Int J Leg Med (1991) 104:181–184

International Journal of
Legal Medicine

Original articles

The distribution of the HLA-DQα alleles and genotypes in the Finnish population as determined by the use of DNA amplification and allele specific oligonucleotides

A. Sajantila[1,2], M. Ström[2], B. Budowle[3], P. J. Tienari[1], C. Ehnholm[1], and L. Peltonen[1]

[1] National Public Health Institute, Mannerheimintie 166, SF-00300 Helsinki, Finland
[2] Department of Forensic Medicine, University of Helsinki, Kytösuontie 11, SF-00300 Helsinki, Finland
[3] Forensic Science Research and Training Center, FBI Academy, Quantico, VA 22135, USA

Received December 10, 1990 / Received in revised form March 18, 1991

Summary. Allele and genotype frequencies for the HLA-DQα locus were determined for use in forensic analyses and paternity tests in Finland. The polymerase chain reaction (PCR) and the reverse dot blot format were employed to detect 6 different HLA-DQα alleles. All 6 HLA-DQα alleles were detected among the 112 unrelated individuals with allele frequencies ranging from 5.8% to 32.6%. The distribution of the observed genotypes is in Hardy-Weinberg equilibrium. Additionally, this Finnish population sample is statistically similar to 2 other Caucasian sample populations. The power of discrimination of this system in the Finnish population sample is 0.92, suggesting this method may prove suitable for identification purposes.

Key words: Polymerase chain reaction (PCR) – HLA-DQα – Allele specific oligonucleotides – Population genetics – Finnish population

Zusammenfassung. Allelfrequenzen und Genotypfrequenzen für den HLA-DQα-Locus wurden bestimmt mit der Zielsetzung der Anwendung in Spurenanalysen und Vaterschaftstests in Finnland. Die Polymerase-Kettenreaktion (PCR) und das reverse Dot-Blot-Format wurden benutzt, um 6 verschiedene HLA-DQα-Allele nachzuweisen. Alle 6 HLA-DQα-Allele wurden bei 112 unverwandten Personen nachgewiesen, die Allelfrequenzen waren zwischen 5,8% und 32,6%. Die Verteilung der beobachteten Genotypen ist im Hardy-Weinberg-Gleichgewicht. Weiterhin ist diese finnische Populationsstudie statistisch ähnlich zu 2 anderen europäischen Bevölkerungsstichproben. Die Diskriminationskraft mit diesem System innerhalb der finnischen Bevölkerung ist 0,92, was daran denken läßt, daß diese Methode sich als brauchbar für Identifikationszwecke erweisen läßt.

Offprint requests to: A. Sajantila at the first address

Schlüsselwörter: Polymerase-Kettenreakion (PCR) – HLA-DQα – Allel-spezifische-Oligonucleotide – Populationsgenetik – Finnische Bevölkerung

Introduction

Genetic characterization of individuals for identity testing is increasingly being performed at the DNA level. The basic method for DNA identification in the forensic arena has so far involved analysis of restriction fragment length polymorphisms (RFLP). Although informative and reliable, RFLP analysis is laborious and time consuming, requiring in addition to DNA purification, enzyme digestion, electrophoresis, Southern blotting and radioactively labeled probes [1, 2].

A new, technically more feasible strategy for individual identification at the DNA level is the use of the polymerase chain reaction (PCR), i.e. the enzymatic amplification of polymorphic DNA loci [3, 4, 5, Sajantila et al. submitted for publication]. The advantages of the PCR technique over the classical RFLP technique for forensic DNA analysis are that it is more sensitive and specific, less time-consuming, less laborious and the analysis does not require radioactively labeled probes. An additional advantage of PCR is that it enables the analysis of minute amounts of DNA from various sources including impure or degraded samples [6, 7, Sajantila et al. submitted for publication].

One approach for the analysis of forensic samples by PCR is based on the human leukocyte antigen (HLA) DQα locus which codes for the polymorphic α-chain of the HLA-DQ molecule. The HLA-DQ molecule is a heterodimer composed of one α-chain and one β-chain and it is expressed in B-lymphocytes, macrophages, thymic epithelium and activated T-cells. It serves as an

integral membrane protein for binding as well as presenting antigen peptide fragments to the T cell receptor of $CD4^+$ T lymphocytes. At the DNA level most of the polymorphism of HLA-DQ α-chain locus is localized on the second exon encoding the NH_2-terminal outer domain [8]. Eight alleles have been identified in the HLA-DQα locus and 6 of these can be distinguished with the method applied here [9].

