Cecelia A. Crouse,[1] Ph.D.; William J. Feuer,[2] M.S.; Daniel C. Nippes,[3] M.S.; S. Celia Hutto,[4] M.D.; Karen S. Barnes,[5] B.S.; David Coffman,[5] B.S.; Suzanne H. Livingston,[5] B.S.; Lonnie Ginsberg,[6] B.S.; and Debra E. Glidewell,[1] B.S.

Analysis of HLA DQα Allele and Genotype Frequencies in Populations from Florida

REFERENCE: Crouse, C. A., Feuer, W. J., Nippes, D. C., Hutto, S. C., Barnes, K. S., Coffman, D., Livingston, S. H., Ginsberg, L., and Glidewell, D. E., **"Analysis of HLA DQα Allele and Genotype Frequencies in Populations from Florida,"** *Journal of Forensic Sciences*, JFSCA, Vol. 39, No. 3, May 1994, pp. 731–742.

ABSTRACT: HLA DQα allele and genotype frequencies for Caucasian, African American, Haitian, and Hispanic populations in Florida have been estimated. The Florida laboratories involved in these studies collected donor samples from a variety of sites including clinical laboratories, victim and suspect standards, blood banks, county jail detainees, and laboratory personnel. We have determined that the Caucasian and African American DQα genotype frequencies do not deviate significantly from Hardy-Weinberg expectations and as a result of this heterogeneity analyses, data from the four Florida Caucasian populations may be combined and data from the four Florida African American populations may be combined to form two large HLA DQα genotype frequency databanks. Further, data from the Florida Haitian population may be combined with the Florida African American population. Comparison of the combined Florida Caucasian populations, combined Florida African American populations, the Palm Beach Sheriff's Office (PBSO) Hispanic, and PBSO Haitian population with other databases does not support combination because allele frequency distributions are heterogenous.

KEYWORDS: pathology and biology, genotype frequencies, population databanks, HLA DQα, allele

Several recent reports have put forth recommendations for the development of DNA analysis protocols for the forensic community [*1–5*]. In addition to providing guidelines for the initiation of DNA programs in crime laboratories, these reports also emphasize

Received for publication 7 June 1993; revised manuscript received 12 Nov. 1993; accepted for publication 15 Nov. 1993.

[1]Forensic Scientists, Palm Beach County Sheriff's Crime Laboratory, West Palm Beach, FL.

[2]Research Associate, Department of Ophthalmology-Biostatistics, Univ. of Miami School of Medicine, Miami, FL.

[3]Laboratory Director, Regional Crime Laboratory at Indian River Community College, Fort Pierce, FL.

[4]Associate Professor, Department of Pediatric Infectious Disease, Univ. of Miami School of Medicine, Miami, FL.

[5]Senior Crime Laboratory Analysts, Florida Department of Law Enforcement Crime Laboratory, Tallahassee, FL.

[6]Senior Crime Laboratory Analyst, Florida Department of Law Enforcement Crime Laboratory, Pensacola, FL.

that interpretation of DNA typing results must be based on valid scientific methods for determining the probability that a random person matches a particular forensic sample. Specifically, it has been suggested that relevant population studies should be conducted for each DNA marker used in a forensic DNA laboratory in order to determine distribution data within racial/ethnic groups.

Variable number tandem repeat (VNTR) polymorphism has been studied extensively in many populations throughout the world [*6*] and allele frequencies compiled and made available for use in forensic case reports. Another DNA technology that is gaining popularity in many forensic laboratories is the PCR or the polymerase chain reaction [*7*]. The AmpliType HLA DQα Forensic DNA Amplification and Typing kit is commercially available and contains the necessary reagents to perform PCR based typing on the DQα region of the human leukocyte antigens (HLA) Class II genes. The AmpliType kit uses "typing strips" to identify six HLA DQα alleles [*7*]. The technology associated with the HLA DQα typing kit has been described in detail elsewhere [*8–14*]. Two reports regarding HLA DQα allele and genotype distribution data from a variety of populations have been published [*15,16*]. However, there are no reports on the HLA DQα allele and genotype frequency distributions in Florida populations.

The purpose of this study was to 1) determine if the Caucasian and African American HLA DQα databases from four Florida crime laboratories may be combined to form a single Caucasian and a single African American database for the state of Florida; 2) to evaluate and compare DQα allele and genotype distributions in a Haitian population from southern Florida with other racial/ethnic populations; 3) to compare combined Florida Caucasian and African American databases with previously published HLA DQα databases to determine if there are differences in HLA DQα allele and genotype distributions; and 4) to evaluate HLA DQα allele and genotype frequencies from South Florida Hispanics in order to ascertain if this population is similar to Hispanic populations from other geographical locations.

Materials and Methods

Nomenclature

The recently published "Nomenclature for factors of the HLA System, 1991" reports a list of all Class I and Class II genes and sequences of alleles officially recognized by the scientific community [*18*]. The DQα allele designations are as follows: DQα 1.1 is now DQA1 * 0101, DQα 1.2 is DQA1 * 0102; and DQα 1.3 is DQA1 * 0103. The DQα 2 locus is now DQA1 * 0201, DQα 3 is referred to as DQA1 * 0301, and DQα 4.1 is DQA * 0401. The HLA DQα locus designations on the AmpliType typing strip correspond to older nomenclature: DQα 1.1, 1.2, 1.3, 2, 3, and 4. This manuscript will maintain these DQα identifiers for the sake of continuity.

