TISSUE ANTIGENS
IMMUNE RESPONSE GENETICS

Tissue Antigens ISSN 0001-2815

HLA-A, HLA-B, and HLA-DRB1 allele distribution in a large Armenian population sample

L. Matevosyan [1,2], S. Chattopadhyay[3], V. Madelian[4], S. Avagyan[1], M. Nazaretyan[1], A. Hyussian[1], E. Vardapetyan[2], R. Arutunyan[2] & F. Jordan[4]

1 Armenian Bone Marrow Donors Registry, Yerevan, Republic of Armenia
2 Department of Genetics and Cytology, State University, Yerevan, Republic of Armenia
3 Department of Microbiology, University of Washington, Seattle, WA 98195, USA
4 Armenian Bone Marrow Donor Registry, Los Angeles, CA 90039, USA

Key words
allele frequency; Armenian populations; DNA typing; genetic distances; human leukocyte antigen; population genetics

Correspondence
V. Madelian
Armenian Bone Marrow Donor Registry
3111 Los Feliz Blvd
Suite 206
Los Angeles
CA 90039
USA
Tel: +818 704 0072
Fax: 909-623-9306
e-mail: vmadelian@yahoo.com

Received 28 March 2010; revised 21 January 2011; accepted 27 February 2011

doi: 10.1111/j.1399-0039.2011.01668.x

Abstract

Human leukocyte antigen (HLA)-A, HLA-B, and HLA-DRB1 gene frequencies were investigated in 4279 unrelated Armenian bone marrow donors. HLA alleles were defined by using PCR amplification with sequence specific primers (PCR-SSP) high- and low-resolution kits. The aim of this study was to examine the HLA diversity at the high-resolution level in a large Armenian population sample, and to compare HLA allele group distribution in Armenian subpopulations. The most frequently observed alleles in the HLA class I were *HLA-A*0201*, *A*0101*, *A*2402*, *A*0301*, *HLA-B*5101*, *HLA-B*3501*, and *B*4901*. Among DRB1 alleles, high frequencies of *DRB1*1104* and *DRB1*1501* were observed, followed by *DRB1*1101* and *DRB1*1401*. The most common three-locus haplotype found in the Armenian population was *A*33-B*14-DRB1*01*, followed by *A*03-B*35-DRB1*01*. Our results show a similar distribution of alleles in Armenian subpopulations from different countries, and from different regions of the Republics of Armenia and Karabagh. The low level of genetic distances between subpopulations indicates a high level of population homogeneity, and the genetic distances between Armenians and other populations show Armenians as a distinct ethnic group relative to others, reflecting the fact that Armenians have been an 'isolated population' throughout centuries. This study is the first comprehensive investigation of HLA–allele group distribution in a subset of Armenian populations, and the first to provide HLA–allele and haplotype frequencies at a high-resolution level. It is a valuable reference for organ transplantation and for future studies of HLA-associated diseases in Armenian populations.

Introduction

Human leukocyte antigen (HLA) polymorphism has been analyzed for several decades to investigate human genetic relationships (1). This study was undertaken to establish the Armenian HLA gene profile on a large population sample and to compare Armenian subpopulations by investigating HLA class I and class II polymorphisms.

Armenia is one of the oldest countries in the world. It has a recorded history of about 3500 years. Over its history, it has seen many invasions, and it, in turn, has had its own periods of expansion, at times reaching and engulfing lands as far away as ancient Rome. Bordered by the Republic of Georgia, the Islamic Republic of Iran, and the Republics of Azerbaijan and Turkey, the present day Republic of Armenia and to its east, the Republic of Mountainous Karabagh are situated in the northeastern portion of the Armenian Highland (Figure 1), a region of jagged volcanic mountains, ravines and narrow passes, an austere climate, and a population comprising >95% ethnic Armenians (2).

Through centuries, many factors have isolated Armenians from their neighboring populations. Chief among these have been its topography, language, and religion. Thus, in spite of repeated invasions, few invaders have braved its harsh weather and settled in permanently. Instead, once invaded, the land was annexed to the invading empire (Persian, Byzantine, Roman, and Ottoman), and was ruled by the locals. The Armenian language, an Indo–European language, and as such radically different from those of its historic and present day

