Int J Leg Med (1993) 105:233–238

International Journal of
Legal Medicine

Population data of the HLA DQα locus in Dutch caucasians

Comparison with other population studies

Ate D. Kloosterman[1], Bruce Budowle[2], and Erin L. Riley[3]

[1]Gerechtelijk Laboratorium van het Ministerie van Justitie, Rijswijk, The Netherlands
[2]Forensic Science Research and Training Center, Laboratory Division, Federal Bureau of Investigation, FBI Academy, Quantico, VA, USA
[3]Los Angeles Police Department, Scientific Investigation Division, Los Angeles, CA, USA

Received July 13, 1992 / Received in revised form September 22, 1992

Summary. The HLA DQα amplification and typing kit has been designed to be used by the forensic community for purposes of identity testing. The introduction of any new DNA marker in forensic identity testing requires the establishment of a population database for the relevant population(s) [1]. To this end allele and genotype frequencies for the HLA DQα locus were determined in a Dutch Caucasian population sample and compared with 7 other population genetic studies. In our population sample the HLA DQα genotype frequencies did not deviate from Hardy-Weinberg expectations and for this locus the power of discrimination is 0.94. A test for homogeneity of the HLA DQα population data based on the allele frequency counts for 8 Caucasian population samples was performed and significant differences were found ($P = 0.007$). The differences in the frequency of the HLA DQα 2 and 3 alleles are the major cause of this deviation. No deviation from population homogeneity was observed when we compared the *genotype frequency* distributions among the 8 Caucasian population samples. Combined with the extensive validation studies from Comey and Budowle [7] and Helmuth et al. [8] this population genetic study will allow HLA DQα typing to be used in forensic identity testing in the Netherlands.

Key words: HLA DQα – Polymerase chain reaction (PCR) – Population genetics – Forensic DNA typing

Zusammenfassung. Der Kit für die Amplifikation und Typisierung von HLA DQα wurde als geeignet erklärt für Identitätstests. Die Einführung jedes neuen DNA-Markersystems in forensischen Identitätsuntersuchungen erfordert die Etablierung einer Datenbank für die relevante(n) Population(en) [1]. Mit dieser Zielsetzung wurden Allel- und Genotypfrequenzen in einer holländischen Kaukasier-Stichprobe untersucht und mit 7 anderen populationsgenetischen Studien verglichen. In unserer Populationsstichprobe wichen die HLA DQα-Genotypfrequenzen nicht von Hardy-Weinberg-Erwartungen ab; und für diesen Locus beträgt die Diskriminations-Kraft 0.94. Eine Untersuchung auf Homogenität der HLA DQα-Populationsdaten, basierend auf den Allelfrequenzen für 8 kaukasische Populationsstichproben, wurde durchgeführt und es fanden sich signifikante Unterschiede ($P = 0.007$). Die Differenzen in der Frequenz der Allele HLA DQα 2 und 3 sind die wesentliche Ursache dieser Abweichung. Keine Abweichung von der Homogenität wurde beobachtet, wenn wir die Genotyp-Frequenzverteilungen zwischen den 8 kaukasischen Populationsstichproben verglichen. In Kombination mit den extensiven Validierungsstudien von Comey und Budowle [7] und Helmuth et al. [8] erlaubt die vorliegende populationsgenetische Studie die Typisierung von HLA DQα bei forensischen Identitätsuntersuchungen in den Niederlanden.

Schlüsselwörter: HLA DQα – Polymerase Kettenreaktion – Populationsgenetik – Forensische DNA-Typisierung

Introduction

The enzymatic amplification of polymorphic DNA loci by the Polymerase Chain Reaction (PCR) offers a number of distinct advantages in forensic DNA typing. Firstly this approach provides a means of rapid and specific typing of genetic markers. Secondly, amplification and subsequent detection of target sequences may allow reliable typing results from very small amounts and even degraded DNA from tissues and body fluids. Currently there is one PCR-based genetic marker system that has

