA1, and D1S80 loci in a Taiwanese population sample. Therefore, the Korean data were compared using a test for homogeneity with the only other Asian population sample, the Taiwanese, where the entire battery

TABLE 7—*Two locus inter-class correlation test for HLA-DQα, PM, and D1S80 loci for 116 unrelated Koreans.*

Loci	Two-sided Probability
LDLR/GYPA	0.706
LDLR/HBGG	0.665
LDLR/D7S8	0.153
LDLR/Gc	0.152
LDLR/DQα	0.085
GYPA/HBGG	0.493
GYPA/D7S8	1.000
GYPA/Gc	0.814
GYPA/DQα	0.577
HBGG/D7S8	0.500
HBGG/Gc	0.967
HBGG/DQα	0.679
D7S8/Gc	0.594
D7S8/DQα	0.500
Gc/DQα	0.731
D1S80/LDLR	0.057
D1S80/GYPA	0.612
D1S80/HBGG	0.180
D1S80/D7S8	0.461
D1S80/Gc	0.816
D1S80/DQα	0.837

[a]Test considered significant if $P < 0.05$.

of the seven PCR-based loci were analyzed (Table 8). There were no significant differences in allele frequencies for the PM and HLA-DQA1 loci for the two Asian groups. However, although the allele frequency distributions had similar trends, the D1S80 data were different. Because of the difference in D1S80 data, the Korean D1S80 data were compared with two other Asian D1S80 databases, a Japanese sample [*33*] and a general Asian sample [*31*]. The Korean D1S80 data were statistically similar to the Japanese data ($p = 0.1970 \pm 0.0126$), but not similar to the general Asian sample ($p = 0.0320 \pm 0.0056$). The difference in D1S80 data between the sample populations may be due to genetic differences or to small sample sizes. Regardless, the Korean and Chinese population data overall are similar. There would be no anticipated substantial differences in DNA profile frequency estimates if either sample population were used as a reference database.

In conclusion, a Korean population database has been established for seven PCR-based polymorphic loci. The distribution of the genotype frequencies for the various loci meet HWE, and there is no evidence for departures from expectations of independence

TABLE 8—*G statistic test (p values) for homogeneity on PM, HLA-DQA1, and D1S80 allele distributions in Koreans and Chinese.*[a]

Locus	G Statistic *P* value
LDLR	0.0810 ± 0.0086
GYPA	0.1570 ± 0.0115
HBGG	0.2730 ± 0.0141
D7S8	0.1050 ± 0.0097
Gc	0.6300 ± 0.0153
HLA-DQA1	0.1520 ± 0.0114
D1S80	0.0080 ± 0.0028

[a]Chinese population data from Huang, et al. [*11*].
[b]Test considered significant if $P < 0.05$.

of alleles across loci. The data demonstrate that estimates of multiple locus profile frequencies can be obtained from the Korean database for identity testing purposes using the product rule under the assumption of independence. Additionally, the Korean and Chinese databases appear more similar to each other than with two other major population group databases. However, comparisons of allele frequencies between different population samples are not very meaningful for forensic applications; instead a comparison of the differences in multiple locus DNA profile frequencies derived using various databases would be more informative [*28,29*]. A study is underway to evaluate the differences in PCR-based multiple locus DNA profile frequencies when various reference databases are employed.

References

[*1*] Saiki, R. K., Scharf, S., Faloona, T., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N., "Enzymatic Amplification of Beta-Globin Genomic Sequences and Restriction Analysis for Diagnosis of Sickle Cell Anemia," *Science*, Vol. 230, 1985, pp. 1350–1354.

[*2*] Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L., and Russell, D. W., "The Human LDL Receptor: A Cysteine-Rich Protein with Multiple Alu Sequences in Its mRNA," *Cell*, Vol. 39, 1984, pp. 27–38.

