ELSEVIER

Sequence-Based Analysis of the HLA-DRB1 Polymorphism in Metalsa Berber and Chaouya Arabic-Speaking Groups From Morocco

Khadija Oumhani, Angelica Canossi, Daniela Piancatelli, Marilena Di Rocco, Tiziana Del Beato, Gabriella Liberatore, Anna Aureli, Abd Elaziz Ben Jouad, Rajae El Aouad, Domenico Adorno, and Carlo Umberto Casciani

ABSTRACT: To examine the genetic diversity in Morocco, the polymorphism at the HLA-DRB1 locus was investigated in two populations: the Metalsa group consisting of Berbers from north Morocco (who speak the Tarifit language and live in the Nador area), and the Chaouya group who are Arabic-speaking people from west Morocco (Atlantic coast) living in the Settat area. The DRB1 alleles of 197 healthy unrelated individuals were identified by direct DNA sequencing of exon 2 using fluorescently-labeled primers. A total of 28 and 29 alleles at DRB1 locus were identified in the Metalsa and Chaouya groups, respectively. The most frequent alleles in the Metalsa group are DRB1*03011 (20.2%), DRB1*0701 (12.12%), and DRB1*1302 (11.11%). In the Chaouya group, DRB1*0701 (16.33%), DRB1*15011 (12.76%), and DRB1*03011 (11.73%) are most common. Each population exhibits some specific variants and some uncommon alleles. The frequency of the DRB1*03011 allele differs significantly between the two populations ($p = 0.0311$). The DRB1 frequency distributions in the two groups suggest the effects of balancing selection. The interpopulation analysis highlighted a strong relatedness, based on genetic distances, between the two Moroccan groups and the other north Africans (the Moroccans from El Jadida area, Moroccan Souss Berbers, Algerians, and Tunisians), and to a lesser extent with the Iberians, French, and Ethiopians. *Human Immunology 63, 129–138 (2002).*

KEYWORDS: HLA-DRB1 polymorphism; sequence-based typing; Morocco; Metalsa group; Chaouya group

ABBREVIATIONS

HLA	human leukocyte antigens
SBT	sequence-based typing
ME	Metalsa
CH	Chaouya
SSP-PCR	sequence-specific primers–polymerase chain reaction
SGD	standard genetic distance

INTRODUCTION

The geographical integrity of Morocco, separated from Black Africa by the Sahara Desert, and bordered by the Mediterranean Sea to the north and the Atlantic Ocean to the west, explains the unicity of its population. Its geographical location has influenced the biologic features of its population [1]. Moroccan populations have preserved within them some of the genetic features of other pop-

From the Istituto C.N.R. Tipizzazione Tissutale e Problemi della Dialisi (A.C., D.P., M.D.R., T.D.B., G.L, A.A., D.A., C.U.C.), L'Aquila, Italy; Laboratoire d'Immunologie (K.O., R.E.A), Institut National D'Hygiene, Rabat, Morocco; and Laboratoire de Biochimie (A.E.B.J.), Faculté des Sciences, Rabat, Morocco.

Address reprint requests to: Dr. Domenico Adorno, Istituto C.N.R. Tipizzazione Tissutale e Problemi della Dialisi, P.le Collemaggio, 67100 L'Aquila, Italy; Tel: +39 (0862) 414275; Fax: +39 (0862) 410758; E-mail: ittpda01@area.ba.cnr.it.

Received April 9, 2001; revised October 11, 2001; accepted October 31, 2001.

Human Immunology 63, 129–138 (2002)

0198-8859/02/$–see front matter
PII S0198-8859(01)00370-6

ulations to whom they are geographically and historically related, the Europeans and Asiatics bordering the Mediterranean area.

The people of Morocco speak either Arabic or the Berber language, and are, in some places, bilingual. The Berber population has an ancient origin (from the Neolithic period), and represents the autochthonous basis of the Moroccan population. It has been successively driven back by the exclusively urban civilizations that have conquered Africa [2, 3]. Genealogists and historians of Africa have ascribed an oriental origin to the Berber population [1], which was brought to Morocco by one or more of the immigrations from Asia during the second millennium BC, and probably also during the following centuries. Morocco has experienced colonization by Phoenicians, Carthaginians, and Romans. The Berber elements were, on the other hand, only weakly affected by the Latin conquest [4–6]. Between the 7^{th} and 9^{th} centuries AD the greatest Arabian invasion occurred that determined a considerable intermingling of Arab and Berber populations. On the other hand, the conquests of Spain and Sicily by Arabs brought a stream of Berbers into these countries. After the Catholic conquest, between the 13^{th} and 15^{th} centuries AD, new waves of immigrants (Muslims and Jews) intermingled with the Berber tribes of Morocco. Between the 16^{th} and 19^{th} centuries AD, various immigrations occurred, especially those of the European traders [6], Spaniards, and Portuguese. Finally, the French and Spanish colonized Morocco in the 19^{th} and 20^{th} centuries.