Source of Samples/Populations

Four Florida Crime Laboratories collected population specimens as follows:

Palm Beach Sheriff's Office (PBSO)—The PBSO laboratory collected whole blood purple top specimens from three regional hospital clinical laboratories for a total of 113 Caucasian, 104 African American and 100 Hispanic samples. Haitian samples (n = 88) were collected from a Palm Beach County hospital clinic and University of Miami Pediatric Infectious Disease Clinic. The PBSO Caucasian database also includes 16 samples from laboratory personnel. The contact at each hospital laboratory was responsible for verifying that the samples were neither duplicates nor from related individuals. The

donors determined racial status. Samples were received as whole blood and stains were made on cotton cloth and stored in individual plastic bags at −20°C.

Regional Crime Laboratory-Indian River (RCL-IR)—The RCL-IR performed DNA analysis on victim and suspect blood standards that were submitted to the laboratory. RCL-IR personnel verified that duplicates were not included in the population databanks. The donor determined his/her race. DNA analysis was performed on Caucasian (n = 136) and African American (n = 62) samples. There were no Hispanic or Haitian samples included in either of these populations. The standards were submitted as whole blood and stains made from the samples and stored in individual paper bags at −20°C.

Florida Department of Law Enforcement-Pensacola (FDLE-P)—FDLE-P coordinated efforts with a local county jail to receive whole blood samples voluntarily from individuals being incarcerated using EDTA Vacutainer blood collection tubes. Donors indicated racial status. Records regarding any personal information of the donor, except racial status, were not maintained in order to guarantee confidentiality. FDLE-P analyzed samples from Caucasian (n = 100) and African American (n = 97) donors. Hispanic and Haitian samples were not available. Blood stains were made from whole blood samples and stored at −20°C.

Florida Department of Law Enforcement-Tallahassee (FDLE-T)—FDLE-T collected samples from the regional blood bank from blood donors over a six week period. The restriction on donating twice within a six week period prevented duplication. FDLE-T also collected samples from laboratory personnel and case standards. Samples from Caucasians (n = 102) and African American (n = 104) donors were analyzed. No Hispanic or Haitian samples were available. The samples were collected as whole blood; stains were made and then stored at −20°C. Buccal swabs were also collected.

DNA Extraction, Amplification and Typing

DNA was extracted from dried blood stains using organic methods (PBSO) followed by ethanol precipitation or by using Chelex-100 (PBSO, RCL-IR, FDLE-P, and FDLE-T) extraction procedures. Both protocols have been described in published reports [*15,16*] and the 1990 AmpliType Users Guide [*17*].

PBSO, RCL-IR and FDLE-P crime laboratories amplified sample DNA using reagents from the AmpliType HLA DQα kit (Perkin-Elmer Corporation, Norwalk, CT) as per manufacturer's recommendations. PBSO and FDLE-P used the Perkin-Elmer 480 Thermal cycler and RCL-IR used the Perkin-Elmer TC-100 Thermal cycler. DNA samples from FDLE-T were amplified in a similar manner except PCR was done at 27 cycles. The FDLE-T cycling parameter was validated by amplifying and typing 50 blood stain samples of unknown DQα types for 27 cycles (FDLE-T) and these same 50 samples were amplified and typed by FDLE-P using 32 cycles. The DQα typing results were identical for all 50 samples typed in both laboratories; 28 of the 50 samples contained the DQα 1 allele and no preferential amplification was seen when the DNA was amplified for 27 cycles. For DNA extracted by Chelex-100, 20μl of extract was amplified. Organically extracted DNA was quantitated using QuantiBlot (Roche Molecular Systems, Alameda, CA) and approximately 10 ng was used in the amplification mixture. Typing was conducted as per manufacturer's suggestion in all four Florida laboratories with no deviation from the manufacturer's protocol.

Statistical Methods

Observed genotype frequencies were compared with Hardy-Weinberg expected values using the Chi-square test. Allele frequencies were compared between different populations using RxC contingency table chi-square analysis. StatXact, computer software that

generates exact P values was used when RxC contingency table analysis produced too many small expected cell frequencies to allow standard asymptotic tests (1992, Snedecor and Cochran, Cytel Software Corporation, Cambridge, MA). Tests of statistical significance were conducted at an alpha level of 0.05.

Results

Comparison of HLA DQα Frequencies in the Caucasian Population

Caucasian populations from four Florida crime laboratories, two from southeastern Florida (PBSO and RCL-IR), one from north Florida (FDLE-T) and one from northwestern Florida (FDLE-P), were typed for HLA DQα alleles using the AmpliType kit (Table 1A). Allele frequencies are similar and no statistically significant differences (P = 0.51) were found among these groups (Table 4A). The DQα observed frequencies and the Hardy-Weinberg expected genotype frequencies for the four Florida laboratories are shown in Table 2. None of the individual populations nor their sum departed statistically from Hardy-Weinberg expected values. The four Caucasian populations are very similar and do not violate assumptions of random assortment. Therefore, the four Caucasian databases have been combined and will be referred to as the Florida Caucasian database.