Correspondence to: A. D. Kloosterman, Netherlands Forensic Science Institute, Volmerlaan 17, 2288 GD Rijswijk, The Netherlands

been frequently used in actual forensic casework analyses [2, 3]. This system is based on the analysis of the high degree of sequence polymorphism at the HLA DQα locus. The HLA DQα locus is part of the human HLA (Human Leucocyte Antigen) class II region and codes for 3 families of membrane-bound polymorphic glycoproteins (HLA DR, DQ and DP antigens). These molecules are expressed as heterodimers consisting of an α- and a β-chain and are encoded by 2 separate genes (α- and β-genes). The polymorphism at the HLA DQα locus is detected using primers specific to the conserved sequences that flank the hypervariable region. A 242 bp fragment (or a 239 bp fragment for alleles 2 and 4) that contains the polymorphic sequences is amplified. Allele specific oligonucleotide (ASO) probes detect the sequence polymorphism. There are 4 major HLA DQα alleles: HLA DQα1, α2, α3 and α4. The HLA DQα1 and HLA DQα4 alleles can be further subtyped into HLA DQα1.1, α1.2, α1.3 and α4.1, α4.2 and α4.3. Extensive oligotyping results suggest that these alleles represent the complete HLA DQα genetic variation present in the human population [4, 8].

Forensic HLA DQα genotyping has been greatly facilitated by the recent development of a PCR-based kit [5, 6, 7]. This kit (Amplitype™, HLA DQα Forensic DNA Amplification and Typing Kit; Roche Molecular Systems, Alameda, CA, USA) is available commercially and has been validated for forensic casework [6, 7].

The detection assay of this kit is a reverse dot blot procedure. The single-stranded ASO probes are immobilized on a nylon membrane. The PCR primers are labeled with biotin at their 5′ ends. The amplified DNA product (which contains biotin on the 5′ end) hybridizes with the immobilized ASO(s) that are complementary. The sequence-specific binding of the biotin labeled PCR products is detected by a streptavidin-peroxidase conjugate. The peroxidase reaction converts tetramethylbenzidine into a blue spot at a defined position on the nylon membrane.

The current Amplitype system is designed to type 6 of the HLA DQα alleles (HLA DQα1.1, α1.2, α1.3, α2, α3 and α4) and 21 genotypes can be detected [6, 8].

HLA DQα genotype frequencies have been determined in various human populations. However, the validation of this new genetic marker for use in forensic analysis in the Netherlands requires the collection of genotype and allele frequency data from relevant population(s) [1]. This paper presents the data on the genotype and allele frequencies in a Dutch Caucasian population sample. The allele frequency distribution was compared with population genetic studies in 7 other Caucasian population groups (Germany [9], United Kingdom [10], Finland [11], United States [7], Australia [12], Switzerland [13] and Spain [14]).

Materials and methods

Population sample. Blood was obtained from 94 unrelated male and 63 unrelated female Caucasian donors (employees and students of the Dutch Forensic Science Laboratory). A sample of 100 µl liquid blood was deposited on clean cotton weave (20 × 20 mm), air dried and stored at room temperature.

DNA extraction and quantification. DNA was isolated from bloodstains by Chelex extraction using previously described procedures [6, 15]. Chelex-extracted DNA from each sample was quantified using the slot-blot procedure described by Waye et al [16]. All DNA-extracts were normalized to a concentration of 1 ng DNA per µl.

Amplification and typing of the HLA DQα locus. Amplification was carried out using 20 ng DNA. Amplification and typing reactions were performed by strictly following the recommended protocols [6, 7]. Positive and negative (without template DNA) control samples were included in each analysis.

Statistics. The frequency of each allele in the population was calculated from the numbers of each genotype in the sample set. From the allele frequency data the expected number of genotype frequencies was calculated under the assumption of Hardy-Weinberg (H-W) equilibrium expectations. Possible divergence of genotype frequencies from the H-W equilibrium expectations was determined by calculating the unbiased estimate of the expected heterozygote frequency (h). This frequency is equivalent to the allelic diversity [18],

$$h = n\,[1 - \Sigma\,(n_i/n)^2]\,/\,(n-1)$$

where $n_1, n_2 \ldots n_6$ are the allele counts of the HLA DQα alleles in a sample of n chromosomes. In a population following H-W equilibrium expectations h should be equivalent to the frequency of observed heterozygotes. The standard error for h was computed as the square root of the variance of a normal distribution:

$$\sqrt{[h \cdot (1-h)\,/\,N]}$$

where h represents the expected heterozygote frequency and N the number of individuals in the sample. The 95% confidence intervals for the individual allele and genotype frequency data were taken from statistical tables [17] taking into consideration that allele and genotype frequencies are binomially distributed.