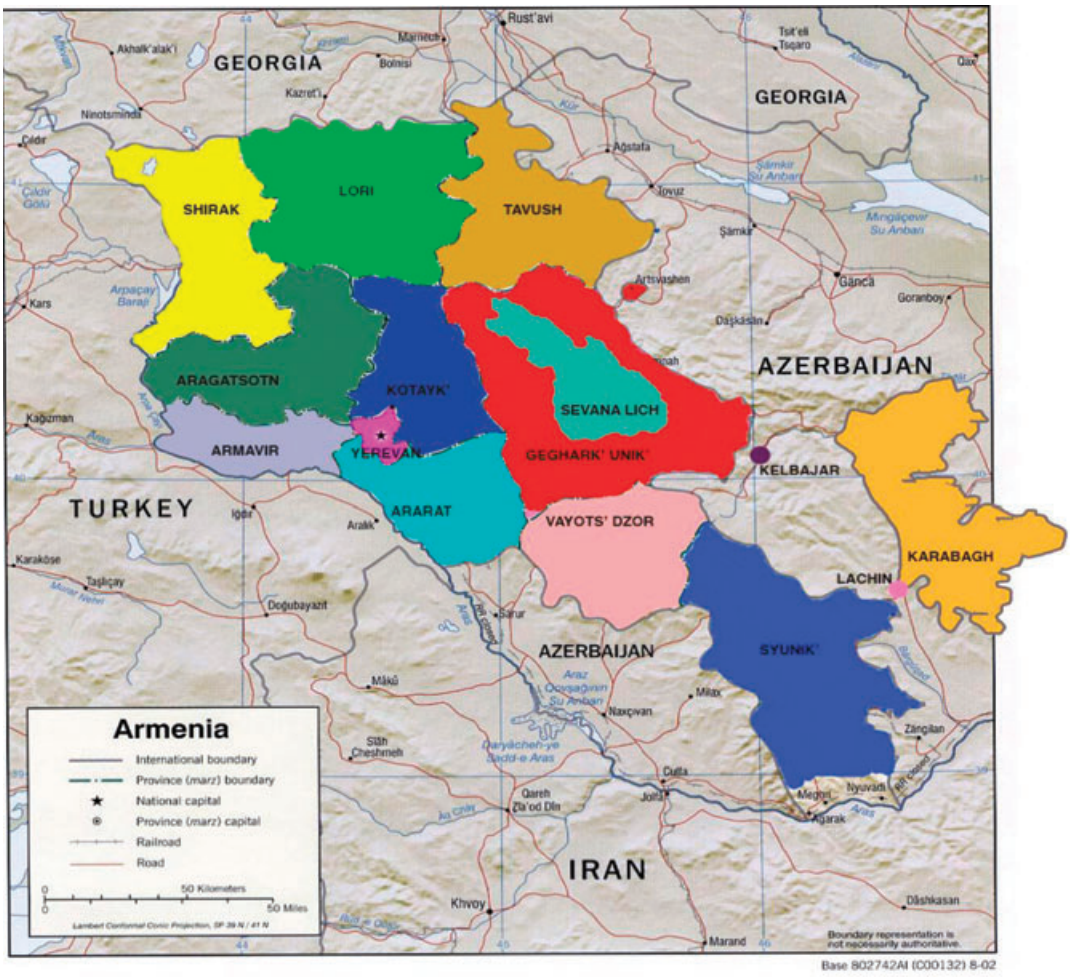

Figure 1 Map of the Republic of Armenia, with different colors indicating different regions, and the Republic of Karabagh (in orange, far right). The neighboring countries are left uncolored. Within Armenia, the regions used in this study include Shirak in yellow, Lori in grass green, Gegharkunik in red (with Lake Sevan shown in contrasting color occupying the center of the state), Syunik in deep bright blue, and the capital region of Yerevan in magenta/purple.

neighbors, has remained separate and distinct from those of its neighbors. Finally, the adoption of Christianity in 301 AD and the concurrent development of its own unique alphabet further isolated Armenians from their non-Christian neighbors (3, 4, 5). The present day Karabagh was once the ancient land of Alwanians, who maintained close contacts with the Armenians, adopted the Armenian brand of Christianity in the fifth century, readily merged with the Armenians, and soon lost their separate identity and became Armenian (3).

Although these forces of geography, language, and religion kept Armenians apart from their neighbors, they also served as the very forces that gave Armenians a strong sense of religious and cultural identity, and kept them united throughout their history, a history rife with invasions and mass migrations. Prompted by countless foreign incursions, people within Armenia and Karabagh migrated extensively from one region to another, and mixed readily amongst themselves. In the 11th century, Seljuk Turkish invasions destroyed much of Armenia, forcing thousands of Armenians to take refuge in the only independent regions left, namely Karabagh and Syunik. At a later date, masses left Karabagh to move west to Syunik. During the 16th and 17th centuries, when Armenia was divided between the Ottoman Empire and Persia, a large number of Armenians were moved to – and settled in – Persia (present day Iran), and late in the 19th century 45,000 Armenians from the Turkish-occupied west and 90,000 from the Persian-occupied east moved north to Shirak and Gegharkunik, respectively. In more recent history, as a result of the genocide and deportation of the Armenians by the OttomanTurks in 1912–1915, a large number of Armenians were forced to leave their homeland and settled in Syria, Lebanon, Iran, Russia, and as far away as Europe and the United States, establishing the present day Armenian Diaspora, which spreads over 65 countries in five continents, and comprises approximately 5 million people. The Republics

of Armenia and Mountainous Karabagh combined, have, in turn, a population of <4 million, almost all Armenians (3).

In this study, we investigate genetic diversity of Armenian populations by performing HLA profiling on a large sample of individuals from the Republic of Armenia, the Republic of Mountainous Karabagh, plus populations from Iran, Lebanon, and the United States, the latter three as representative samples of the Armenian Diaspora.

Materials and methods

Population samples

HLA-A, HLA-B, and HLA-DRB1 typing were performed on 4279 unrelated Armenian bone marrow donors, all healthy, randomly selected Armenians from different provinces of the Republic of Armenia, the Republic of Karabagh, and three countries of the Diaspora – Iran, Lebanon (all in Beirut), and the United States (primarily in Los Angeles and Boston). Within this large general population, the genealogy of 2113 donors could be established based on the birthplace of the ancestors going back three generations. These donors were grouped into nine subpopulations. Five were from within the Republic of Armenia: Yerevan ($n = 445$), Lori ($n = 102$), Syunik ($n = 117$), Shirak ($n = 76$), and Gegharkunik ($n = 242$), and the others from outside the Republic of Armenia: Karabagh ($n = 445$), Iran ($n = 85$), United States ($n = 233$), and Lebanon ($n = 368$). Donors whose ancestors came from different provinces or whose ancestry could not be traced to the same region for three generations were included in the general population studies, but were excluded from the subpopulation comparison.

HLA typing

Genomic DNA was extracted from whole venous blood using QIAamp DNA Blood Mini Kits (27220 Turnberry Lane, Valencia, CA). HLA class I (HLA-A and HLA-B) and HLA class II (HLA-DRB1) alleles were typed using PCR–SSP low- and high-resolution kits (One Lambda, Canoga Park, CA and Genovision, West Chester, PA). For a group of 100 samples selected at random from the 2113 individuals with properly traced ancestry (total of all nine subpopulations), high-resolution typing was performed according to manufacturers' protocol (Genovision).

Statistical analysis

HLA-A, HLA-B and HLA-DRB1 allele frequencies were analyzed using the ARLEQUIN software (Institute of Ecology and Evolution, University of Bern, Switzerland; http://cmpg.unibe.ch/software/arlequin35), to obtain maximum-likelihood estimates (6, 7). Standard deviations were calculated from 10 bootstrap iterations. ARLEQUIN software was also used to calculate maximum-likelihood of three-locus haplotype frequencies from genotyping data through an expectation–maximization (EM) algorithm. Linkage disequilibrium (LD) analysis for pairs of loci and Hardy–Weinberg equilibrium (HWE) for each locus was performed using the same software.

The relationships between subpopulations were calculated using the principal coordinates analysis based on Reynolds distances (8) (number of permutations × 1000) using the NTSYS (Exeter software v.2.11, Setauket, NY 11733-2870, USA; http://www.exetersoftware.com/cat/ntsyspc/ntsyspc.html). High-resolution data for non-Armenian populations were obtained using the New Allele Frequency database: (http://www.allelefrequencies.net.) (9).