Table 1B shows the HLA DQα allele distribution for the Florida Caucasian database and previously published Caucasian population allele frequencies from Helmuth [*16*], Roche Biomedical Laboratories (RBL) [*16*], and the FBI [*15*]. Heterogeneity Chi Square analysis of these allele distributions show that differences among the Caucasian allele frequencies approach statistical significance (p = 0.051). In order to determine if there was a published Caucasian population allele distribution that was similar to the Florida Caucasian database, the Helmuth, RBL and FBI Caucasian databases were individually compared to the Florida Caucasian database allele distribution (Table 4A). The Florida Caucasian allele frequency distribution was very similar to those of RBL and FBI (p = 0.253). In contrast, when the Helmuth Caucasian database is compared to the FBI, Florida and RBL Caucasian databases, the differences approach statistical significance (Table 4A, p = 0.044).

Table 3 shows the observed and expected genotype frequencies for the combined Florida Caucasian databases and the Helmuth, RBL and FBI reported Caucasian databases. The FBI reported the observed number of genotype frequencies and the percent observed frequencies [*15*]. We have calculated the expected frequencies from the FBI data in order to compare all of the populations discussed herein. The results of this analysis shows that the genotype frequencies in all of the Caucasian populations studied are very similar (Table 4B, p = 0.415).

The power of discrimination (PD) is based on the genotype distribution for each population (Table 2 and 3). PD was determined from the genotype frequencies for each of the populations studied where PD = $1 - \_P_j^2$. (P_j is the frequency of each HLA DQα frequency), [*19*]. The PD values range from 0.91 (FDLE-P) to 0.94 (FDLE-T) for the Florida Caucasian populations (Table 2) and 0.93 for the combined Florida Caucasian population databases which is very consistent with previously published PD values (Table 3). Analysis of the allelic diversity ($h = [1 - \_X_i^2]\ [n/n - 1]$), which is an unbiased estimate of the genetic diversity within a population and is based on genotype frequencies, was also calculated for each Caucasian database (Table 2). The range of diversity was from 0.79 to 0.81 in the Florida Caucasian populations with an h value of 0.93 for the combined Florida Caucasian population.

Comparison of HLA DQα Allele and Genotype Frequencies in African American and Haitian Populations

African American populations from the four Florida Crime Laboratories were evaluated for HLA DQα allele distribution (Table 1A) and genotype frequencies (Table 2).

TABLE 1—*The distribution of HLA DQα alleles are described as % observed for both 1A (Florida populations) and 1B (combined Florida Caucasian or African American populations from 1A, and previously reported Caucasian and African American databases). The number of samples evaluated is indicated by* n.

A. HLA DQ alpha allele distribution(%) in Florida population databases

	Population									
	Caucasian				African American				Haitian	Hispanic
HLA DQ α allele	PBSO (n = 238)	RCL-IR (n = 272)	FDLE-P (n = 200)	FDLE-T (n = 204)	PBSO (n = 208)	RCL-IR (n = 124)	FDLE-P (n = 194)	FDLE-T (n = 208)	PBSO (n = 176)	PBSO (n = 200)
1.1	12.2	10.7	14.0	16.7	13.0	13.7	14.9	19.2	15.3	9.5
1.2	16.8	23.2	22.0	17.6	32.7	29.8	25.8	29.3	33.0	11.0
1.3	5.9	6.3	3.0	5.4	7.2	0.8	4.6	2.4	2.3	8.5
2	14.7	12.1	13.0	15.2	9.6	11.3	9.8	10.6	9.1	11.5
3	18.9	17.6	17.5	21.1	7.7	9.7	8.8	9.6	9.7	19.0
4	31.5	30.1	30.5	24.0	29.8	34.7	36.1	28.8	30.7	40.5

B. HLA DQ alpha allele distribution (%) in Florida and published population databases

	Population							
	Caucasian				African American			
HLA DQ α allele	Florida (n = 914)	Helmuth (n = 826)	RBL (n = 348)	FBI (n = 300)	Florida (n = 734)	Helmuth (n = 448)	RBL (n = 344)	FBI (n = 386)
1.1	13.1	13.7	12.6	12.3	15.4	15.0	11.3	13.7
1.2	20.0	19.7	24.4	18.0	29.4	26.3	29.4	25.6
1.3	5.3	8.5	4.3	8.7	4.1	4.5	3.8	4.4
2	13.7	10.9	13.5	14.7	10.2	12.1	11.1	11.7
3	18.7	20.1	16.7	15.0	8.9	11.8	12.2	10.1
4	29.2	27.1	28.5	31.3	32.0	30.4	32.3	34.5

TABLE 2—*The HLA DQα genotypes for the populations studied in four Florida crime laboratories are indicated as % observed (% expected). The % expected was calculated from Hardy-Weinberg equilibrium assumptions.* n = *number of individuals. The power of discrimination (PD) and allelic diversity (h) is described in the text.*

Observed and (expected) HLA DQ alpha genotype frequencies (%) in Florida populations