The null hypothesis defines that the allele frequencies in the Caucasian population samples are not significantly different. This hypothesis was tested by pairwise comparisons in the chi-squared test of homogeneity ($R \times C$ contingency table [20]) using a computer program provided by G. Carmody (Carleton University, Ottowa, Ontario, Canada).

Population homogeneity of particular alleles was examined using a two allele system in 2 × 2 contingency tables, i.e. the allele of interest in a particular population is category 1, while the sum of all other alleles in that population is category 2.

The power of discrimination (PD) was calculated as $1 - \Sigma(P_i)^2$, were P_i represents the frequency of each genotype.

Results and discussion

The amount of extracted DNA from 3 × 3 mm bloodstains was in the range 50–100 ng. All 157 extracted DNA samples could be typed for the HLA DQα locus.

The distributions of observed genotpye and allele frequencies for the HLA DQα locus in the Dutch Caucasian population sample are shown in Tables 1 and 2. All 21 genotypes were encountered in this population sample. The most common genotypes were HLA DQα1.2–4 ($f = 0.127$) and HLA DQα1.1–4 ($f = 0.108$) and the least frequent genotypes were HLA DQα1.3–1.3, HLA DQα 1.3–2 and HLA DQα2–2 ($f = 0.013$). The most common allele was HLA DQα4 ($f = 0.270$) and the least frequent allele was HLA DQα1.3 ($f = 0.096$).

Table 1. Point estimates and 95% Upper Confidence Limits (UCL) for HLA DQα genotypes in a population sample of 157 Dutch Caucasian individuals. Expected genotype frequencies (shown in parentheses) were calculated on the basis of H-W equilibrium from the allele frequency data in Table 2. Genotype frequencies with expected values of less than 5 were combined into two cells ([1] and [2]) for χ^2 calculation. The values of the power of discrimination and allelic diversity are also shown

HLA-DQα genotype frequencies in the Dutch Caucasian population (N = 157 individuals)			
Genotype	n (n exp)	Frequency	95% UCL
1.1–1.1	7 (6.6)	0.045	0.077
1.1–1.2	13 (12.6)	0.083	0.126
1.1–1.3	4 (6.2)	0.025	0.049
1.1–2	9 (7.1)	0.057	0.093
1.1–3	7 (7.8)	0.045	0.077
1.1–4	17 (17.3)	0.108	0.157
1.2–1.2	7 (6.1)	0.045	0.077
1.2–1.3	6 (5.9)	0.038	0.068
1.2–2	3 (6.9)	0.019	0.038
1.2–3	6 (7.5)	0.038	0.068
1.2–4	20 (16.7)	0.127	0.179
1.3–1.3	2 (1.4)[1]	0.013	0.025
1.3–2	2 (3.3)[1]	0.013	0.025
1.3–3	4 (3.7)[1]	0.025	0.049
1.3–4	10 (8.1)	0.064	0.102
2 –2	2 (1.9)[2]	0.013	0.025
2 –3	5 (4.2)[2]	0.032	0.060
2 –4	12 (9.4)	0.076	0.117
3 –3	3 (2.3)[2]	0.019	0.038
3 –4	10 (10.3)	0.064	0.102
4 –4	8 (11.4)	0.051	0.085
H-W equilibrium (10 degrees of freedom): $\chi^2 = 6.50$ ($0.75 < P < 0.90$)			
Homozygosity (n = 29):		0.185	
Heterozygosity (n = 128):		0.815	
Expected heterozygosity (= allelic diversity):		0.812 ± 0.031	
Power of discrimination:		0.943	