PYPOP 0.6.0 (10) was used to perform Ewens–Watterson homozygosity (EWH) test of neutrality (11–14). The observed (F_{obs}) and expected (F_{exp}) homozygosity (under neutral selection) were calculated respectively as sum of the squares of haplotype frequencies and through simulation, for the same sample size and number of unique haplotypes. The difference between F_{obs} and F_{exp}, divided by the square root of the variance of F_{exp} provides the normalized deviate of the homozygosity (F_{nd}) (11–15). A significantly negative F_{nd} value with the observed homozygosity level lower than the expected level suggests the presence of balancing selection, whereas a significantly positive value implies the presence of directional selection. The probability of the observed sample population to be obtained under neutrality is given by the P-value of F. As a two-tailed test of the null hypothesis of neutrality, the P values at either extreme of the distribution are considered significant: $P < 0.025$ or >0.975 at 0.05 significance level, while $P < 0.05$ or >0.95 at 0.10 significance level.

Results and discussion

HLA diversity of the Armenian population

In the total Armenian population ($n = 4279$), the most frequent alleles (frequency > 0.04) for HLA-A, HLA-B, and HLA-DRB1 loci were **02*, **03*, **24*, **01*, **11*, **26*, **32*, and **23* for the HLA-A locus; **35*, **51*, **44*, **49*, **18*, **38*, and **50* for the HLA-B locus; and **11*, **04*, **13*, **15*, **03*, **07*, **14*, and **01* for the DRB1 locus (Table 1). In agreement with our results, an earlier study of a smaller sample (16) reported that the most frequent DRB1 alleles in Armenians as *DRB1*11* and *DRB1*04*.

High-resolution typing of three loci in 100 unrelated donors yielded 23 HLA-A, 33 HLA-B, and 31 HLA-DRB1 alleles. The highest level of allelic heterogeneity was observed within *HLA-A*02* (four alleles: 0201, 0205, 0206, and 0208), *HLA-B*15* (four alleles: 1501, 1508, 1510, and 1517), *HLA-B*35* (four alleles: 3501, 3502, 3503, and 3508), *HLA-B*51* (four alleles: 5101, 5102, 5105, and 5107), *DRB1*04* (eight alleles: 0401, 0402, 0403, 0404, 0405, 0406, 0407, and 0408), *DRB1*11* (four alleles: 1101, 1102, 1103, and 1104), and *DRB1*13* (four alleles: 1301, 1302, 1303, and 1305). Other

Table 1 HLA-A, HLA-B, and HLA-DRB1 gene frequencies in a large Armenian population sample (low-resolution level, $n = 4279$)

HLA-A (P-value = 0.187)	Allele frequencies	HLA-B (P-value = 0.373)	Allele frequencies	HLA-DRB1 (P-value = 0.067)	Allele frequencies
*A*01*	0.113	*B*07*	0.036	*DRB1*01*	0.055
*A*02*	0.197	*B*08*	0.035	*DRB1*03*	0.074
*A*11*	0.075	*B*13*	0.026	*DRB1*04*	0.210
*A*23*	0.044	*B*14*	0.037	*DRB1*07*	0.072
*A*24*	0.150	*B*15*	0.029	*DRB1*08*	0.013
*A*25*	0.002	*B*1517*	0.011	*DRB1*09*	0.008
*A*26*	0.049	*B*18*	0.055	*DRB1*10*	0.020
*A*29*	0.026	*B*24*	0.000	*DRB1*11*	0.262
*A*03*	0.151	*B*27*	0.028	*DRB1*12*	0.012
*A*30*	0.037	*B*29*	0.000	*DRB1*13*	0.105
*A*31*	0.019	*B*35*	0.184	*DRB1*14*	0.058
*A*32*	0.046	*B*37*	0.009	*DRB1*15*	0.085
*A*33*	0.034	*B*38*	0.055	*DRB1*16*	0.027
*A*34*	0.000	*B*39*	0.010	—	—
*A*66*	0.006	*B*40*	0.014	—	—
*A*68*	0.038	*B*41*	0.025	—	—
*A*69*	0.013	*B*42*	0.000	—	—
*A*74*	0.001	*B*44*	0.082	—	—
—	—	*B*45*	0.001	—	—
—	—	*B*46*	0.000	—	—
—	—	*B*47*	0.002	—	—
—	—	*B*48*	0.000	—	—
—	—	*B*49*	0.064	—	—
—	—	*B*50*	0.049	—	—
—	—	*B*51*	0.143	—	—
—	—	*B*52*	0.034	—	—
—	—	*B*53*	0.004	—	—
—	—	*B*54*	0.000	—	—
—	—	*B*55*	0.027	—	—
—	—	*B*56*	0.001	—	—
—	—	*B*57*	0.016	—	—
—	—	*B*58*	0.018	—	—
—	—	*B*73*	0.002	—	—

HLA, human leukocyte antigen.

allele groups were characterized with a limited polymorphism (one or two alleles detected) in the Armenian population (Table 2).

We also calculated the frequencies of three-loci haplotypes for alleles at high resolution. We detected 165 HLA-A-B-DRB1 haplotypes. The most frequent ones are listed in Table 3. The complete results of all haplotypes are not presented because of space limitation, but could be made available upon request. The highest frequencies have the following three-loci haplotypes: *A*3301*, *B*1402*, and *DRB1*0102*, followed by *A*0301*, *B*3501*, and *DRB1*0101*, and *A*0101*, *B*4901*, and *DRB1*1104*. Our results are in agreement with previously published results that showed the *A*33*, *B*14*, *Cw8*, *DRB1*0102*, *DQA*0101*, and *DQB1*0501* haplotypes to occur with the highest frequency in Armenians (17). LD analysis using 1000 permutations for each of the three pairs of loci showed the presence of strong LD with P-values of <0.0001 (HLA-A/HLA-B pair), 0.007 (HLA-A/HLA-DRB1 pair), and <0.0001 (HLA-B/HLA-DRB1 pair). The impact of such pronounced LD was also supported by the presence of low allele frequencies for all the three loci. Although there were 200 gene copies in each locus from 100 donors, the numbers of observed alleles were found to be only 23 for HLA-A, 34 for HLA-B, and 31 for HLA-DRB1.