	Population									
	Caucasian				African American				Haitian	Hispanic
Genotype	PBSO (n = 119)	RCL-IR (n = 136)	FDLE-P (n = 100)	FDLE-T (n = 102)	PBSO (n = 104)	RCL-IR (n = 62)	FDLE-P (n = 97)	FDLE-T (n = 104)	PBSO (n = 88)	PBSO (n = 100)
1.1,1.1	0.8 (1.5)	0.7 (1.1)	0.0 (2.0)	2.9 (2.8)	1.0 (1.7)	1.6 (1.9)	2.1 (2.2)	6.7 (3.7)	0.0 (2.4)	2.0 (0.9)
1.1,1.2	2.5 (4.1)	4.4 (4.9)	9.0 (6.2)	5.9 (5.9)	9.6 (8.5)	12.9 (8.2)	6.2 (7.7)	8.7 (11.3)	6.8 (10.1)	1.0 (2.1)
1.1,1.3	0.8 (1.4)	1.5 (1.3)	1.0 (0.8)	1.0 (1.8)	1.9 (1.9)	0.0 (0.2)	1.0 (1.4)	0.0 (0.9)	1.1 (0.7)	1.0 (1.6)
1.1,2	3.4 (3.6)	2.2 (2.6)	5.0 (3.6)	4.9 (5.1)	2.9 (2.5)	3.2 (3.1)	5.2 (2.9)	2.9 (4.1)	4.5 (2.8)	1.0 (2.2)
1.1,3	5.0 (4.6)	3.7 (3.8)	5.0 (4.9)	5.9 (7.0)	2.9 (2.0)	1.6 (2.7)	2.1 (2.6)	1.9 (3.7)	5.7 (3.0)	7.0 (3.6)
1.1,4	10.9 (7.7)	8.1 (6.4)	8.0 (8.5)	9.8 (8.0)	6.7 (7.7)	6.5 (9.5)	11.3 (10.8)	11.5 (11.1)	12.5 (9.4)	5.0 (7.7)
1.2,1.2	5.0 (2.8)	5.1 (5.4)	3.0 (4.8)	4.9 (3.1)	11.5 (10.7)	11.3 (8.9)	4.1 (6.6)	11.5 (8.6)	12.5 (10.9)	2.0 (1.2)
1.2,1.3	0.8 (2.0)	3.7 (2.9)	0.0 (1.3)	1.0 (1.9)	2.9 (4.7)	0.0 (0.5)	1.0 (2.4)	1.0 (1.4)	1.1 (1.5)	1.0 (1.9)
1.2,2	9.2 (4.9)	4.4 (5.6)	7.0 (5.7)	7.8 (5.4)	4.8 (6.3)	4.8 (6.7)	3.1 (5.0)	5.8 (6.2)	6.8 (6.0)	6.0 (2.5)
1.2,3	4.2 (6.4)	12.5 (8.2)	13.0 (7.7)	3.9 (7.4)	2.9 (5.0)	3.2 (5.8)	5.2 (4.5)	3.8 (5.6)	6.8 (6.4)	4.0 (4.2)
1.2,4	6.7 (10.6)	11.0 (14.0)	9.0 (13.4)	6.9 (8.5)	22.1 (19.5)	16.1 (20.7)	27.8 (18.6)	16.3 (16.9)	19.3 (20.2)	6.0 (8.9)
1.3,1.3	0.0 (0.3)	0.7 (0.4)	0.0 (0.1)	0.0 (0.3)	1.0 (0.5)	0.0 (0.0)	0.0 (0.2)	0.0 (0.1)	0.0 (0.1)	0.0 (0.7)
1.3,2	2.5 (1.7)	0.7 (1.5)	2.0 (0.8)	1.0 (1.6)	1.9 (1.4)	0.0 (0.2)	1.0 (0.9)	1.0 (0.5)	2.3 (0.4)	2.0 (2.0)
1.3,3	2.5 (2.2)	1.5 (2.2)	0.0 (1.1)	3.9 (2.3)	1.9 (1.1)	0.0 (0.2)	3.1 (0.8)	1.0 (0.5)	0.0 (0.4)	1.0 (3.2)
1.3,4	5.0 (3.7)	3.7 (3.8)	3.0 (1.8)	3.9 (2.6)	3.8 (4.3)	1.6 (0.6)	3.1 (3.3)	1.9 (1.4)	0.0 (1.4)	12.0 (6.9)
2,2	1.7 (2.2)	2.2 (1.5)	0.0 (1.7)	2.0 (2.3)	1.9 (0.9)	0.0 (1.3)	0.0 (1.0)	2.9 (1.1)	0.0 (0.8)	0.0 (1.3)
2,3	4.2 (5.6)	5.9 (4.3)	0.0 (4.6)	6.9 (6.4)	0.0 (1.5)	4.8 (2.2)	2.1 (1.7)	1.9 (2.0)	3.4 (1.8)	1.0 (4.4)
2,4	6.7 (9.3)	6.6 (7.3)	12.0 (7.9)	5.9 (7.3)	5.8 (5.7)	9.7 (7.8)	8.2 (7.1)	3.8 (6.1)	1.1 (5.6)	13.0 (9.3)
3,3	4.2 (3.6)	2.2 (3.1)	3.0 (3.1)	4.9 (4.4)	0.0 (0.6)	1.6 (0.9)	1.0 (0.8)	1.0 (0.9)	0.0 (0.9)	3.0 (3.6)
3,4	13.4 (11.9)	7.4 (10.6)	11.0 (10.7)	11.8 (10.1)	7.7 (4.6)	6.5 (6.7)	3.1 (6.3)	8.7 (5.5)	3.4 (5.9)	19.0 (15.4)
4,4	10.1 (9.9)	11.8 (9.1)	9.0 (9.3)	4.9 (5.9)	6.7 (8.9)	14.5 (12.0)	9.3 (13.0)	7.7 (8.3)	12.5 (9.4)	13.0 (16.4)
P.D.	0.926	0.926	0.913	0.935	0.901	0.898	0.878	0.912	0.900	0.900
h	0.8	0.8	0.79	0.81	0.77	0.75	0.76	0.77	0.77	0.76