[*3*] Siebert, P. D. and Fukuda, M., "Molecular Cloning of Human Glycophorin B cDNA: Nucleotide Sequence and Genomic Relationship to Glycophorin A," *Proceedings of the National Academy of Sciences U.S.A.*, Vol. 84, 1987, pp. 6735–6739.

[*4*] Slightom, J. L., Blechl, A. E., and Smithies, O., "Human Fetal $^{G}\gamma$- and $^{A}\gamma$-globin Genes: Complete Nucleotide Sequences Suggest That DNA can be Exchanged Between These Duplicated Genes," *Cell*, Vol. 21, 1980, pp. 627–638.

[*5*] Horn, G. T., Richards, B., Merrill, J. J., and Klinger, K. W., "Characterization and Rapid Diagnostic Analysis of DNA Polymorphisms Closely Linked to the Cystic Fibrosis Locus," *Clinical Chemistry*, Vol. 36, 1990, pp. 1614–1619.

[*6*] Yang, F., Brune, J. L., Naylor, S. L., Apples, R. L., and Naberhaus, K. H., "Human Group-specific Component (Gc) is a Member of the Albumin Family," *Proceedings of the National Academy of Sciences U.S.A.*, Vol. 82, 1985, pp. 7994–7998.

[*7*] Gyllensten, U. B. and Erlich, H. A., "Generation of Single-Stranded DNA by the Polymerase Chain Reaction and its Application to Direct Sequencing of the HLA-DQ Alpha Locus," *Proceedings of the National Academy of Sciences U.S.A.*, Vol. 85, 1988, pp. 7652–7656.

[*8*] Saiki, R. K., Walsh, S., Levenson, C. H., and Erlich, H. A., "Genetic Analysis of Amplified DNA with Immobilized Sequence-Specific Oligonucleotide Probes," *Proceedings of the National Academy of Sciences U.S.A.*, Vol. 86, 1989, pp. 6230–6234.

[*9*] Kasai, K., Nakamura, Y., and White, R., "Amplification of a Variable Number of Tandem Repeat (VNTR) Locus (pMCT118) by the Polymerase Chain Reaction (PCR) and Its Application to Forensic Science," *Journal of Forensic Sciences*, Vol. 35, 1990, pp. 1196–1200.

[*10*] Budowle, B., Chakraborty, R., Giusti, A. M., Eisenberg, A. J., and Allen, R. C., "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.

[*11*] Huang, N. E. and Budowle, B., "Chinese Population Data on the PCR-based Loci HLA-DQ∞, LDLR, GYPA, HBGG, D7S8, and Gc," *Human Heredity* (in press).

[*12*] Huang, N. E., Chakraborty, R., and Budowle, B., "D1S80 Allele Frequencies in a Chinese Population," *International Journal of Legal Medicine* (in press).

[*13*] Budowle, B. and Baechtel, F. S., "Modifications to Improve the Effectiveness of Restriction Fragment Length Polymorphism Typing," *Applied and Theoretical Electrophoresis*, Vol. 1, 1990, pp. 181–187.

[*14*] Waye, J. S., Presley, L., Budowle, B., Shutler, G. G., and Fourney, R. M., "A Simple Method for Quantifying Human Genomic DNA in Forensic Specimen Extracts," *Biotechniques*, Vol. 7, 1989, pp. 852–855.

[*15*] Budowle, B., Baechtel, F. S., Smerick, J. B., Presley, K. W., Giusti, A. M., Parsons, G., Alevy, M., and Chakraborty, R., "D1S80 Population Data in African Americans, Caucasians, Southeastern Hispanics, Southwestern Hispanics, and Orientals," *Journal of Forensic Sciences* (in press).

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

[*17*] Chakraborty, R., Smouse, P. E., and Neel, J. V., "Population Amalgamation and Genetic Variation: Observations on Artificially Agglomerated Tribal Populations of Central and South America," *American Journal of Human Genetics*, Vol. 43, 1988, pp. 709–725.