Analysis of the HLA diversity based on HWE

Testing HWE based on the high-resolution typing data of the entire population showed that individual HLA-A, HLA-B, and HLA-DRB1 phenotypes were in HWE, no significant influx of HLA alleles of heterologous origin into the gene pool. When HWE analysis was performed on nine subpopulations separately, the observed phenotype frequencies of the three loci did not deviate from the expected proportions for the majority of the subpopulations, suggesting the phenotypes being in HWE. However, for Karabakh and Lebanon subpopulations, HLA-A and HLA-DRB1 showed significant HW departures ($P < 0.05$). Although the reasons of this deviation for the Karabagh subpopulation are not clear to us, the deviation for the Lebanese Armenians probably reflects a certain degree of intermarriage between this subpopulation and the religiously

Table 2 HLA-A, HLA-B, and HLA-DRB1 alleles and their frequencies among Armenians – high resolution ($n = 100$)

HLA-A (P-value = 0.686)	Allele frequencies	HLA-B (P-value = 0.402)	Allele frequencies	HLA-DRB1 (P-value = 0.554)	Allele frequencies
*A*0101*	0.125	*B*0702*	0.030	*DRB1*0101*	0.045
*A*0201*	0.155	*B*0801*	0.010	*DRB1*0102*	0.040
*A*0205*	0.005	*B*1302*	0.035	*DRB1*0301*	0.035
*A*0206*	0.005	*B*1401*	0.005	*DRB1*0401*	0.045
*A*0208*	0.010	*B*1402*	0.060	*DRB1*0402*	0.045
*A*0301*	0.105	*B*1501*	0.015	*DRB1*0403*	0.035
*A*0302*	0.050	*B*1508*	0.005	*DRB1*0404*	0.045
*A*1101*	0.080	*B*1510*	0.005	*DRB1*0405*	0.010
*A*2301*	0.050	*B*1517*	0.020	*DRB1*0406*	0.005
*A*2402*	0.120	*B*1801*	0.045	*DRB1*0407*	0.020
*A*2403*	0.010	*B*2702*	0.015	*DRB1*0408*	0.010
*A*2601*	0.040	*B*2705*	0.005	*DRB1*0701*	0.050
*A*2901*	0.010	*B*3501*	0.095	*DRB1*0803*	0.015
*A*2902*	0.025	*B*3502*	0.040	*DRB1*0804*	0.005
*A*3001*	0.030	*B*3503*	0.035	*DRB1*0901*	0.010
*A*3101*	0.035	*B*3508*	0.025	*DRB1*1001*	0.025
*A*3201*	0.035	*B*3801*	0.040	*DRB1*1101*	0.080
*A*3301*	0.045	*B*3901*	0.015	*DRB1*1102*	0.005
*A*3303*	0.010	*B*4001*	0.005	*DRB1*1103*	0.010
*A*6601*	0.005	*B*4402*	0.050	*DRB1*1104*	0.115
*A*6801*	0.030	*B*4403*	0.010	*DRB1*1201*	0.005
*A*6901*	0.015	*B*4901*	0.095	*DRB1*1301*	0.040
*A*7403*	0.005	*B*5001*	0.045	*DRB1*1302*	0.015
—	—	*B*5101*	0.175	*DRB1*1303*	0.010
—	—	*B*5102*	0.010	*DRB1*1305*	0.015
—	—	*B*5105*	0.005	*DRB1*1401*	0.080
—	—	*B*5107*	0.005	*DRB1*1404*	0.005
—	—	*B*5201*	0.030	*DRB1*1501*	0.115
—	—	*B*5501*	0.020	*DRB1*1502*	0.035
—	—	*B*5701*	0.015	*DRB1*1601*	0.020
—	—	*B*5806*	0.015	*DRB1*1602*	0.010
—	—	*B*41XX*	0.015	—	—
—	—	*B*53XX*	0.005	—	—

HLA, human leukocyte antigen.

and culturally similar Lebanese. The only significant case of homogeneity was noticed in HLA-DRB1 locus of Gegharkunik subpopulation, undoubtedly a case of assortative mating in this region (18).

Relatedness between different Armenian subpopulations based on the Reynolds distances

Comparison of HLA frequencies between the nine subpopulations of Armenians studied is presented in Table 4. Relatedness between different Armenian subpopulations according to HLA-A, HLA-B, and HLA-DRB1 low-resolution allele frequency data and based on the Reynolds distances is presented through the principal coordinates plot (two-dimensional representation) in Figure 2. Using high-resolution data, a similar plot showing the genetic distances between the general Armenian population and several European populations based on HLA-A, HLA-B, and HLA-DRB1 allele frequencies is presented in Figure 3, and the genetic distances between the general Armenian population and some 'Mediterranean' populations using high-resolution HLA-DRB1 allele frequency data only is presented in Figure 4. The latter analysis was performed using HLA-DRB1 alone, because high-resolution data for HLA-A and HLA-B were unavailable for some of the populations we wanted to include because of their reported closeness to the Armenians. Results indicate that all nine Armenian subpopulations studied, including those geographically isolated from the Republics of Armenia and Karabagh, fall within the same cluster (Figure 2). Not surprising, given the fact that these communities were formed by Armenians from practically all of historic Armenia, and have only existed in the Diaspora for two to three generations at the most – not long enough to genetically differentiate significantly from the original population. Within the Armenian cluster, it was interesting to find the subpopulations of Karabagh and Syunik to be almost identical, probably reflecting the fact that present day Syunik is populated almost entirely by descendents of Armenians who migrated from Karabagh (see Introduction). When

Table 3 HLA-A, HLA-B, and HLA-DRB1 haplotype frequencies among Armenians ($n = 100$)

Haplotype	Frequency	SD′
*A*3301 B*1402 DRB1*0102*	0.030	0.013
*A*0301 B*3501 DRB1*0101*	0.025	0.011
*A*0101 B*4901 DRB1*1104*	0.020	0.012
*A*0201 B*0702 DRB1*1501*	0.015	0.009
*A*0201 B*5101 DRB1*1501*	0.015	0.011
*A*0302 B*4402 DRB1*0402*	0.015	0.010
*A*2402 B*3801 DRB1*1401*	0.015	0.009
*A*2402 B*5101 DRB1*0404*	0.015	0.010
*A*2601 B*3801 DRB1*0301*	0.015	0.009
*A*0101 B*1517 DRB1*1104*	0.010	0.007
*A*0101 B*3508 DRB1*1101*	0.010	0.007
*A*0101 B*5201 DRB1*1502*	0.010	0.008
*A*0201 B*5001 DRB1*0701*	0.010	0.009
*A*0201 B*5101 DRB1*1301*	0.010	0.007
*A*0301 B*3503 DRB1*1401*	0.010	0.007
*A*0301 B*5001 DRB1*0701*	0.010	0.008
*A*0301 B*5101 DRB1*0101*	0.010	0.009
*A*0302 B*1402 DRB1*1501*	0.010	0.007
*A*1101 B*1801 DRB1*1501*	0.010	0.007
*A*1101 B*4901 DRB1*0403*	0.010	0.008
*A*2301 B*0702 DRB1*1501*	0.010	0.007
*A*2301 B*4901 DRB1*1101*	0.010	0.009
*A*2402 B*4402 DRB1*1104*	0.010	0.008
*A*2402 B*5101 DRB1*1104*	0.010	0.009
*A*3101 B*3501 DRB1*1104*	0.010	0.009
*A*3101 B*5101 DRB1*1501*	0.010	0.008
*A*6801 B*5101 DRB1*0404*	0.010	0.008

HLA, human leukocyte antigen; SD′, standard deviation.