TABLE 3—*The % observed HLA DQα genotype frequencies are indicated as are the % expected genotype frequencies calculated from Hardy-Weinberg Equilibrium. The HLA DQα genotype frequencies from the four Florida Caucasian populations were combined to give a Florida Caucasian database and the four Florida African American populations were combined to give a Florida African American database. The Helmuth and RBL data was previously published. The FBI data was published as the number of observed individuals and % observed; for the sake of clarity, we have presented the FBI Caucasian and African American HLA DQα genotype data as % observed and % expected. The PD and* h *values were calculated as described in Table 2.* n = *number of samples.*

Observed and (expected) HLA DQ alpha genotype frequencies(%) from Florida and previously published population databanks

	Population							
	Caucasian				African American			
Genotype	Florida (n = 457)	Helmuth (n = 413)	RBL (n = 174)	FBI (n = 150)	Florida (n = 367)	Helmuth (n = 224)	RBL (n = 172)	FBI (n = 193)
1.1,1.1	1.1 (1.7)	2.2 (1.9)	1.2 (1.6)	2.0 (1.5)	3.0 (2.4)	3.6 (2.3)	0.0 (1.3)	2.1 (1.9)
1.1,1.2	5.3 (5.3)	3.6 (5.4)	5.2 (6.2)	6.0 (4.4)	9.0 (9.1)	7.6 (7.9)	12.2 (6.6)	7.8 (7.0)
1.1,1.3	1.1 (1.4)	2.9 (2.3)	1.7 (1.1)	2.0 (2.1)	0.8 (1.3)	0.9 (1.4)	0.0 (0.9)	1.0 (1.2)
1.1,2	3.7 (3.6)	1.9 (3.0)	4.0 (3.4)	5.3 (3.6)	3.5 (3.1)	3.6 (3.6)	2.9 (2.5)	3.1 (3.2)
1.1,3	4.8 (4.9)	5.3 (5.5)	2.9 (4.2)	1.3 (3.7)	2.2 (2.7)	2.7 (3.5)	1.2 (2.8)	2.6 (2.8)
1.1,4	9.2 (7.7)	9.2 (7.4)	9.2 (7.2)	6.0 (7.7)	9.3 (9.9)	8.0 (9.1)	6.4 (7.3)	8.8 (9.5)
1.2,1.2	4.6 (4.0)	4.6 (3.9)	8.6 (6.0)	3.3 (3.2)	9.5 (8.7)	8.5 (6.9)	5.8 (8.6)	4.7 (6.6)
1.2,1.3	1.5 (2.1)	3.4 (3.4)	1.2 (2.1)	2.0 (3.1)	1.4 (2.4)	2.2 (2.4)	1.2 (2.2)	1.6 (2.3)
1.2,2	7.0 (5.5)	4.6 (4.3)	4.6 (6.6)	6.0 (5.3)	4.6 (6.0)	4.0 (6.4)	6.4 (6.5)	5.7 (6.0)
1.2,3	8.5 (7.5)	8.2 (7.9)	7.5 (8.2)	4.0 (5.4)	3.8 (5.2)	7.1 (6.2)	8.1 (7.2)	6.2 (5.2)
1.2,4	8.5 (11.7)	10.4 (10.7)	13.2 (13.9)	11.3 (11.3)	21.0 (18.8)	14.7 (16.0)	19.2 (19.0)	20.7 (17.7)
1.3,1.3	0.2 (0.3)	1.2 (0.7)	0.0 (0.2)	0.7 (0.8)	0.3 (0.2)	0.0 (0.2)	0.6 (0.1)	0.5 (0.2)
1.3,2	1.5 (1.4)	1.5 (1.9)	0.6 (1.2)	2.0 (2.5)	1.1 (0.8)	2.2 (1.1)	1.2 (0.8)	2.6 (1.0)
1.3,3	2.0 (2.0)	1.7 (3.4)	0.6 (1.4)	2.7 (2.6)	1.6 (0.7)	1.3 (1.1)	0.6 (0.9)	1.0 (0.9)
1.3,4	3.9 (3.1)	5.1 (4.6)	4.6 (2.5)	7.3 (5.4)	2.7 (2.6)	2.2 (2.7)	3.5 (2.5)	1.6 (3.0)
2,2	1.5 (1.9)	2.2 (1.2)	2.3 (1.8)	1.3 (2.2)	1.4 (1.0)	2.2 (1.5)	0.6 (1.2)	1.0 (1.4)
2,3	4.4 (5.1)	4.8 (4.4)	4.6 (4.5)	4.7 (4.4)	1.9 (1.8)	1.3 (2.9)	5.2 (2.7)	1.0 (2.4)
2,4	7.7 (8.0)	4.6 (5.9)	8.6 (7.7)	8.7 (9.2)	6.5 (6.5)	8.5 (7.4)	5.2 (7.2)	8.8 (8.0)
3,3	3.5 (3.5)	4.4 (4.0)	3.5 (2.8)	4.7 (2.3)	0.8 (0.8)	0.9 (1.4)	0.0 (1.5)	0.5 (1.0)
3,4	10.7 (10.9)	11.4 (10.9)	10.9 (9.5)	8.0 (9.4)	6.5 (5.7)	9.4 (7.2)	9.3 (7.9)	8.3 (7.0)
4,4	9.2 (8.5)	6.8 (7.3)	5.2 (8.1)	10.7 (9.8)	9.0 (10.3)	8.9 (9.2)	10.4 (10.4)	10.4 (11.9)
P.D.	0.93	0.94	0.93	0.93	0.91	0.92	0.91	0.90
h	0.80	0.81	0.80	0.80	0.77	0.79	0.77	0.77