This paper presents data on the frequencies of 6 different HLA-DQα alleles (DQA 1.1, 1.2, 1.3, 2, 3 and 4) and the distribution of the different genotypes in a Finnish population sample. The power of discrimination (PD) was determined to evaluate its applicability for forensic identity testing. To establish the use of the observed genotypes in the calculations of the probability of a random match, the database was assessed to determine whether or not it was in Hardy-Weinberg equilibrium. Also heterozygosity and homozygosity were addressed.

Materials and methods

Nomenclature. According to the most recent nomenclature system adopted by the World Health Organization (WHO) HLA Nomenclature Committee [10], the HLA-DQα locus is now designated as DQA1 and the previously known DXα locus is now termed DQA2. To avoid confusion between this recently proposed nomenclature and the designation generally used in forensic DNA HLA-DQα genotyping the following allele designations were used as suggested by Helmuth et al. [11]: DQA1.1 (WHO nomenclature DQA1*0101), DQA1.2 (WHO nomenclature DQA1*0102), DQA1.3 (WHO nomenclature DQA1*0103), DQA2 (WHO nomenclature DQA1*0201), DQA3 (WHO nomenclature DQA1*0301) and DQA4 (WHO nomenclature DQA1*0401, DQA1*0501 and DQA1*0601).

DNA amplification and HLA-DQα typing. DNA was extracted from fresh peripheral blood leukocytes from 112 unrelated Finns using a slightly modified method from that described by Vandenplas et al. [12]. The amplification of the HLA-DQα locus and the analysis of the alleles were performed using the AmpliType HLA-DQα Forensic DNA Amplification and Typing Kit (Cetus Corporation, CA, USA). Basically, the procedure involves amplification of the HLA-DQα locus by PCR using biotinylated primers. The amplified product is subsequently hybridized to a filter strip carrying immobilized allele specific oligonucleotide DNA probes. Detection is enzymatically mediated by a streptavidin-horseradish peroxidase conjugate resulting in a visually detectable dye. This technique is known as a "reverse dot blot" format (Fig. 1).

PCR amplification of 25 ng of template DNA was performed using a thermal cycler (Perkin-Elmer Cetus) programmed for the following steps: denaturation at 94°C for 1 min, annealing at 60°C for 30 s and extension at 72°C for 30 s. A total of 32 cycles was performed in a volume of 50 µl. To ensure the quality of the amplification product each sample was analyzed on an agarose gel containing ethidium bromide. One positive and one negative (no template DNA) control sample were included in each analysis.

Hybridization was performed in a solution containing 5 × SSPE (1:4 dilution of 3.6 *M* NaCl-200 m*M* NaH_2PO_4*H_2O-20 m*M* EDTA, pH 7.4) – 0.5% w/v SDS (sodium dodecyl sulfate), streptavidin-horseradish peroxidase conjugate and 35 µl of the amplified DNA at 55°C for 20 min. After hybridization 3 washing steps were performed with 2.5 × SSPE – 0.1% SDS solution. The filterstrips were first rinsed at room temperature followed by a stringent wash at 55°C for 12 min and finally at 5 min wash at room temperature. The color development was accomplished in a solution containing 0.1 *M* sodium citrate pH 5.0, 0.003% H_2O_2 and 1 mg of 3,3',5,5'-tetramethylbenzine (TMB).

Statistical evaluations. The power of discrimination (PD) was calculated from the genotype data on the basis of an equation derived by Fisher ($PD = 1 - \sum P_j^2$, where P_j is the frequency of each genotype) [13]. The allelic diversity values (h) were calculated as described by Nei and Roychoudhury [14] ($h = [1 - \sum X_i^2][n/n - 1]$, where X_i = allele frequencies and n = number of alleles). Chi-square test was performed to assess whether our population database conformed to Hardy-Weinberg equilibrium expectations. A program kindly provided by G. Carmody (Carleton University, Ottawa, Canada) to test population data homogeneity by chi-square analysis and G-statistic was used.