Table 2. Point estimate frequencies (f) and 95% confidence limits for HLA DQα allele frequencies in the Dutch Caucasian Population (n = 314 chromosomes)

HLA DQα allele frequencies in the Dutch Caucasian population			
Allele	f	95% LCL	95% UCL
HLA DQ α1.1	0.207	0.162	0.257
HLA DQ α1.2	0.198	0.150	0.247
HLA DQ α1.3	0.096	0.066	0.136
HLA DQ α2	0.112	0.077	0.151
HLA DQ α3	0.121	0.091	0.170
HLA DQ α4	0.270	0.220	0.322

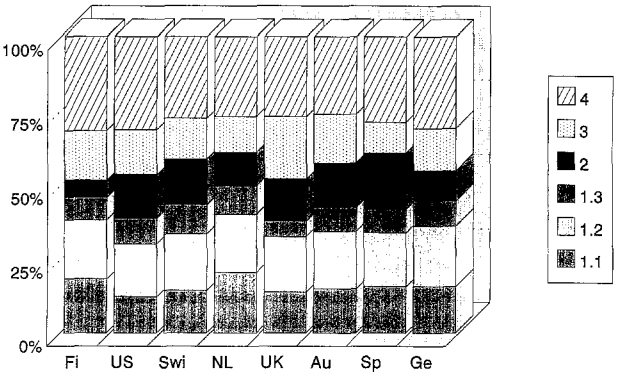

Fig. 1. HLA DQα allele frequencies in 8 different Caucasian populations
Fi = Finland (N = 112)
US = United States (N = 150)
Swi = Switzerland (N = 200)
Nl = The Netherlands (N = 157)
UK = United Kingdom (N = 201)
Au = Australia (N = 280)
Sp = Spain (N = 120)
Ge = Germany (N = 212)
N = number of individuals typed
Bars represent point estimate frequency data

The observed number of genotypes as well as the expected number of genotypes based on the assumption of H-W expectations for this population are shown in Table 1. The chi-squared test was used to compare observed and expected genotype numbers. For this population sample the P value was $0.75 < P < 0.90$ ($chi^2 = 6.5$; $df = 10$). Thus, this population sample was in agreement with the expectations of H-W equilibrium and indicates that genotype frequencies can be reliably estimated from the allele frequency data. It also suggests that there is no evidence for null (undetected) alleles in this population sample and therefore there is no evidence for mistyping caused by the allelic dropout phenomenon (the failure to amplify one of the alleles in a heterozygote). Dropout of all the HLA DQα1 alleles in some HLA DQα1–4 heterozygotes has been reported [7].

The allelic diversity (an estimate of the expected heterozygote frequency) was 0.812 ± 0.03. The observed heterozygosity (0.815) did not deviate from the expected value of h. The power of discrimination for forensic identity testing was 0.943 in the Dutch population.

A test for homogeneity of HLA DQα population data based on *genotype observations*, among 8 Caucasian population samples (i.e., Dutch, Finnish, American, Swiss, English, Australian, Spanish and German) was performed using an $R \times C$ contingency table. It was found that the combined groups are statistically similar ($P = 0.438 \pm 0.019$). Furthermore, since there was no statistical difference between observed and expected genotype frequencies among the 8 population samples, these populations can be considered homogeneous.

Even though there was statistical homogeneity, some allele frequencies appeared to be different. A test for homogeneity of the HLA DQα population data based on the *allele frequency* counts was also performed for the 8 Caucasian population samples using a $R \times C$ contingency