TABLE 4—*This table represents a summary of the* p*-values calculated for HLA DQα allele (4A) and genotype distributions (4B) among all of the populations studied. The statistical methods are described in Materials and Methods.*

A. Race	Allele Frequencies: Population Source	*p* value
Caucasian	PBSO, RCL-IR, FDLE-P, FDLE-T	0.51
African American	PBSO, RCL-IR, FDLE-P, FDLE-T	0.286
Haitian (PBSO)	vs. PBSO African American	0.35
	vs. All African American populations	0.43
Caucasian	Florida, RBL, Helmuth, FBI	0.051
	Florida, Helmuth, FBI	0.079
	RBL, Helmuth, FBI	0.044*
	Florida, RBL, FBI	0.252
	Florida, RBL, Helmuth	0.052
African American	Florida, RBL, Helmuth, FBI	0.755
Hispanic	PBSO, RBL, Helmuth, FBI	0.000*
	All three way groups	<0.005*
	RBL vs. Helmuth/Mex	<0.0001*
	Helmuth/Mex vs. PBSO	<0.0001*
	Helmuth/Mex vs. FBI	<0.0001*
	PBSO vs. RBL	0.0062*
	PBSO vs. FBI	0.230
	RBL vs. FBI	0.380

B. Race	Genotype frequencies: Population Source	*p* value
Caucasian	PBSO, RCL-IR, FDLE-P, FDLE-T	0.433
African American	PBSO, RCL-IR, FDLE-P, FDLE-T	0.69
Haitian (PBSO)	vs. PBSO African American	0.179
	vs. All 4 Florida African American	0.176
Caucasian	Florida, RBL, Helmuth, FBI	0.415
African American	Florida, RBL, Helmuth, FBI	0.745
Hispanic	RBL, Helmuth, FBI, PBSO	<0.0001*
	RBL, PBSO, FBI	0.15
	RBL, HELMUTH, FBI	<0.0001
	RBL, HELMUTH, PBSO	<0.0001
	HELMUTH, PBSO, FBI	<0.0001

All four databases were found to be very similar, showed no statistical evidence of differences (Table 4A, $p = 0.286$), and met Hardy-Weinberg assumptions both singly and combined. The PBSO, RCL-IR, FDLE-P and FDLE-T African American populations were therefore combined to form a Florida African American database which contains a total of 734 HLA DQα alleles. The allele (Table 1B) and genotype frequencies (Table 3) for the combined Florida African American database were compared to previously published African American data from Helmuth, RBL and FBI [*15,16*] and showed no evidence of statistical differences.

We have also evaluated a Haitian population from South Florida and determined that the HLA DQα allele (Table 1A) and genotype (Table 2) frequencies are very similar to the combined Florida African American and all previously published African American populations (Table 4B) and there is no evidence of violating Hardy-Weinberg assumptions. As a result, the Haitian population may be combined with all African American databases.

The PD values for the African American populations range from 0.88 (FDLE-P) to 0.91 (FDLE-T). The Haitian population has a PD value of 0.90 which is very similar to all populations studied. The combined African American population databases also had a PD value that was identical to two of the previously published databases (PD=0.77 for RBL and FBI databases) and similar to the Helmuth PD value of 0.79. The allelic diversity values ranged from of 0.75 to 0.77 in the Florida African American populations and 0.77 for the PBSO Haitian population.