Results

Analysis of the HLA-DQα allele distribution in the Finnish population revealed that the most common HLA-DQα allele was HLA-DQA 4 (frequency, $f = 0.33$) and the least frequent allele was HLA-DQA 2 ($f = 0.06$). The other alleles HLA-DQA1.1, DQA1.2, DQA1.3, and DQA3 occurred at frequencies of 17.4%, 20.1%, 6.7% and 17.4%, respectively (Fig. 2). The most common genotypes in the Finnish population were DQA1.2,4, DQA1.1,4 and DQA3,4. The frequencies of these genotypes were 15.2%, 13.4% and 11.6%, respectively. The genotypes HLA-DQA1.3,1.3 and HLA-

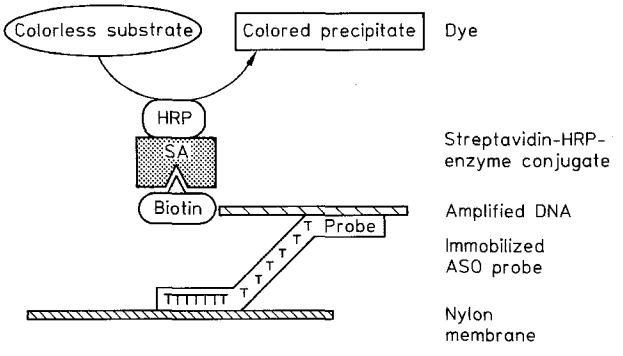

Fig. 1. A schematic presentation (Adopted from ref. [19]) of the generation of a dot blot format using immobilized allele-specific oligonucleotide probes (ASO). HRP = horseradish peroxidase

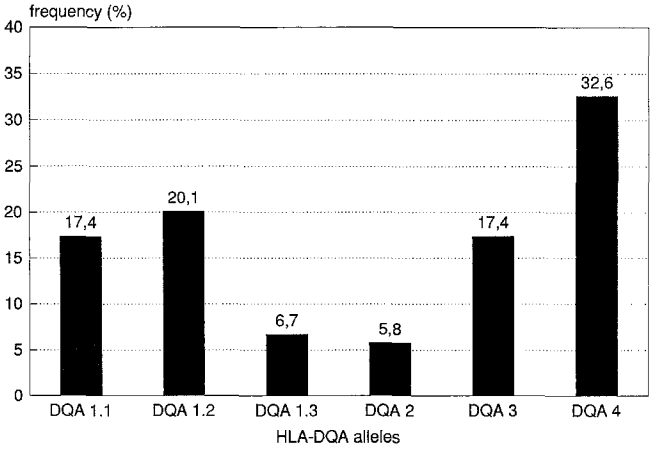

Fig. 2. Distribution of HLA DQα alleles in the Finnish population

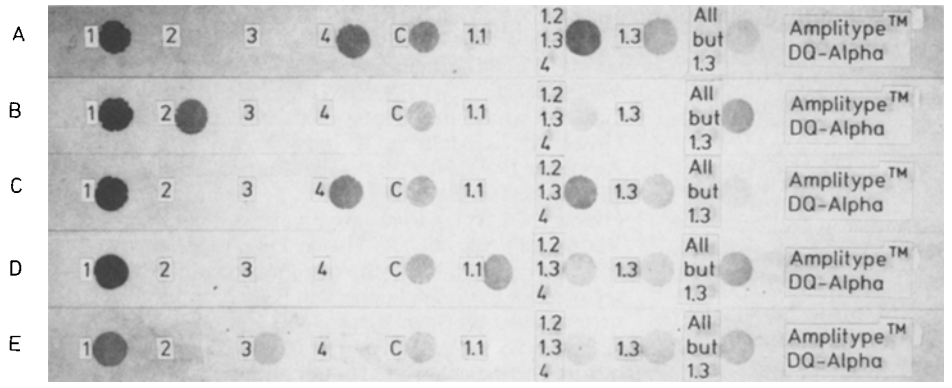

Fig. 3. Examples of HLA-DQα genotyping on the dot blot nylon membrane strips. The HLA-DQα genotypes analyzed here are *A* = HLA-DQA1.3,4; *B* = HLA-DQA1.2,2; *C* = 1.3,4; *D* = 1.1,1.3 and *E* = 1.3,3