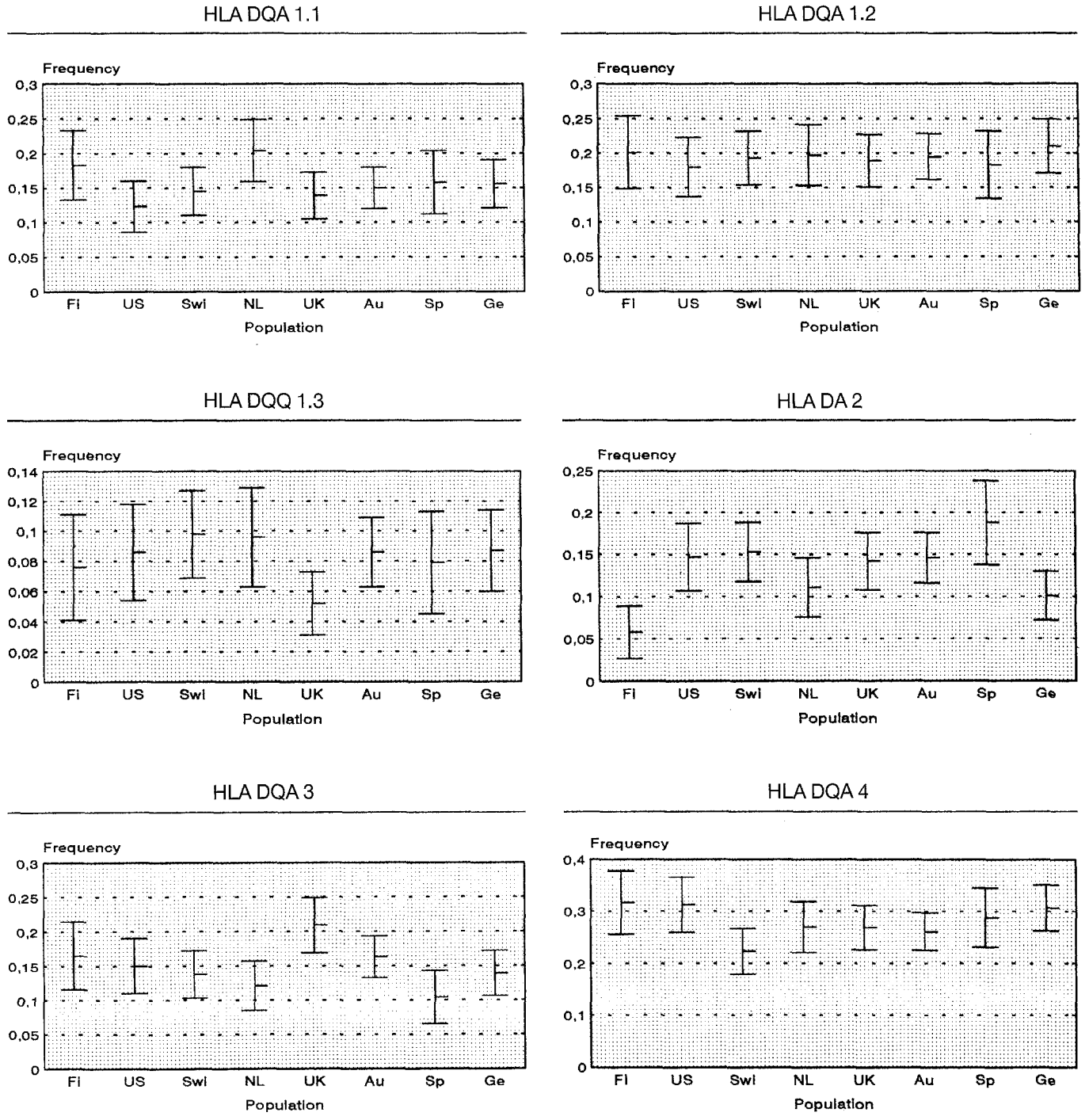

Fig. 2. HLA DQα allele frequencies with confidence intervals in 8 different Caucasian populations (see Fig. 1 for abbreviations and sizes of population samples). Line segments represent 95% confidence intervals of *f*

table. It was found that the combined groups were significantly different in this test ($P = 0.007 \pm 0.003$).

On closer examination it appeared that the differences in the frequency of the HLA DQα2 ($P < 0.001$) and the HLA DQα3 ($P = 0.020 \pm 0.004$) alleles are the major cause of this deviation. Figure 1 shows a qualitative comparison with respect to the distribution of the HLA DQα allele frequency data for the 8 population samples. In most cases the allele frequencies overlapped, based on their 95% confidence intervals (Fig. 2). However, the HLA DQα2 allele in the Finnish population deviated from the 2 allele frequency in all other Caucasian population samples except the German population sample.