Comparison of PBSO HLA DQα Hispanic Population Frequencies to Reported Databases

The PBSO Hispanic HLA DQα genotype frequencies were evaluated for violation of Hardy-Weinberg equilibrium and no significant effects were found ($p = 0.16$). The PBSO Hispanic DQα allele distribution (Table 1A) was compared to Helmuth/Mexican, RBL and FBI Hispanic populations. The DQα allele distributions among these Hispanic populations were significantly different (Table 4A, $p = <0.0001$). A comparison of the Hispanic allele distribution from each of Helmuth/Mexican, RBL and FBI Hispanic populations with the PBSO Hispanic allele frequency (Table 4A) showed that the Helmuth/Mexican allele frequencies were different from all other Hispanic databases. This is most likely due to a higher frequency of the HLA DQα 3 allele (43.5%) in the Helmuth/Mexican population compared to the PBSO, 19%; FBI, 22%; and RBL, 23.6% [*15,16*]. The Helmuth/Mexican samples were collected from individuals in Mexico and the HLA DQα allele and genotype frequencies were previously reported to be different than the RBL and FBI Hispanic populations [*15,16*]. The PBSO Hispanic allele distribution is also statistically different from RBL Hispanic allele frequencies (Table 4A, $p = 0.0062$) although in this case, the source of the difference appears to be the cumulative effect of small differences for each allele. The PBSO Hispanic allele frequencies are not significantly different from the FBI Hispanic population which consists of DNA samples from Baylor University School of Medicine (Table 4A, $p = 0.199$). A comparison of the Hispanic genotype frequencies (Table 4B) found that Helmuth/Mexicans were different from the PBSO, RBL, and FBI populations, but there were no significant differences among the latter three. As a result of these analysis, the PBSO Hispanic databases may be combined with the FBI Hispanic database but the PBSO Hispanic database cannot be combined with the RBL or Helmuth Hispanic databases.

Discussion

We have evaluated the applicability of combining population DQα typing results from four Florida crime laboratories in order to provide a single Caucasian database and a single African American database for the State of Florida. We found no evidence of differences in the DQα allele distribution or genotype frequencies from the four Florida Caucasian populations tested or evidence of subpopulations that violate Hardy-Weinberg assumptions and have, therefore, combined these Caucasian databases. We have also determined that the allele distribution and genotype frequencies from the four Florida African American databases are similar and as a result have combined the four African American populations into a single database. The combined Florida Caucasian population DQα genotype frequencies are significantly different from the four Florida African American genotype frequencies and, as a result, cannot be combined to provide a single HLA DQα database.

Upon comparing the Florida Caucasian database with other published results of Helmuth, Roche Biomedical Laboratories, and the FBI, there was some evidence of a statistical difference between the Helmuth Caucasian group and the Florida, RBL and FBI

groups [*15,16*]. However, it is not strong. The differences between allele frequencies are small and the large number of significant tests conducted in this study make it likely that a few spurious statistically significant comparisons may emerge. The Caucasian samples reported in Helmuth et al. [*16*] were collected from several sources including 324 DNA samples from Ed Blake (Forensic Science Associates, Richmond, CA), 80 samples from CEPH (Centre D'Etude du Polymorphisme Humaine), five from the University of California and four from Cetus employees. The sources of the samples from the combined Florida Caucasian populations were diverse as well, yet the four Florida databases allele frequencies are very similar (Table 4A, $p = 0.433$).

There is no statistical evidence of differences among the African American databases evaluated and as a result, these populations may be combined.

Due to the fact that there is a significant proportion of Hispanics in southern Florida, the PBSO Crime Laboratory has also analyzed Hispanic HLA DQα types and has found that the genotype frequencies are significantly different from the Florida Caucasian and Florida African American populations and as a result cannot be combined with either of these two databases.

Comparison of the PBSO Hispanic data with other published studies demonstrates that the Helmuth allele and genotype distributions are quite different from the other groups. The PBSO and RBL Hispanic groups differ statistically. The frequencies of the FBI alleles 1.2 and 4 are similar to the PBSO frequencies while the FBI alleles 1.3 and 2 are similar to the RBL frequencies. This indicates that further studies of regional differences among Hispanic populations is merited. Helmuth suggested that since the RBL Hispanic samples were identified by surname, the population may be heterogenous due to inclusion of DNA samples from Puerto Rico, Latin America, Mexico, Spain and other geographical locations. In contrast, in all of the studies presented, with the exception of the Helmuth/Mexican database, the genotype frequencies were very similar and in accord with Hardy-Weinberg expectations within each major Hispanic population database.

The Haitian population is also represented in the Florida community and our data shows that the Haitian allele and genotype frequencies are very similar to the HLA DQα frequencies found in the African American populations and there is no need to have a separate Haitian and African American database for HLA DQα frequency information.

A fundamental assumption in computing estimates and comparing them statistically is that the sampled individuals have been selected at random. However, the difficulty of identifying a sampling frame for the communities we studied and obtaining blood from individuals in a population based study made using a random sample impractical. Although biases may exist in our estimates, it is reassuring that allele frequencies collected from the four Florida laboratories were similar within the Caucasian and African American races and that all of the populations studied met assumptions of Hardy-Weinberg equilibrium regardless of the sample collection sources; clinical laboratories, victim/suspect standards, blood banks, county jail inmate volunteers or laboratory personnel. Further, technical differences between the four laboratories regarding DNA extraction protocols, Thermal cycler models, or amplification parameters did not have an effect on the similarity of the population DQα genotype frequencies from the Caucasian and African American populations. Differences identified in comparisons with HLA DQα alleles among the Florida and other published studies were not always evident in comparisons with HLA DQα genotypes. This is likely due to the higher statistical power of the allele based comparisons which use fewer categories (six as opposed to twenty-one).