The historic perspective attests to the great heterogeneity of the Moroccans. The present Moroccan population comprises different ethnic groups (Arabic speaking or Berber speaking). Although in the past the Moroccan groups mingled, the present population is classified on the basis of language. In fact, as elsewhere described [7], the most strongly represented is the Arabic-speaking group, which lives mainly in northern Morocco and in the south, on the Atlantic Coast. Saharians and Jews are also present and are Arabic-speaking in Morocco.

Moroccan Berbers are represented by three main communities: those from the north, living in the Rif mountains (Tarifit language); those from the High Plateau, Middle Atlas, and Moulouya regions (Tamazight language); and those from southwestern Morocco (the Chleuh group), including the Anti-Atlas, western High Atlas, and Souss Valley regions (Tachelhit language).

One of the most important genetic markers for characterizing populations and their relationships is the human leukocyte antigen (HLA) system, because of its extensive polymorphisms. Until now, only a few studies have been available for HLA analysis in Moroccans. Available data include one study of a Souss Berber population sample from the Agadir city area [8], and another population sample from El Jadida area [9]. We present here, for the first time, a high-resolution study of polymorphisms at the HLA-DRB1 locus, using sequence-based typing (SBT), in the following two Moroccan populations: the Metalsa (ME), Berbers from the north speaking the Tarifit language and living in the Nador area; and the Chaouya (CH), an Arabic-speaking group from the west and living in the Settat area. The allelic differences and similarities between these two groups were evaluated and compared with other African groups and European populations from the Mediterranean area, to provide genetic information with which to clarify the historical origins of these populations.

MATERIALS AND METHODS

Population Samples

Blood samples were obtained from 197 healthy unrelated volunteers: 99 from the ME group and 98 from the CH group. Both parents and the grandfathers of all subjects belonged to the same ethnic group as the subject.

DNA Preparation

EDTA-blood samples were lysed, and pellets were stored frozen at -20 °C until DNA extraction. DNA samples were extracted by a salting-out procedure.

HLA-DRB1 Sequence-Based Typing

DRB1 SBT was carried out after preliminary sequence-specific primer–polymerase chain reaction (SSP-PCR) screening of the major DRB1 allele groups (DRB1G1, DRB1G2, DRB1G3/5/6/8, DRB1G4, DRB1G7, DRB1G9, and DRB1G10), using primers designed at the 12^{th} International Histocompatibility Workshop and described elsewhere [10]. The group-specific primers complementary to the 5′-region of the gene were 21M13-tailed, and the 3′ primer was M13-tailed. PCR products were sequenced using one set of fluorescently-labeled primers for each direction, complementary to the 21M13 and M13 tails (Applied Biosystems, Foster City, CA, USA). Electrophoresis was performed on an ABI 377 DNA sequencer, on a 5% Long Ranger denaturing gel (FMC BioProducts, Rockland, ME, USA). Further specific PCR and sequencing reactions were performed, using selective primers for codon 86, to distinguish between some ambiguous DRB1 allele heterozygotes. The sequence data were processed using the HLA Matchtools and MT Navigator software (Perkin-Elmer Applied Biosystems, Norwalk, CT, USA), which detects the heterozygous positions within each electropherogram, and assesses the typing on the basis of an alignment of the processed sequence with an updated HLA sequence library.

TABLE 1 Populations analyzed in this study (high-resolution DRB1 typing)

Reference number	Ethnic group	Number of individuals
35	Algerians (Algiers)	106
36	Arab Emirates Bedouin	108
37	Asian Indians	54
38	Cretans	135
39	Ethiopians (Amhara)	98
39	Ethiopians (Oromo)	83
40	French	130
41	Greeks (Northern Greece)	181
42	Italians (Abruzzo)	325
43	Kenians (Nairobi)	299
44	Lebanese NS, KZ, YH	257
45	Macedonians	80
Present study	Moroccans (Metalsa, Berbers)	99
Present study	Moroccans (Chaouya, Arabic-speaking)	98
46	Morocco Jews (Non-Ashkenazi)	120
9	Moroccans (EI Jadida area, Arabic speaking)	98
8	Moroccans (Souss, Berbers)	98
47	Sardinians (central-south Sardinia)	80
48	Senegalese (Eastern Senegal)	198
49	Spanish	120
50	Spanish Basques (Province of Vizcaya)	103
51	Tunisians	101
52	Turks (Istanbul)	250