The HLA DQα3 allele in the English population sample deviated from the 3 allele frequency in the Swiss, Dutch, Spanish and German population samples. Table 3 gives the P values for the pairwise comparisons of the HLA DQα allele frequencies among the 8 Caucasian population samples.

Table 3. *P* values for pairwise comparisons of HLA DQα allele frequencies among 8 Caucasian population samples. The particular alleles that deviate significantly from population homogeneity are given in the bottom rows (*** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$)

Fi = Finland	($N = 112$)	Sajantila et al. [11]
US = United States	($N = 150$)	Comey et al. [7]
Swi = Switzerland	($N = 200$)	Budowle et al. [13]
Nl = The Netherlands	($N = 157$)	this study
UK = United Kingdom	($N = 201$)	Sullivan et al. [10]
Au = Australia	($N = 280$)	Harrington et al. [12]
Sp = Spain	($N = 120$)	Lorente et al. [14]
Ge = Germany	($N = 212$)	Schneider et al. [9]

N refers to the number of individuals typed

Finland						
US	Swi	Nl	UK	Au	Sp	Ge
0.032 2***	0.010 2***	0.160 2*	0.012 2***	0.016 2***	0.002 2***	0.045
United States						
Fi	Swi	Nl	UK	Au	Sp	Ge
0.032	0.840 1.3*	0.084 1.1*	0.150	0.670	0.414	0.434
Switzerland						
Fi	US	Nl	UK	Au	Sp	Ge
0.010 2***	0.840 1.3*	0.296 1.1*	0.050 1.3* 3**	0.874	0.624	0.326
The Netherlands						
Fi	US	Swi	UK	Au	Sp	Ge
0.160 2*	0.084 1.1*	0.296 1.1*	<0.001 1.1* 1.3* 3**	0.190 1.1*	0.190	0.548
United Kingdom						
Fi	US	Swi	Nl	Au	Sp	Ge
0.012 2***	0.150	0.050 3**	<0.001 1.1* 1.3* 3**	0.236	0.020 3**	0.012 3*
Australia						
Fi	US	Swi	Nl	UK	Sp	Ge
0.016 2***	0.670	0.874 3**	0.190 1.1*	0.236	0.264 3*	0.248
Spain						
Fi	US	Swi	Nl	UK	Au	Ge
0.002 2***	0.414	0.624	0.190	0.020 3**	0.264 3*	0.054 2**
Germany						
Fi	US	Swi	Nl	UK	Au	Sp
0.045	0.434	0.326	0.548	0.012 3*	0.248	0.054 2**

The following examples may illustrate the effects of the HLA DQα allele frequency differences among the Caucasian population samples.

1. This maximum possible discrepancy among the 8 population groups for the expected most common HLA DQα genotype frequency (DQα1.2–4) was observed between the United Kingdom ($f_{exp} = 0.101$; 1 in 9.9) and the German ($f_{exp} = 0.127$; 1 in 7.9) population samples.

2. The highest discrepancy among the 8 population groups was observed when the expected HLA DQα2–2 genotype frequencies were compared and occurred between the Finnish ($f_{exp} = 0.0034$; 1 in 294) and the Spanish ($f_{exp} = 0.035$; 1 in 28.6) population samples. It should be noted here that the observed 2–2 genotype frequencies are rare in both the Finnish (1 observation in 112 individuals; $f_{obs} = 0.0089$) and the Swiss (3 observations in 400 individuals; $f_{obs} = 0.0075$) population samples. Furthermore, it should be emphasized that the numbers of individuals in the sample sets are relatively small which can increase the probability of sampling errors. However, other population studies also have shown that both the distribution of HLA-class I and class II allele frequencies vary between various human populations from the same racial group [8, 19]. By obtaining more data from population genetic studies we will gain additional insight into the population structure of the human race with respect to this genetic marker system.

In conclusion, the Dutch data for HLA DQα are comparable with other Caucasian population studies and is therefore appropriate for the general case situation in the Netherlands.

Combined with the extensive validation studies from Comey and Budowle [7] and Helmuth [8] this population genetic study should allow the use of HLA DQα typing in forensic identity testing in the Netherlands.