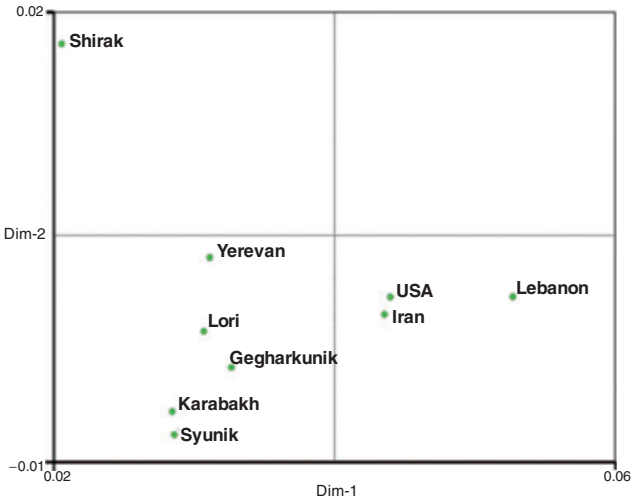

Figure 2 Principal coordinates plot (two-dimensional representation), based on the Reynolds distances, showing the relationship between different Armenian subpopulations according to low-resolution human leukocyte antigen (HLA)-A, HLA-B, and HLA-DRB1 allele frequency data.

viewed in terms of other ethnic groups, our analysis confirms previous findings regarding the relationship of Armenians with various European populations (Figure 3), and their relative closeness to Bulgarians (19, 20) and Italians (21, 22). In addition, based on HLA-DRB1 allele frequencies, our analysis also agrees with previous reports that place Armenians within the 'Mediterranean' substratum (19–23), and show them to be distinct from – but related to – Ashkenazi (24–26) and non-Ashkenazi (27) Jews, Turks (26, 28, 29), and Cretans (25, 26, 30).

EWH test of neutrality

Finally, we used EWH test of neutrality for comparative analysis of selective processes in Armenian subpopulations based on HLA alleles diversity. The EWH test compares the observed homozygosity with the homozygosity expected with no selection. In cases when the difference between observed and expected homozygosity (F_{nd}) is significant, presence of directional ($F_{nd} > 0$) or balancing ($F_{nd} < 0$) selection is likely.

All three loci have negative F_{nd} values for all subpopulations indicating an overall direction toward balancing selection, i.e. selection for sustaining high diversity of the HLA alleles in the populations (Table 5). It has been previously shown (15) that there is some evidence of balancing selection in DRB1 locus from 23 human populations across the globe that was attributed to selection for variability in the amino acid residues that form antigen-binding/presenting pocket. Our data confirm the previous observation and expand it to other HLA alleles. However, the difference between the observed and expected homozygosity was not statistically significant in all of the populations (Table 5).

The evidence of balancing selection in three loci combined (as observed by the two-tailed P-values for corresponding F_{nd} values) is most prominent for Karabagh and, then, Yerevan regions, whereas no deviation from neutral selection was noted in populations from Shirak, Iran, and Lori regions.

As we look at each locus separately, in case of HLA-A, four regions – Karabagh, Lebanon, Yerevan, and United States – showed strong evidence of balancing selection for HLA-A. On the other hand, in Ghegharkunik, Lori, Syunik, and Iran, HLA-A heterozygosity showed moderate, although insignificant, deviations from neutrality, indicative of some balancing selection pressures. However, in population from Shirak, the deviation from neutrality was found to be the lowest, and the Shirak HLA-A F_{nd} value was calculated to be significantly lower than all other F_{nd} values (t-test, $P < 0.01$), suggesting a relatively homogeneous subpopulation.

In case of HLA-B, the regions showing footprints of balancing selection were again Karabakh and Yerevan, along with Ghegharkunik, whereas Syunik and US diversity suggested marginal evidence of similar selective forces. Iran, Shirak, Lori, and Lebanon subpopulations, on the contrary, were with significantly low F_{nd} values compared with that for the other

Table 4 HLA-A, HLA-B, and HLA-DRB1 allele frequencies in nine Armenian subpopulations