Statistical Analysis

The analysis of molecular polymorphisms at the HLA-DRB1 locus within populations was performed using the Arlequin v.1.1 population genetics software (http://anthropologie.unige.ch/arlequin; University of Geneva, Geneva, Switzerland) [11]. HLA-DRB1 allele frequency data in both Moroccan groups were calculated. Results were compared with those of previous studies of other north African and Mediterranean populations, using a χ^2 analysis with Yates' correction. A list of the populations compared in the present study is given in Table 1.[35–47] The level of significance was set at $p = 0.05$. The Bonferroni correction was applied in the comparison of DRB1 frequencies among populations [12]. To test the hypothesis that the observed diploid genotypes were the product of a random union of gametes, the Hardy-Weinberg equilibrium was calculated [13]. An Ewens-Watterson homozygosity statistic F [14, 15] was used to examine the presence of selective forces influencing the HLA allele frequency distribution at the DRB1 locus. The observed allele frequency distributions were compared with those expected under the neutral model (mutation and random genetic drift). Homozygosity values lower than the neutral value by the Ewens-Watterson model indicate the action of balancing selection, and values higher than the neutral value reflect the action of directional selection. As described [16], a balancing selection tends to maintain multiple polymorphic alleles in a population at appreciable frequencies and a directional selection determines an allele distribution with one common allele, significantly more frequent than the neutral expectation The significance was calculated by using the Slatkin exact test [17, 18], which compares the probabilities of the random samples with that of the observed sample. The normalized deviate homozygosity statistic (F_{nd}) was calculated to compare the homozygosity F statistics across populations [19]. Negative values of F_{nd} suggest a tendency to balancing selection, whereas positive values suggest directional selection. The null distribution of F is generated by simulating random neutral samples having the same number of genes and the same number of haplotypes using the algorithm of Stewart [20]. Gene diversity (h = expected heterozygosity) [21] and maximal heterozygosity were analyzed (calculated according to the formula [k–1]/k, where k is the number of alleles at a locus and all the alleles have the same frequency of 1/k). Standard genetic distances (SGD) were computed between populations, and a phylogenetic tree (dendrogram) was constructed with the allelic frequencies using the neighbor-joining method and the software Dispan, with the programs GNKDST and TREEVIEW [22, 23]. This method constructs a tree by successive

TABLE 2 Comparison of DRB1 allele frequencies between the Moroccan Metalsa and Chaouya populations

DRB1*	Metalsa af%[2], $n = 99$	Chaouya af%, $n = 98$	p Value[1]
0101	0,00	1,02	
01021	4,04	2,55	
15011	9,60	12,76	
15021	0,51	0,00	
1503	0,51	1,53	
1601	0,51	1,02	
03011	20,20	11,73	**0.031**
03021	0,00	1,02	
0401	0,00	1,02	
0402	6,57	7,14	
0403	2,02	3,57	
0404	1,01	1,02	
04051	6,06	4,08	
0406	4,04	2,55	
11011	4,55	2,04	
1102	1,52	5,10	
1103	0,00	1,02	
11041	1,52	1,02	
1201	0,51	0,00	
1301	3,54	3,06	
1302	11,11	5,61	
13031	1,01	0,51	
13032	1,01	4,59	
1304	0,51	0,00	
1401	1,01	1,53	
1406	0,51	1,02	
0701	12,12	16,33	
0801	0,51	0,51	
08032	0,00	0,51	
08041	0,51	0,00	
0806	0,51	1,02	
09012	2,02	0,51	
1001	2,53	4,59	

[1] P probability values (p values) were calculated using a Chi-square test with Yates' correction. The level of significance was set at $p = 0.05$ after correction.

[2] Frequency is defined as the percentage of the number of times in which an allele is present in 198 and 196 alleles in total, respectively. Alleles for presumed homozygous samples are counted twice.

clustering of lineages, setting branch lengths as the lineages join. The software was also used to perform bootstrap tests for phylogenetic tree.