In summary, the State of Florida has a single Caucasian and a single African American database available. Aside from an obvious larger sample size, the advantage of providing Florida with comprehensive HLA DQα databases is that it is not necessary to conduct Caucasian and African American population frequency studies in every Florida laboratory that elects to utilize HLA DQα typing thus saving time and money implementing

this PCR DNA analysis technique. Although the Caucasian and African American DQα genotype frequencies should show little, if any deviation, as the sample size increases, our current objective is to continue to accumulate HLA DQα typing information from participating Florida laboratories and update our statistical analysis of the populations on an annual basis. Further, these studies show it may not be possible to provide a national HLA DQα Caucasian or Hispanic database. However, the previously published African American databases as well as the Florida African American databases are all very similar and if combined would provide a database with a substantial number of HLA DQα genotypes.

Acknowledgments

The authors wish to thank Martin Tracy and Bruce Weir for their helpful discussion and review of the manuscript. We are grateful to Earl Ritzline, Magda Clanton, and Jack Remus for their technical assistance in the preparation of data.

References

[1] "Guidelines for a Quality Assurance Program for DNA RFLP Analysis," *Crime Laboratory Digest*, Vol. 16, No. 2, 1989, pp. 40–59.

[2] "Guidelines for a Proficiency Testing Program for DNA RFLP Analysis," *Crime Laboratory Digest*, Vol, 17, No. 3, 1990, pp. 59–64.

[3] "Guidelines for a Quality Assurance Program for DNA Analysis," *Crime Laboratory Digest*, Vol. 18, No. 2, 1991, pp. 44–75.

[4] Committee on DNA Technology in Forensic Science, The National Research Council, "DNA Technology in Forensic Science," National Academy Press, Washington, D.C. 1992.

[5] Fourney, R. M., "Forensic Reality and DNA Typing," *Proceedings of the Third International Symposium on Human Identification*, April 1992.

[6] Federal Bureau of Investigation, "VNTR Population Data: A Worldwide Study," Vol. 1A, 1B, 2, 3, and 4, 1993.

[7] *AmpliType-Users Guide*, Version 2, Cetus Corporation, 1990.

[8] Saiki, R. K., Scharf, S. J., Faloona, F., Mullis, K. B., Horn, G. T., et al., "Enzymatic Amplification of β-globin Genomic Sequence and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia," *Science*, Vol. 230, 1985, pp. 1350–1354.

[9] Gorski, J., "Analysis of HLA-DQα Polymorphism Using Sequence-Specific Oligonucleotide Probe Hybridization and Gene Amplification," *The HLA System, A New Approach*, J. Lee, Ed., Springer-Verlag, New York, 1990, pp. 73–106.

[10] Horn, G. T., Bugawan, T. L., Long, C., and Erlich, H. A., "Allelic Sequence Variation of the HLA DQα Loci: Relationship to Serology and Insulin-Dependent Diabetes Susceptibility," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 85, 1988, pp. 6012–6016.

[11] Scharf, S., Horn, G. T., and Erlich, H. A., "Direct Cloning and Sequencing Analysis of Enzymatically Amplified Genomic Sequences," *Science*, Vol. 233, 1986, pp. 1076–1078.

[12] 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 DQα Locus," *Proceedings of the National Academy of Science USA*, Vol. 85, 1988, pp. 7652–7556.

[13] Saiki, R., Walsh, P. 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 Science USA*, Vol. 86, 1989, pp. 6230–6234.

[14] Saiki, R., Bugawan, T. L., Horn, G. T., Mullis, K. B., and Erlich, H. A., "Analysis of Enzymatically Amplified β-Globin and HLA DQα DNA with Allele-Specific Oligonucleotide Probes," *Science*, Vol. 324, 1986, pp. 163–166.

[15] Comey, C. and Budowle, B., "Validation Studies on the Analysis of the HLA DQα Locus Using the Polymerase Chain Reaction," *Journal of Forensic Sciences*, Vol. 36, No. 6, November 1991, pp. 1633–1648.

[16] Helmuth, R., Fildes, N., Blake, E., Luce, M. C., Chimera, J., Madej, R., Gorodezky, C., Stoneking, M., Schmill, N., Klitz, W., Higuchi, R., and Erlich, H. A., "HLA DQα Allele and Genotype Frequencies in Various Human Populations, Determined by Using Enzymatic Am-

plification and Oligonucleotide Probes," *American Journal of Human Genetics*, Vol. 47, 1990, pp. 515–523.
[17] Walsh, P., Metzger, D. A., and Higuchi, R., "Chelex 100 As a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material," *Biotechniques*, Vol. 10, No. 4, 1991, pp. 506–513.
[18] Bodmer, J. and committee members, "Nomenclature for Factors of the HLA System, 1991," *Immunogenetics*, Vol. 36, 1992, pp. 135–148.
[19] Fisher, R. A., *Heredity*, 1951, Chapter 5, pp. 95–102.

Address requests for reprints or additional information to
Cecelia A. Crouse, Ph.D.
Palm Beach County Sheriff's Crime Laboratory
3228 Gun Club Road
West Palm Beach, FL 33406