Table 1. Observed and expected HLA-DQα genotypes in the Finnish population

Genotype	Observed		Expected		χ^2
	n	%	*n*	%	
1.1–1.1	5	4.5	3.4	3.0	0.75
1.1–1.2	7	6.3	7.8	7.0	0.08
1.1–1.3	3	2.7	2.6	2.3	0.06
1.1–2	3	2.7	2.3	2.0	0.21
1.1–3	3	2.7	6.8	6.1	2.12
1.1–4	15	13.4	12.7	11.3	0.42
1.2–1.2	3	2.7	4.5	4.0	0.50
1.2–1.3	4	3.6	3.0	2.7	0.33
1.2–2	3	2.7	2.6	2.3	0.06
1.2–3	8	7.1	7.8	6.9	0.01
1.2–4	17	15.2	14.7	13.1	0.36
1.3–1.3	0	0	0.5	0.4	*
1.3–2	0	0	0.9	0.8	*
1.3–3	4	3.6	2.6	2.3	0.75
1.3–4	6	5.4	4.9	4.4	0.25
2–2	1	0.9	0.4	0.3	*
2–3	1	0.9	2.3	2.0	*
2–4	4	3.6	4.2	3.8	0.01
3–3	4	3.6	3.4	3.0	0.11
3–4	13	11.6	12.7	11.3	0.01
4–4	8	7.1	11.9	10.6	1.28
* Classes 1.3–1.3, 1.3–2, 2–2 and 2–3 pooled together (*n* observed = 2, *n* expected = 4.1)					1.08
				Total	8.39**
** $\chi^2 = 8.39$, $0.75 < P < 0.90$, $df = 12$					
Heterozygotes	$n = 91$	(81.0%)			
Homozygotes	$n = 21$	(19.0%)			
Allelic diversity (h)	= 0.78				
Power discrimination (PD)	= 0.92				

DQA1.3,2 were not observed among the 112 individuals analyzed (Table 1).

The chi-square test was used to compare the observed and expected genotypes (calculated on the basis of Hardy-Weinberg equilibrium from allele frequencies in Fig. 2.). The population data genotypes are in agreement with the expected genotypes and conform to Hardy-Weinberg equilibrium expectations ($0.75 < P < 0.90$). The genotype frequencies used in this analysis are given in Table 1.

The allelic diversity (h) in the Finnish population was 0.78 (Table 1). This is an unbiased estimate of the expected heterozygosity calculated as described by Nei and Roychoudhury [14]. The observed heterozygosity in our population survey was 0.81 (Table 1).

The power of discrimination (PD) for forensic identity testing using the HLA-DQα genotyping alone was 0.92 in the Finnish population (Table 1).

A statistical test for significance of a 2-way R×C contingency table to determine population sample homogeneity was used. The data suggests that the Finnish sample is statistically similar to the two U.S. Caucasian samples described by Helmuth et al. [11], ($\chi^2 = 41.87$, $P = 0.3930 \pm 0.0154$ (S.E.); G-statistic = 45.70, $P = 0.3650 \pm 0.0152$ (S.E.)).

Discussion

The evaluation of the significance of a detected match in forensic problems and paternity tests between genetically typed samples is less meaningful unless the population distribution of alleles and phenotypes have been determined. When the sample analyzed represents a distinctly different phenotype than the reference sample, the "exclusion" is independent of the phenotype frequencies obtained in the particular case. In contrast, in cases where an "inclusion" has been determined the probability that the phenotype from another unrelated individual could also match the evidence sample is equal to the frequency of that phenotype in the appropriate population.

The Finnish population represents an example of an isolated population and genetic isolation has been demonstrated with different polymorphic protein markers [15]. At the DNA level this genetic isolation has been studied by the RFLP technique with highly polymorphic DNA probes. Both multilocus and singlelocus probes have failed to show any significant differences in DNA length polymorphisms between the Finnish population and more heterogeneous Caucasian populations [16, 17, Sajantila et al. submitted for publication].

The HLA-DQα data fails to show any significant differences when compared to U.S. Caucasian data. In

contrast, the frequency distribution of HLA-DQα alleles varies significantly between other racial groups [11], consequently differences in the genotype frequencies can be expected. The most striking difference in allele distributions is seen with the Japanese population confirming the observations of Helmuth et al. [11]. The frequency of HLA-DQA1.3 in Japanese is significantly higher (22.8%) than in the Finnish population (6.75%) and other populations studied (0–8.5%). Furthermore, the frequency of the HLA-DQA2 allele is lower in Japanese (0.5%) than in the Finnish population (5.8%) and other populations studied (5–13.5%), as is the frequency of the HLA-DQA4 allele (11.4% in the Japanese population, 25–40% in other populations, 32.6% in the Finnish population).