RESULTS

We have analyzed the DRB1 allele polymorphisms in two Moroccan populations, one Berber and one Arabic speaking. We analyzed the DRB1 gene polymorphisms using the most recently updated HLA-DRB1 library (273 DRB1 alleles) [24]. The DRB1 allele frequencies (af) of the two populations are reported in Table 2. A total of 28 and 29 alleles at the DRB1 locus were found in the ME and CH groups, respectively. In the ME group, the most frequent alleles (present in at least 10% of the population) are DRB1*03011 (af = 20.2%), DRB1*0701 (af = 12.12%), and DRB1*1302 (af = 11.11%). In the CH group, DRB1*0701 (af = 16.33%), DRB1*15011 (af = 12.76%), and DRB1*03011 (af = 11.73%) are most frequent. Each population exhibits some specific variants (ME: DRB1*15021, 1201, 1304, 08041; CH: DRB1*0101, 03021, 04011, 1103, 08032) and some uncommon alleles (DRB1*1503, 1406, 0806; ME: DRB1*1304; CH: DRB1*03021). The frequency of the DRB1*03011 allele differs significantly between the two populations (ME, af = 20.2%; CH, af = 11.7%, $p = 0.031$; Figure 1). The expected and observed genotype frequency values for the DRB1 locus do not differ significantly, and both the populations are in Hardy-Weinberg equilibrium (ME, $p = 0.232$; CH, $p = 0.986$). By comparing the observed homozygosity F values with those expected under the Ewens-Watterson neutral model, a trend in the direction of balancing selection in both populations was observed (ME: observed F = 0.0942, expected F = 0.1029; CH: observed F = 0.0789, expected F = 0.1017), although these were not statistically significant (Slatkin's exact p value). The normalized deviate F ($F_{nd}A$) is negative for both populations (ME $F_{nd}A = -0.364$; CH $F_{nd}A = -0.762$), suggesting that balancing selection is acting on these populations, perhaps together with other evolutionary forces. Furthermore, gene diversity was 0.9104 for the ME group and 0.9259 for the CH group; these values approached the maximum heterozygosity levels expected for 28 alleles (0.964) and 29 alleles (0.966). This is in keeping with the relatively low frequencies observed for most alleles, especially in the CH group, where the most frequent allele, DRB1*0701, occurs at a frequency of 16.33%.

Comparison of the Most Significant DRB1 Subtypes in the ME and CH Groups With Other Populations

The most significant DRB1 subtypes in the ME and CH groups were compared with other African populations (Algerians, Tunisians, and Oromo Ethiopians), Europeans (Spanish, Spanish-Basques, Sardinians, French, and Italians), and neighboring populations (Moroccans from Souss, Moroccans from El Jadida city, and Moroccan Jews).

*DRB1*01 subtyping.* The DRB1*0101 allele, only detected in the CH group, occurs at a lower frequency than in the European populations ($p < 0.045$ compared with Spain, France, and Spanish Basques). On the other hand, the DRB1*01021 allele, observed in both groups (ME, 4.04%; CH, 2.55%), is uncommon in European populations (except in Spain, Sardinia, and Greece), but is found in other north African groups, and among Leba-