References

1. DNA commission of the International Society for Forensic Haemogenetics (1992) Second DNA recommendations. Int J Leg Med 104:361–364
2. Higuchi R, von Beroldingen CH, Sensabough GF, Erlich HA (1988) DNA typing from single hairs. Nature 332:543–546
3. von Beroldingen CH, Blake ET, Higuchi R, Sensabaugh GF, Erlich HA (1989) Applications of PCR to the analysis of biological evidence. In: Erlich HA (ed) PCR Technology, principles and applications for DNA Amplification. Stockton Press. Chapter 17.
4. Saiki RK, Walsh PS, Levenson CH, Erlich HA (1989) Genetic analysis of amplified DNA with immobilized sequence-specific olignucleotide probes. Proc Natl Acad Sci USA 86:6230–6234
5. Bugawan TL, Saiki RK, Levenson CH, Watson RM, Erlich HA (1988) The use of non-radioactive oligonucleotide probes to analyze enzymatically amplified DNA for prenatal diagnosis and forensic HLA typing. Biotechnology 6:943–947
6. Amplitype™ User Guide – Version 2 (1990) Cetus Corporation, Emmeryville, CA, USA
7. Comey CT, Budowle B (1991) Validation studies on the analysis of the HLA DQα locus using the polymerase chain reaction. J Forensic Sci 36:1633–1648
8. Helmuth R, Fildes N, Blake E, Luce MC, Chimera J, Madej R, Gorodezky C, Stoneking M, Schmill N, Klitz W, Higuchi

R, Erlich HA (1990) HLADQα allele and genotype frequencies in various human populations, determined by using enzymatic amplification and oligonucleotide probes. Am J Hum Genet 47:515–523
9. Schneider PM, Prager-Eberle M, Rittner C (1991) Zur Anwendung der Polymerase Kettenreaktion (PCR) des HLADQα-Systems in der forensischen Spurenkunde. Arch Kriminol 188: 167–174
10. Sullivan KM, Gill P, Lingard D, Lygo JE (1992) Characterization of HLADQα for forensic purposes. Allele and genotype frequencies in British Caucasian, Afro-Caribbean and Asian populations. Int J Leg Med 105:17–20
11. Sajantila A, Ström M, Budowle B, Tienari PJ, Ehnholm C, Peltonen L (1991) The distribution of the HLADQα alleles and genotypes in the Finnish population as determined by the use of DNA amplification and allele specific oligonucleotides. Int J Leg Med 104:181–184
12. Harrington CS, Dunaiski V, Williams K, Fowler C (1991) HLADQα typing of forensic specimens by amplification restriction fragment polymorphism (ARFP) analysis. Forensic Sci Int 51:147–157
13. Budowle B, Sajantilla A, Hochmeister MN, Comey CT (1992) The application of PCR to forensic science. In: Ferre F, Mullis K and Gibbs R (eds). The Polymerase chain reaction. Birkhauer Verlag, Basel (in press)
14. Lorente JA, Lorente M, Budowle B, Comey CT, Villanueva E (1992) HLADQα types in the Spanish population. Proceedings of Promega 1992 International Symposium on Human Identification. Scottsdale, Ariz, USA
15. Jung JM, Comey CT, Bear DB, Budowle B (1991) Extraction strategy for obtaining DNA from bloodstains for PCR amplification and typing of the HLADQα gene. Int J Leg Med 104:145–148
16. Waye JS, Presley LA, Budowle B, Shutler GG, Fourney RM (1990) A simple and sensitive method for quantifying human genomic DNA in forensic specimen extracts. Biotechniques 7:852–855
17. Documenta Geigy; Wissenschaftliche Tabellen; 7th edn (1968). Geigy Basel, pp 85–103
18. Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76:379–390
19. Piazza A, Menozzi P, Cavalli-Sforza LL (1980) The HLA-A,B gene frequencies in the world: migration or selection? Hum Immunol 1:297–304
20. Roff DA, Bentzen P (1989) The statistical analysis of mitochondrial DNA-polymorphisms and the problem of small samples. Mol Biol Evol 6:539–545