As data derived from forensic and paternity DNA profiling tests could have an impact on a case, it is important to establish a database for the appropriate population for a genetic marker that can potentially be used for identification. It is also desirable to evaluate (when possible) whether the population is in Hardy-Weinberg equilibrium for the genetic marker in question. This type of evaluation has been difficult to carry out for RFLP analyses with single-locus probes due to the difficulties in distinguishing between different alleles [18]. As PCR analysis of the HLA-DQα locus provides discrete allele determination, it is possible to demonstrate that the distribution of genotypes in different populations is generally in agreement with Hardy-Weinberg equilibrium [11]. This is also the case in the Finnish population sample.

Additionally, the observed frequency of homozygotes (19%) is not significantly different from that expected (22%).

In conclusion, we have established a population database of the HLA-DQα locus for forensic use. This locus is in Hardy-Weinberg equilibrium and the discrimination power of this method alone is 0.92 in the Finnish population. Additionally, the distribution of the genotypes in Finns is similar to the U.S. Caucasian population. The reverse dot blot analysis is simple, quick and does not require special technical skills. Furthermore, small amounts of crude sample can be used. The method appears to be a valuable screening test and may prove useful as an approach to resolve forensic and other identification problems.

Acknowledgements. We are grateful to Dr. A.-C. Syvänen and Dr. A. F. Vuorio for reading the manuscript and for critical comments. We would also like to thank Mr. J. Pulliainen for photographs, and Dr. C. T. Comey for technical advice in the DNA analyses and Professor K. Aho for constant help during the study.

References

1. Southern EM (1975) Detection of sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503–517
2. Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13
3. Mullis KB, Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. In: Wu R, Grossman L, Moldave K (eds) Methods in Enzymology. Academic Press, New York, vol 155, pp 335–350
4. Higuchi R, von Beroldingen CH, Sensabaugh GF, Erlich HA (1988) DNA typing from single hairs. Nature 332:543–546
5. Vuorio AF, Sajantila A, Hämäläinen T, Syvänen AC, Ehnholm C, Peltonen L (1990) Amplification of the hypervariable region close to the apolipoprotein B gene: application to forensic problems. Biochem Biophys Res Commun 170:616–620
6. Westwood SA, Werrett DJ (1990) An evaluation of the polymerase chain reaction method for forensic applications. Forensic Sci Int 45:201–215
7. 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
8. Kappes D, Strominger JL (1988) Human class II major histocompatibility complex genes and proteins. Annu Rev Biochem 57:991–1028
9. Saiki RK, Walsh PS, Lewenson CH, Erlich HA (1989) Genetic analysis of amplified DNA with immobilized sequence specific oligonucleotide probes. Proc Natl Acad Sci USA 86:6230–6234
10. Bodmer JG, Marsh SGE, Parham P, Erlich HA, Albert E, Bodmer WF, Dupont B, Mach B, Mayr WR, Sasazuki T, Schreuder GMTH, Strominger JL, Sveigaard A, Terasaki PI (1990) Nomenclature for factors of the HLA system, 1990. Tissue Antigens 35:1–8
11. 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) HLA-DQα allele and genotype frequencies in various human populations, determined using enzymatic amplification and oligonucleotide probes. Am J Hum Genet 47:515–523
12. Vandenplas S, Wiid I, Grobler-Rabie A, Brebner K, Ricketts M, Wallis G, Bester A, Boyd C, Mathew C (1984) Blot hybridization analysis of genomic DNA. J Med Genet 21:164–172
13. Fisher RA (1951) Standard calculations for evaluating a blood group system. Heredity 5:95–102
14. Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76:379–390
15. Nevanlinna HR (1972) The Finnish population structure – A genetic and genealogical study. Hereditas 71:195–236
16. Helminen P, Ehnholm C, Lokki M-J, Jeffreys AJ, Peltonen L (1988) Application of DNA "fingerprints" to paternity determinations. Lancet 574–576
17. Smith JC, Anwar J, Riley D, Jenner B, Hopkins B, Markham AF (1990) Highly polymorphic minisatellite sequences: allele frequencies and mutation rates for five locus specific probes in a Caucasian population. J Forensic Sci Soc 30:19–32
18. Odelberg SJ, Plaetke R, Eldridge JR, Ballard L, O'Connell P, Nakamura Y, Leppert M, Lalouel J-M, White R (1989) Characterization of eight VNTR loci by agarose gel electrophoresis. Genomics 5:915–924
19. Erlich HA, Bugawan TL (1989) HLA class II gene polymorphism: DNA typing, evolution, and relationship to disease susceptibility. In: Erlich HA (ed) PCR technology principles and applications for DNA amplification. Stockton Press, New York, pp 193–208