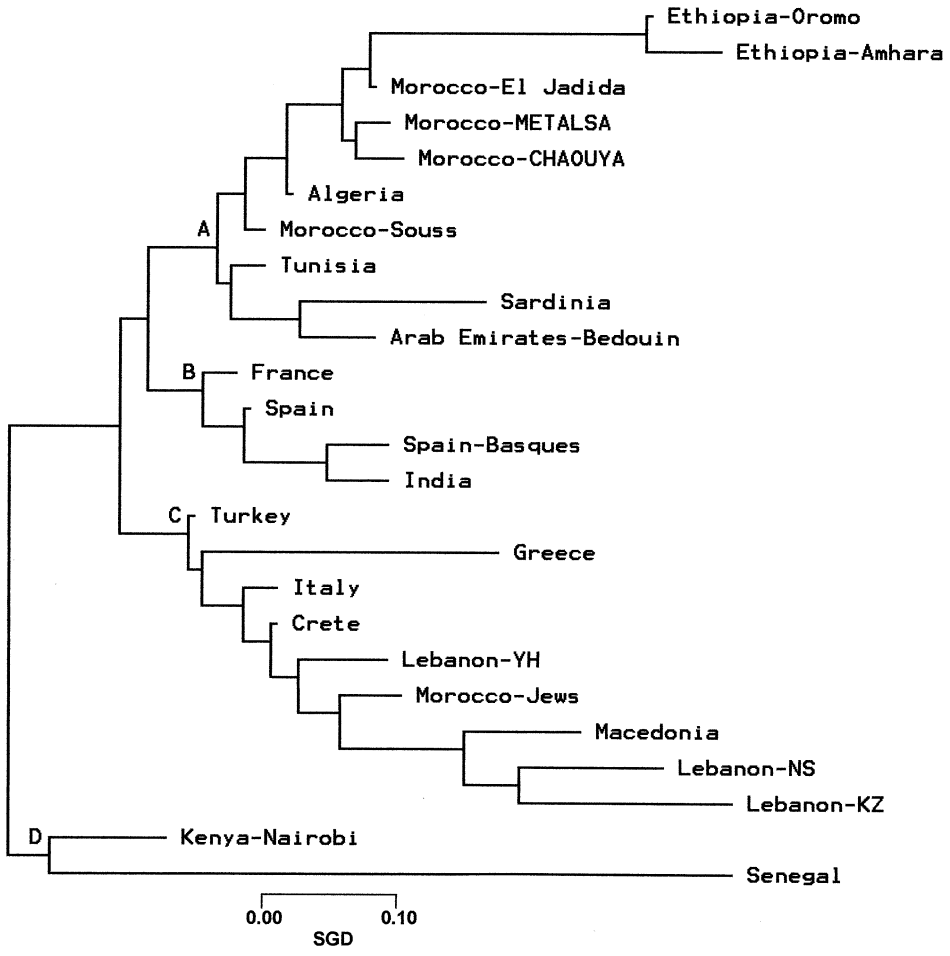

FIGURE 1 Neighbor-joining dendrogram illustrating the relatedness of the Metalsa and Chaouya Moroccans, and other populations. Standard genetic distances (SGD) were calculated using only data from high-resolution DRB1 typing. Data for other populations were taken from the references cited in Table 1. Bootstrap values from 1000 replicates are illustrated.

nese (4.8%) and Turkish (1.2%) populations. Nevertheless, the frequency of DRB1*01021 in CH is significantly reduced compared with Algerians ($p = 0.029$).

*DRB1*15,16 subtyping.* The DRB1*15011 allele is one of the most frequent alleles in the CH (12.76%) and ME (9.6%) populations. It is the most frequent subtype in the DR2 allele group, occurring at a significantly higher frequency in the ME and CH populations than in Italy or Sardinia, or among Moroccan Jews and Oromo Ethiopians ($p < 0.018$). The DRB1*15021 allele is rare or absent (ME, 0.51%; CH, 0%), as it is among the Moroccan Souss and other African populations (Tunisians, Moroccans from El Jadida, Senegalese, and Oromo and Amhara Ethiopians), except Moroccan Jews ($p < 0.023$ compared with ME or CH). The DRB1*1503 allele, present at low frequencies in both the Moroccan populations (ME, 0.51%; CH, 1.53%), was not observed among the Moroccan Souss or the Moroccan Jews, but has been reported in the Moroccans from El Jadida and other African groups (Ethiopian, Kenyan, Cameroon, and Ivory Coast populations) [25].

*DRB1*03 subtyping.* The DRB1*03011 allele is one of the most frequent alleles in the ME (20.2%) and CH (11.7%) groups. Its frequency in the ME group is significantly higher than in Spanish, French, Italian, and Moroccan Jews ($p < 0.025$). The DRB1*03021 allele, absent in the ME group and uncommon in the CH (1.02%) groups, is a typical allele in central (Kenya and Senegal) and South African populations, but is also present in other Moroccan populations (Souss and from El Jadida area), in Algeria, and in Tunisia.

*DRB1*04 subtyping.* A high frequency (ME, 19.70%; CH, 19.38%) and wide variability in DR4 subtypes was observed. Of these, the DRB1*0402 allele has the highest frequency (ME, 6.57%; CH, 7.14%), occurring sig-

nificantly more often in both groups than it does in France, Italy, or Oromo population ($p < 0.023$). The DRB1*04051 allele (ME, 6.6%; CH, 4.08%) occurs more frequently in ME than in some European populations (Italy, Spanish Basques, $p < 0.042$). The DRB1*0406 allele is a relatively frequent variant in both of the studied populations (ME, 4.04%; CH, 2.55%), in contrast with its low incidence in European populations (France, Italy). Its occurrence in the ME group is also more frequent than in Tunisia ($p = 0.013$).

*DRB1*11 subtyping.* The DRB1*1102 allele is the most frequent variant of the DR11 allele in the CH group (5.1%), which is also the case only among the Spanish Basques. Its occurrence is significantly higher than in Moroccan Jews ($p = 0.043$). The DRB1*11011 allele has the highest frequency of the DR11 subtypes in the ME group (4.55%). In the CH group the frequency of this allele is significantly lower than in Tunisia ($p = 0.008$). Conversely, the DRB1*11041 allele is significantly underepresented in both populations (ME, 1.52%; CH, 1.02%) compared with Italians and the Moroccan Jews ($p < 0.002$).