[*18*] Nei, M. and Roychoudhury, A. K., "Sampling variances of heterozygosity and genetic distance," *Genetics*, Vol. 76, 1974, pp. 379–390.

[*19*] Nei, M., "Estimation of Average Heterozygosity and Genetic Distance From a Small Number of Individuals," *Genetics*, Vol. 89, 1978, pp. 583–590.

[*20*] Chakraborty, R., Fornage, M., Guegue, R., and Boerwinkle, E., "Population Genetics of Hypervariable Loci: Analysis of PCR Based VNTR Polymorphism within a Population," In T. Burke, G. Dolf, A. J. Jeffreys, R. Wolff, eds, DNA Fingerprinting: Approaches and Applications, Birkhauser, Verlag, Berlin, 1991, pp 127–143.

[*21*] Weir, B. S., "Independence of VNTR Alleles Defined by Fixed Bins," *Genetics*, Vol. 130, 1992, pp. 873–887.

[*22*] Guo, S. W. and Thompson, E. A., "Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles," *Biometrics*, Vol. 48, 1992, pp. 361–372.

[*23*] Karlin, S., Cameron, E. C., and Williams, P. T., "Sibling and Parent-Offspring Correlation Estimation with Variable Family Size," *Proceedings of the National Academy of Sciences U.S.A.*, Vol. 78, 1981, pp. 2664–2668.

[*24*] Brown, A. H. D., Feldman, M. W., and Nevo, E., "Multilocus Structure of Natural Populations of Hordeum Spontaneum," *Genetics*, Vol. 96, 1980, pp. 523–536.

[*25*] Chakraborty, R., "Detection of Nonrandom Association of Alleles From the Distribution of the Number of Heterozygous Loci in a Sample," *Genetics*, Vol. 108, 1984, pp. 719–731.

[*26*] Roff, D. A. and Bentzen, P., "The Statistical Analysis of Mitochondrial DNA Polymorphisms: χ^2 and the Problem of Small Samples," *Molecular Biology and Evolution*, Vol. 6, 1989, pp. 539–545.

[*27*] Sokal, R. R. and Rohlf, F. J., Biometry, 2nd Edition, W. H. Freeman and Company, San Francisco, 1981.

[*28*] Budowle, B., Monson, K. L., Giusti, A. M., and Brown, B., "The Assessment of Frequency Estimates of Hae III-Generated VNTR Profiles in Various Reference Databases," *Journal of Forensic Sciences*, Vol. 39, 1994, pp. 319–352.

[*29*] Budowle, B., Monson, K. L., Giusti, A. M., and Brown, B., "Evaluation of Hinf I-Generated VNTR Profile Frequencies Determined Using Various Ethnic Databases," *Journal of Forensic Sciences*, Vol. 39, 1994, pp. 988–1008.

[*30*] Devlin, B., Risch, N., and Roeder, K., "Statistical Evaluation of DNA Fingerprinting: a Critique of the NRC's Report," *Science*, Vol. 259, 1993, pp. 748–750.

[*31*] Budowle, B., Baechtel, F. S., Smerick, J. B., Presley, K. W., Giusti, A. M., Parsons, G., Alevy, M., and Chakraborty, R., "D1S80 Population Data in African Americans, Caucasians, Southeastern Hispanics, Southwestern Hispanics, and Orientals," *Journal of Forensic Sciences* (in press).

[*32*] Budowle, B., Lindsey, J. A., DeCou, J. A., Koons, B. W., Giusti, A. M., and Comey, C. T., "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," *Journal of Forensic Sciences* (in press).

[*33*] Nagai, A., Yamada, S., Bunai, Y., and Ohya, I., "Analysis of the VNTR Locus D1S80 in a Japanese Population," *International Journal of Legal Medicine*, Vol. 106, 1994, pp. 268–270.

Address requests for reprints or additional information to
Bruce Budowle, Ph.D.
FBI Academy
FSRTC
Quantico, VA 22135