Allele	Yerevan ($n = 445$)	Shirak ($n = 76$)	Ghegharkunik ($n = 242$)	Lori ($n = 102$)	Syunik ($n = 117$)	Iran ($n = 85$)	Karabakh ($n = 445$)	USA ($n = 233$)	Lebanon ($n = 368$)
HLA-A									
P-value	0.17	0.85	0.18	0.44	0.50	0.69	0.00	0.22	0.00
**01*	0.107	0.112	0.118	0.127	0.124	0.147	0.116	0.129	0.113
**02*	0.180	0.250	0.188	0.167	0.188	0.188	0.193	0.200	0.185
**03*	0.158	0.118	0.182	0.162	0.137	0.129	0.144	0.137	0.140
**11*	0.068	0.092	0.062	0.064	0.085	0.082	0.062	0.054	0.082
**23*	0.050	0.033	0.031	0.049	0.051	0.053	0.045	0.049	0.038
**24*	0.154	0.158	0.159	0.167	0.145	0.165	0.162	0.118	0.173
**25*	0.002	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000
**26*	0.054	0.033	0.056	0.049	0.038	0.041	0.048	0.062	0.043
**29*	0.037	0.033	0.039	0.020	0.030	0.029	0.022	0.017	0.018
**30*	0.042	0.039	0.037	0.039	0.038	0.018	0.042	0.036	0.042
**31*	0.012	0.026	0.021	0.025	0.017	0.029	0.026	0.021	0.015
**32*	0.043	0.046	0.023	0.044	0.056	0.035	0.046	0.060	0.061
**33*	0.025	0.013	0.031	0.044	0.051	0.041	0.042	0.054	0.038
**66*	0.008	0.013	0.008	0.015	0.009	0.000	0.001	0.004	0.008
**68*	0.047	0.033	0.037	0.020	0.013	0.024	0.030	0.041	0.031
**69*	0.009	0.000	0.006	0.010	0.013	0.018	0.019	0.015	0.014
**74*	0.003	0.000	0.000	0.000	0.004	0.000	0.002	0.002	0.000
HLA-B									
P-value	0.07	0.21	0.20	0.09	0.65	0.37	0.23	0.97	0.19
**07*	0.033	0.039	0.033	0.015	0.038	0.035	0.039	0.041	0.045
**08*	0.031	0.039	0.054	0.049	0.030	0.029	0.028	0.032	0.030
**13*	0.037	0.013	0.027	0.025	0.017	0.024	0.011	0.036	0.037
**14*	0.030	0.026	0.039	0.044	0.051	0.071	0.040	0.071	0.038
**15*	0.039	0.039	0.025	0.025	0.034	0.018	0.036	0.030	0.020
**16*	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
**18*	0.051	0.066	0.066	0.054	0.060	0.012	0.044	0.047	0.067
**27*	0.028	0.033	0.033	0.015	0.030	0.024	0.036	0.024	0.012
**35*	0.189	0.171	0.165	0.191	0.175	0.159	0.165	0.161	0.198
**37*	0.008	0.007	0.010	0.025	0.009	0.000	0.004	0.015	0.008
**38*	0.064	0.079	0.048	0.054	0.043	0.065	0.058	0.079	0.056
**39*	0.012	0.013	0.010	0.025	0.000	0.012	0.010	0.015	0.007
**40*	0.016	0.013	0.008	0.005	0.017	0.006	0.015	0.006	0.038
**41*	0.026	0.020	0.017	0.039	0.034	0.024	0.025	0.019	0.045
**44*	0.075	0.112	0.093	0.093	0.073	0.071	0.078	0.064	0.084
**45*	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
**46*	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
**47*	0.000	0.000	0.000	0.005	0.000	0.006	0.002	0.000	0.003
**48*	0.000	0.000	0.000	0.000	0.000	0.071	0.000	0.002	0.001
**49*	0.068	0.066	0.058	0.059	0.073	0.065	0.069	0.069	0.048
**50*	0.043	0.059	0.052	0.064	0.056	0.182	0.047	0.054	0.052
**51*	0.144	0.092	0.110	0.113	0.141	0.041	0.161	0.137	0.121
**52*	0.037	0.033	0.029	0.020	0.038	0.000	0.049	0.028	0.031
**53*	0.003	0.000	0.002	0.000	0.000	0.000	0.007	0.002	0.001
**54*	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000
**55*	0.022	0.013	0.050	0.039	0.030	0.024	0.025	0.030	0.015
**56*	0.002	0.007	0.000	0.000	0.000	0.000	0.000	0.002	0.003
**57*	0.016	0.000	0.012	0.015	0.017	0.018	0.011	0.013	0.016
**58*	0.017	0.026	0.041	0.015	0.013	0.012	0.025	0.006	0.007
**63*	0.006	0.026	0.014	0.015	0.021	0.029	0.013	0.015	0.014
**67*	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
**73*	0.000	0.000	0.002	0.000	0.000	0.006	0.001	0.000	0.001

Table 4 *Continued*

Allele	Yerevan ($n = 445$)	Shirak ($n = 76$)	Ghegharkunik ($n = 242$)	Lori ($n = 102$)	Syunik ($n = 117$)	Iran ($n = 85$)	Karabakh ($n = 445$)	USA ($n = 233$)	Lebanon ($n = 368$)
HLA-DRBI									
P-value	0.29	0.14	0.03	0.08	0.23	0.73	0.05	0.53	0.00
**01*	0.054	0.053	0.045	0.025	0.094	0.082	0.065	0.058	0.068
**03*	0.062	0.072	0.091	0.069	0.077	0.059	0.064	0.071	0.079
**04*	0.196	0.217	0.238	0.206	0.188	0.235	0.197	0.236	0.185
**07*	0.064	0.066	0.062	0.064	0.090	0.059	0.070	0.067	0.084
**08*	0.012	0.013	0.014	0.015	0.004	0.006	0.011	0.013	0.007
**09*	0.005	0.013	0.012	0.005	0.004	0.006	0.015	0.009	0.004
**10*	0.022	0.039	0.010	0.034	0.021	0.006	0.031	0.015	0.020
**11*	0.287	0.257	0.289	0.294	0.226	0.241	0.235	0.227	0.306
**12*	0.014	0.033	0.008	0.015	0.013	0.006	0.009	0.015	0.022
**13*	0.109	0.086	0.083	0.137	0.077	0.118	0.089	0.127	0.087
**14*	0.064	0.033	0.060	0.049	0.077	0.071	0.063	0.034	0.053
**15*	0.093	0.059	0.068	0.074	0.098	0.106	0.122	0.107	0.076
**16*	0.019	0.059	0.019	0.015	0.030	0.006	0.029	0.021	0.010

HLA, human leukocyte antigen.

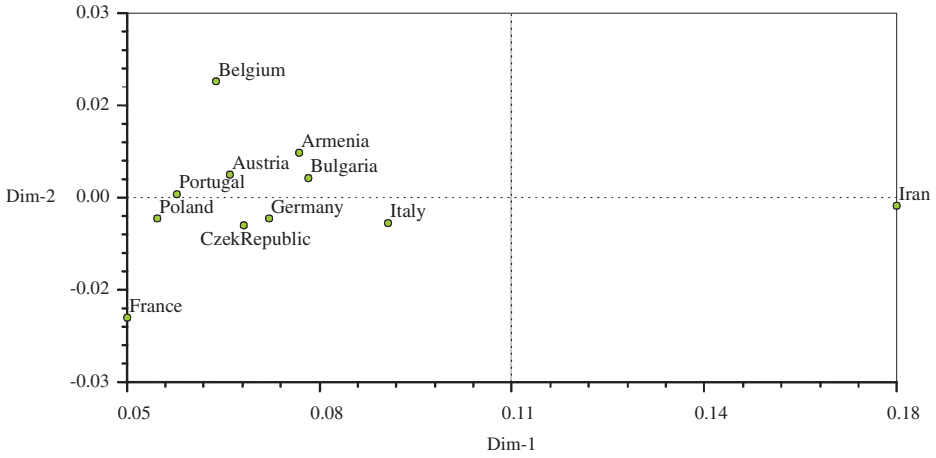

Figure 3 Principal coordinates plot (two-dimensional representation), based on the Reynolds distances, showing the relationship between Armenians and other populations according to high-resolution human leukocyte antigen (HLA)-A, HLA-B, and HLA-DRB1 allele frequency data.

five regions (*t*-test, $P < 0.01$), again suggesting no deviation from neutrality for HLA-B.