*DRB1*13 subtyping.* The DRB1*1302 allele has the highest frequency of the DR13 subtypes in both groups (ME, 11.11%; CH, 5.61%). A significantly higher frequency was found in the ME group compared with Tunisian ($p = 0.014$) and European populations (Spanish Basques, Spanish, Italian, and Sardinian) ($p < 0.01$). The DRB1*1301 allele has a relatively low incidence (ME, 3.54%; CH, 3.06), whereas it represents the most common subtype among Europeans. We observed an increased frequency in the DRB1*13032 subtype in the CH group compared with the Moroccan Jews ($p = 0.029$). There is a predominance of the DRB1*13032 allele (4.59%) compared with the DRB1*13031 allele (0.51%) in the CH population, whereas these frequencies were equal in the ME population. It is noteworthy that DRB1*1304, a very uncommon allele, occurs in the ME population, although it does not occur in other African populations, except in Senegal.

*DRB1*14 subtyping.* DRB1*1401 occurs at a low frequency (ME, 1.01%; CH, 1.53%). The DRB1*1406 allele, an uncommon allele, was detected in ME (0.51%) and CH (1.02%) populations. This allele has never been reported in Morocco or other African countries, and has only been detected previously in France and Turkey.

*DRB1*07 subtyping.* The DRB1*0701 allele is one of the most frequent alleles in both populations (ME, 12.12%; CH, 16.33%). It occurs in both ME and CH at a significantly lower frequency than in Ethiopians ($p = 0.023$).

*DRB1*08 subtyping.* The variants of this subgroup (DRB1*0801, 08032, 0804, and 0806) occur at low frequencies, as reported for other north African populations. We stress the presence in both groups (ME, 0.51%; CH, 1.02%) of DRB1*0806, an uncommon allele previously found in the population from El Jadida in Morocco, and first described in Algeria [26] of the north Africa, and subsequently in Kenya, Senegal, and France.

*DRB1*09012 allele.* The frequency of this allele is higher in the ME group (2.02%) than in other north African populations or among Italian caucasoids ($p =$ NS).

*DRB1*1001 allele.* This allele occurs at a frequency similar to those of other north African populations (ME, 2.53%; CH, 4.59%), and is more common in the CH group than in Italy ($p = 0.003$).

Analysis of the DRB1-based genetic distances reveals a high degree of genetic relatedness between the ME and CH ethnic groups and other north African populations (Algerian, Moroccan from El Jadida, Moroccan Souss Berber, and Tunisian) (Figure 1; Table 3). A specific HLA-DRB1 neighbor-joining tree is presented in Figure 1, and reveals the clustering of the two Moroccan samples, together with the other Moroccan populations (from El Jadida area, Souss), populations of the north African countries (Algeria, Tunisia), Ethiopian and Sardinian populations (in spite of their large genetic distance), and the Arab Emirates on one side, because outgroups among the more homogeneous groups are closer in dendrograms. The populations of France, Spain, Spanish Basques, and Asians are grouped on the other side of one cluster. On the other side of the tree are grouped the populations of Turkey (relatively close to both the ME and CH groups), Greece, Italy, Crete, Lebanon, Macedonia, and the Moroccan Jews. The populations of Kenya and Senegal (the most distant groups) occur in a distinct cluster. The Metalsa Berber group demonstrated the closest genetic relationship with the Moroccans from El Jadida (SGD = 1.62), followed by the Chaouya, Algerian, Tunisian, and Moroccan Souss populations. The Chaouya group, on the other hand, is most closely related to the Algerians (SGD = 3.19), the ME, and the other Moroccan populations described here (the groups from El Jadida and Souss).

DISCUSSION

The HLA system can be used as a genetic marker for anthropologic studies because of its extensive polymor-

TABLE 3 Standard genetic distances between the Metalsa (Berber) and Chaouya (Arabic-speaking) Moroccans and other populations ($\times 10^2$)

	DRB1 Metalsa	Genetic distances (10^2)		DRB1 Chaouya	Genetic distances (10^2)
1.	Moroccans (EI Jadida)	1.62	1.	Algerians	3.19
2.	Moroccans (Chaouya)	5.49	2.	Moroccans (Metalsa)	5.49
3.	Algerians	6.59	3.	Moroccans (EI Jadida)	7.84
4.	Tunisians	7.98	4.	Moroccans (Souss)	11.92
5.	Moroccans (Souss)	9.31	5.	Tunisians	14.26
6.	Ethiopians (Oromo)	18.37	6.	Spanish Basques	15.19
7.	Turks	24.02	7.	Spaniards	20.59
8.	Spanish Basques	24.39	8.	Ethiopians (Oromo)	20.94
9.	French	25.48	9.	Asian Indians	23.59
10.	Arab Emirated Bedouins	25.89	10.	Turks	25.72
11.	Spaniards	28.65	11.	Arab Emirated Bedouins	28.39
12.	Ethiopians (Amhara)	29.54	12.	French	28.60
13.	Kenians (Nairobi)	30.73	13.	Ethiopians (Amhara)	33.51
14.	Sardinians	31.04	14.	Cretans	34.44
15.	Cretans	36.96	15.	Lebanese (YH)	35.61
16.	Asian Indians	38.04	16.	Italians	36.41
17.	Lebanese (YH)	42.90	17.	Moroccans (Jews)	39.01
18.	Italians	44.60	18.	Kenians (Nairobi)	44.90
19.	Moroccans (Jews)	51.40	19.	Sardinians	45.24
20.	Greeks	61.55	20.	Greeks	63.85
21.	Macedonians	62.29	21.	Macedonians	64.81
22.	Lebanese (KZ)	75.43	22.	Lebanese (NS)	73.20
23.	Lebanese (NS)	81.82	23.	Lebanese (KZ)	90.32
24.	Senegalese	85.28	24.	Senegalese	98.91

phisms. DRB1 data are very informative and discriminatory in the comparison of populations because of this extensive variability, and the availability of high-resolution typing information for this locus. The present high-resolution HLA analysis describes, for the first time, DRB1 polymorphisms in two groups from Morocco, distinct both geographically and ethnically (the Berber and Arabic-speaking populations). The Chaouya population, which has never before been studied at the HLA level, is composed totally of Arabic-speaking people, the most highly represented linguistic group in Morocco. These results suggest that the two populations have high levels of genetic diversity, a feature of African populations. In fact, a total of 28 and 29 DRB1 alleles were observed in the Metalsa and Chaouya populations, respectively. Each population exhibited some specific variants and uncommon alleles. In particular, the DRB1 polymorphisms suggest a tendency to balancing selection (heterozygote advantage), perhaps acting together with other evolutionary forces. This trend is more evident in the Chaouya group, where the most frequent alleles (DRB1*0701, DRB1*15011, and DRB1*03011) occur at frequencies below 16.3%. Comparing the allele frequencies in the two studied groups, we found significant differences in the DRB1*03011 allele and some specific allele variants. These differences could be indicative of the distinct origins of the Berber and Arabic-speaking groups, whether or not these ethnic groups cluster phylogenetically. The present study confirms, with genetic distance data, the strong relatedness of the two Moroccan groups, first with other north African populations (the Moroccans from El Jadida, Algerians, Tunisians, and Souss Berbers), and to a lesser degree with the Iberians, French, and Ethiopians. These data are supported by the occurrence, in the ME and CH groups, of the DRB1*01021, 03021, 0806, and 1001 alleles, which are in common with other north African populations; the DRB1*0806, 1406 alleles, which are in common with the French population; the DRB1*01021 allele, which is in common with the Spanish population; and the high frequency of the DRB1*1102 allele, as found among the Basques. The Metalsa group is closely related to the El Jadida population and is less closely related to the Souss population, Berbers from around Agadir who speak a different language (Tachelhit). As for Chaouya group, it was previously reported [27] that the original meaning of the term "Chaouya" described the area inhabited by herdsmen. Since the 15th century this word has become an ethnic name. "Chaouya" is also the name of other ethnic groups (a mixture of Berber- and Arabic-speaking people) inhabiting different sites in north Africa, including Algeria. These groups share a common Berber origin. The Chaouya of western Morocco derive from the Zenata (8th century Berber-speaking people) and Houwara (13th