For HLA-DRB1 locus, significant signature of balancing selection was found only for Karabagh (marginal evidence was noted for Yerevan and Syunik). Balancing selection on the Karabakh HLA loci maintaining high allelic diversity is consistent with the origin of the Karabakh population as merge of ethnically distinct Armenians and Caucasian Albanians in ancient history and, then, because of massive migration into the region of Armenians from different regions during the invasion of Seljuk Turks. Similarly, the Yerevan population is also rich in diversity as a result of high level of migration into this capital from across Armenia, especially in recent times. Such high diversity of stable population might be reflected in increased allelic diversity of Yerevan HLA loci undergoing balancing selection. In contrast, the striking low levels of allelic diversity values and their closeness with the expected values under neutrality for Shirak HLA loci possibly reflect the founders effect (relatively recent growth expansion from a small population) as this region has been populated primarily by a single, relatively small migration wave (see *Introduction*).

Concluding remarks

This study provides the first comprehensive investigation of HLA–allele distribution in a subset of nine Armenian subpopulations and of the HLA–allele and haplotype frequencies at high resolution combined with established and novel types of population genetic analysis. Regarding frequency of different HLA alleles and their combination in the Armenian population in general, our study confirms on a large sample, and

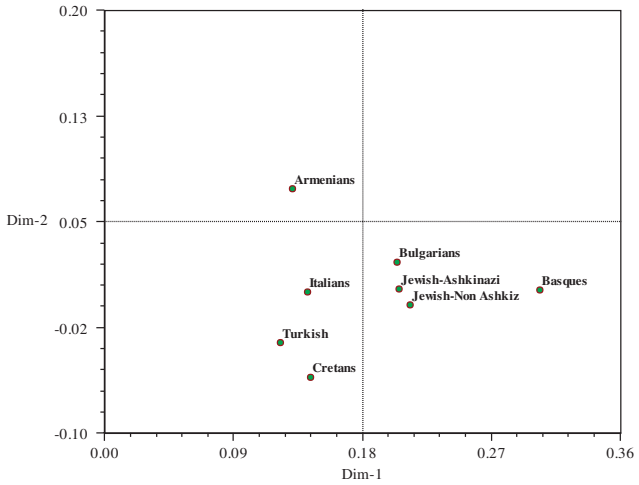

Figure 4 Principal coordinates plot (two-dimensional representation), based on the Reynolds distances, showing the relationship between Armenians and other populations according to high-resolution human leukocyte antigen (HLA)-DRB1 allele frequency data.

expands the conclusions made previously by other smaller low-resolution studies (16, 31, 32). The study firmly shows that Armenian people are relatively homogeneous genetically. On the basis of HWE data, we show the lack of any significant admixture of genes of heterologous origin, indicating the preservation of the gene pool. This is not surprising, given the relative isolation of the nation historically even during a long-term dominance by other ethnic groups, and even after the population fractured and dispersed into a widespread Diaspora.

Most importantly, our study provides evidence for detectable genetic differences between certain subpopulations of Armenians. Specifically, the EWH test shows, at one extreme, the high diversity in one subpopulation (i.e. Karabagh) as evidenced by strong balancing selection on HLA loci, and at the other extreme, the presence of a relatively homogeneous subpopulation (i.e. Shirak). The genetic data appears to reflect the history of intra-Armenian migration and, possibly, could be used for further in-depth studies to refine the history of migration.

Table 5 Results of Ewens–Watterson homozygosity test of neutrality in Armenian subpopulations

Region	Locus	F_{obs}	F_{exp}	Variance of F_{exp}	F_{nd}	P	P-value (two-tailed)
Yerevan [Y] ($n = 445$)	HLA-A	0.111	0.222	0.007	−1.319	0.010	<0.05
	HLA-B	0.085	0.150	0.003	−1.227	0.024	<0.05
	HLA-DRB1	0.158	0.286	0.012	−1.177	0.045	<0.10
Shirak [Sh] ($n = 76$)	HLA-A	0.133	0.175	0.003	−0.769	0.202	ns
	HLA-B	0.073	0.089	0.001	−0.716	0.228	ns
	HLA-DRB1	0.144	0.191	0.004	−0.772	0.206	ns
Ghegharkunik [G] ($n = 242$)	HLA-A	0.122	0.208	0.006	−1.141	0.045	<0.10
	HLA-B	0.075	0.128	0.002	−1.284	0.016	<0.05
	HLA-DRB1	0.171	0.256	0.009	−0.898	0.149	ns
Lori [L] ($n = 102$)	HLA-A	0.113	0.178	0.003	−1.117	0.050	ns
	HLA-B	0.078	0.106	0.001	−0.930	0.118	ns
	HLA-DRB1	0.168	0.208	0.005	−0.579	0.315	ns
Syunik [S] ($n = 117$)	HLA-A	0.110	0.172	0.003	−1.105	0.053	ns
	HLA-B	0.081	0.125	0.001	−1.158	0.036	<0.10
	HLA-DRB1	0.130	0.216	0.005	−1.162	0.043	<0.10
Iran [I] ($n = 85$)	HLA-A	0.119	0.181	0.003	−1.061	0.066	ns
	HLA-B	0.086	0.098	0.001	−0.480	0.365	ns
	HLA-DRB1	0.155	0.237	0.006	−1.036	0.086	ns
Karabakh [K] ($n = 445$)	HLA-A	0.114	0.236	0.008	−1.366	0.008	<0.05
	HLA-B	0.083	0.150	0.003	−1.267	0.017	<0.05
	HLA-DRB1	0.136	0.286	0.012	−1.379	0.009	<0.05
USA [U] ($n = 233$)	HLA-A	0.109	0.206	0.006	−1.308	0.014	<0.05
	HLA-B	0.079	0.126	0.002	−1.155	0.041	<0.10
	HLA-DRB1	0.151	0.254	0.009	−1.105	0.064	ns
Lebanon [L] ($n = 368$)	HLA-A	0.116	0.242	0.008	−1.380	0.007	<0.05
	HLA-B	0.085	0.121	0.002	−0.904	0.126	ns
	HLA-DRB1	0.163	0.277	0.011	−1.086	0.073	ns

HLA, human leukocyte antigen; ns, not significant.

Finally and most importantly, we expect that the current data would serve as a valuable reference for organ transplantation and future HLA disease association studies in Armenians in general and in the specific subpopulations, as well as of Armenians in conjunction with other ethnic groups.

Conflict of interest

The authors have declared no conflicting interests.

References

1. Terasaki PI. *Histocompatibility Testing. Copenhagen: Munksgaard, 1970*.
2. Hewsen R. The Geography of Armenia. In: Hovanissian, RG. ed. The Armenian People from ancient to modern times, Vol. 1. New York: *St. Martin's Press*, 1997, 1–18.
3. Russell J. The formation of the Armenian nation. In: Hovanissian, RG. ed. The Armenian People from ancient to modern times, Vol. 1, New York: *St. Martin's Press*, 1997, 19–35.
4. Garsoian N. The emergence of Armenia. In: Hovanissian, RG. ed. The Armenian People from ancient to modern times, Vol. 1, New York: *St. Martin's Press*, 1997, 36–62.
5. Garsoian N. The Arsakuni Dynasty. In: Hovanissian, RG. ed. The Armenian People from ancient to modern times, Vol. 1, New York: *St. Martin's Press*, 1997, 63–94.
6. Excoffier L, Laval G, Schneider S. Arlequin (Version 3.0): and integrated software package for population genetics data analysis. *Evol Bioinform Online* 2005: **1**: 47–50.
7. Sanchez-Mazas A. HLA data analysis in anthropology: basis, theory and practice. *16th European Histocompatibility Conference*, *Strasbourg, March 19–22, 2002*.
8. Reynolds J, Weir BS, Cockerham CC. Estimation for coancestry coefficient: basis for a short term genetic distance. *Genetics* 1983: **105**: 767–79.
9. Middleton D, Menchaca L, Rood H, Komerofsky R. New allele frequency database: http://www.allelefrequencies.net. *Tissue Antigens* 2003: **61**: 403–7.
10. Lancaster A, Nelson MP, Single RM, Meyer D, Thomson G. PyPop: a software framework for population genomics: analyzing large-scale multi-locus genotype data. In: Altman RB, Dunker AK, Hunter L, Jung TA, Klien TE, eds. Pacific Symposium on Biocomputing, Vol. 8. Singapore: *World Scientific*, 2003, 514–25.
11. Ewens WJ. The sampling theory of selectively neutral alleles. *Theor Popul Biol* 1972: **3**: 87–112.
12. Watterson GA. The homozygosity test of neutrality. *Genetics* 1978: **88**: 405–17.
13. Slatkin M. An exact test for neutrality based on the Ewens sampling distribution. *Genet Res* 1994: **64**: 71–4.
14. Slatkin M. A correction to the exact test based on the Ewens sampling distribution. *Genet Res* 1996: **68**: 259–60.
15. Salamon H, Klitz W, Easteal S et al. Evolution of HLA class II molecules: allelic and amino acid site variability across populations. *Genetics* 1999: **152**: 393–400.
16. Alexeev LP, Patranji G. HLA in eight ethnic groups from the former USSR. *Proceedings of the Eleventh International Histocompatibility Workshop and Conference, Yokohoma, Japan, vol. 1, 1991, Oxford Science Publications. Vol. 1. 1991, 666–7*.
17. Imanishi T, Akaza T, Kimura A, Tokunaga K, Gojobori T. Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: Tsuji K, Aizawa M, Sasazuki T, eds. *HLA 1991. Oxford: Oxford University Press, 1992, 1065–220*.
18. Cavalli-Sforza LL, Menozzi P, Piazza A. *The History and Geography of Human Genes. Princeton: Princeton University Press, 1994*.
19. Ivanova M, Rozemuller E, Tyufekchiev N, Michailova A, Tilanus A, Naumova E. HLA polymorphism in Bulgarians defined by high resolution typing methods in comparison with other populations. *Tissue Antigens* 2002: **60**: 496–504.
20. Ivanova M, Spassova P, Michailova A, Naumova E. Distributions of HLA class I alleles and haplotypes in Bulgarians – contribution to understanding the origin of the population. *Tissue Antigens* 2001: **57**: 208–15.
21. Testi M, Mariani M. HLA-A, -B, -Cw, -DPB1, -DQA1, -DQB1 and DRB1 allele frequencies in a population from Rome, Italy. *Hum Immunol* 2004: **65**: 987–91.
22. Sanchez-Velasco P, Leyva-Cobian F. The HLA class I and class II allele frequencies studied at the DNA level in the Svanetian population (Upper Caucasus) and their relationships to Western European populations. *Tissue Antigens* 2001: **58**: 223–33.
23. Arnaiz-Villena A, Karin M, Bendikuze N et al. HLA alleles and haplotypes in the Turkish population: relatedness to Kurds, Armenians and other Mediterraneans. *Tissue Antigens* 2001: **57**: 308–17.
24. Martinez Laso J, Gazit E, Gomez-Casado E et al. HLA-DQA1, -DQB1 and – DRB1 Alleles in Ashkenazi Jews from Israel. *Hum Immunol* 2004: **65**: 972.
25. Moscoso J, Crawford MH, Vicario JL et al. HLA genes of Aleutian Islanders living between Alaska (USA) and Kamchaka (Russia) suggest a possible southern Siberia origin. *Mol Immunol* 2008: **45**: 1018–102.
26. Arnaiz-Villena A, Reguera R, Ferri A et al. The peopling of Madeira Archipelago (Portugal) according to HLA genes. *Int J Immunogenet* 2008: **36**: 9–14.
27. Martinez Laso J, Gazit E, Gomez-Casado E et al. HLA-DQA, -DQB and – DRB1 Alleles in a non-Ashkenazi population from Israel. *Hum Immunol* 2004: **65**: 972–3.
28. Saruhan-Direskeneli C, Ugar FA. HLA-A, -B, -Cw, -DQA1, -DQB1 and DRB1 allele frequencies in a population from Turkey. *Hum Immunol* 2004: **65**: 1183–5.
29. Musabak U, Ilikci Sagkan R, Misiroglu H et al. Regional distribution of HLA antigens in Turkey. *Tissue Antigens* 2009: **73**: 473. (P152)
30. Arnaiz-Villena A, Iliakis P, Gonzalez-Hevilla Met al. The origin of Cretan population as determined by characterization of HLA alleles. *Tissue Antigens* 1999: **53**: 213–26.
31. Bonne-Tamir B, Bodmer JG, Young D, Ashbel S. HLA polymorphism in Israel. 8. The Armenian community in Jerusalem. *Tissue Antigens* 1978: **11**: 230–4.
32. Nersisian VM, Martirosian IG. Frequency of gene and haplotype occurrence of HLA system A and B loci in Armenians. *Probl Gematol Pereliv Krovi* 1982: **27**: 31–3 (in Russian).