century Berbers who were bilingual) populations, who came from Africa (especially Libya) and spread throughout north Africa. Despite their common Berber origins, some Chaouya populations of north Africa remained Berber-speaking, others became bilingual, and others became totally Arabic-speaking, as did the Moroccan Chaouya. The genetic affinity reported here between the Chaouya group and the Algerians confirms the historically documented common primitive origin of these populations. It is noteworthy that the Oromo and Amhara groups cluster together with the Berbers and other north African groups, consisting of Berbers, Egyptians (shared features include an absence of the DRB1*15021 allele and a high frequency of the DRB1*1302 allele, in common with the Metalsa group). This clustering may reflect these groups common ethnic origin from the ancient (pre-Neolithic) Hamites (Hamite-speaking people) [28] on the Mediterranean and Red Sea coasts. Bedouins of the Arab Emirates and Sardinians are grouped together in the same main cluster with the other north Africans (cluster A), though at a greater genetic distance. This suggests the importance of the Arabian genetic influence (from the 7th century AD) in north Africa. The Spanish Basques are genetically related to the Berbers and the Moroccan groups, and cluster with the population of Spain (cluster B). As previously reported, the Iberian gene pool derives from north Africa; the Basque, Iberian, Libyan, and Berber languages are all related, and belong to the Na-Dene Caucasian group of languages, widely spoken in Eurasia and north Africa (8000–6000 BC) [29, 30]. Some paleolithic north Africans (Hamites) reached the Iberian Peninsula, whereas others may have settled in Sardinia, Crete, and Etruria (Italy) [31]. The correlations found in the dendrogram analysis of the DRB1 locus confirm the relatedness of the Moroccan–north African populations with these populations. Moroccans are more closely related to Africans and western Europeans (clusters A and B) than to eastern Mediterranean people (Greeks, Cretans, Lebanese, and Macedonians) (cluster C). However, we found evidence of some genetic similarity between the Moroccan groups and the Turkish population that clusters phylogenetically with Italy, the eastern Mediterranean, Macedonia, and Lebanon. This suggests that Turkey has been an important genetic influence in these populations, due probably to their invasion of north Africa until Algeria in the 16th century. Among the Metalsa Berbers the genetic DRB1-based affinity with Turkey is even stronger than with Spain. A weak Asian influence in north Africa could be inferred from some genetic features. These include the relatively frequent DRB1*0406 allele, also present in several Asian countries, which reaches frequencies of 4%–5% [32, 33], and the intermediate genetic distance with Asian Indians. The genetic correlation is weak with other populations (Romans and Phoenicians) because the influence was very ancient, and has not been maintained over the centuries, as confirmed by historical data [4–6]. A significant correlation exists between genetic distances, measured in terms of HLA sequence data, and geographic distances. The present study, based on an analysis of the distribution of HLA allele frequencies, specifically of the DRB1 locus, confirms the strong correlation between genetic distances and geographic-historical factors. As previously reported [34], the effects of selective forces influencing allele frequency distributions at the DRB1 locus are very small, and this makes this locus an informative genetic marker for this population analysis. This study is also interesting in its clarification of the differences and the similarities between two Moroccan populations: the Metalsa Berbers (Berber-speaking until now) and the Chaouya (Berbers by origin, but subsequently Arabic speaking). These groups have experienced different immigration events over the centuries, and have consequently evolved different lifestyles that caused them to become Arabic speaking or to remain Berber speaking. This study will be a useful background for clinical analyses (of HLA-associated diseases and transplantation, for example) in the Moroccan population, and for the correlation of genetic traits with anthropologic data in the Mediterranean area.

ACKNOWLEDGMENTS

K.O. provided and reviewed historical data on Moroccan populations, and worked on HLA typing and collected blood samples. A.C. and D.P. wrote the manuscript, analyzed sequence data, and performed statistical calculations. M.D.R., T.D.B., G.L., and A.A. performed HLA typing. A.E.B.J. reviewed historical data. R.E.A. identified populations to be tested and reviewed the manuscript. D.A. and C.U.C. provided founding and reviewed the manuscript. This study was supported in part by CNR/CNCPRST Joint Research Project 1997-'99 and by C.N.R., Targeted Project "Biotecnologie."

REFERENCES

1. Kossovitch N: Anthropologie et groupes sanguins des populations du Maroc. Paris, Masson et Cie, 1953.
2. Murdok GP: Africa, Its Peoples and Their Cultural History. New York: McGrawHill, 1959.
3. Odinot P: Tribus berbères. Revue de géographie marocaine 211, 1928.
4. Brett M, Fentress E: The Berbers. Oxford: Blackwell Publishers, 1997.
5. Gaid M: Les Berbères dans l'historie. Tome 1. Alger: Editions Mimouni, 1990.
6. Julien CA: Historie de l'Afrique du Nord. Paris: Payot, 1961.
7. Grimes BF (eds): Ethnologue, 13th edition. Summer In-

stitute of Linguistics Inc., 1996. Web site: www.sil.org/ethnologue/countries/Moro.html#MEY.

8. Izaabel H, Garchon HJ, Caillat-Zucman S, Beaurain G, Akhayat O, Bach JF, Sanchez-Mazas A: HLA class II DNA polymorphisms in a Moroccan population from the Souss, Agadir area. Tissue Antigens 51:106, 1998.
9. Gomez-Casado E, Del Moral P, Martinez-Laso J, Garcia-Gomez A, Allende L, Silvera-Redondo C, Longas J, Gonzales-Hevilla M, Kandil M, Zamora J, Arnaiz-Villena A: HLA genes in Arabic-speaking Moroccans: close relatedness to Berbers and Iberians. Tissue Antigens 55:239, 2000.
10. Fernandez-Vina M, Bignon JD: Primers and oligonucleotide probes (SSOP) used for DNA typing of HLA class II alleles. In Charron D (ed.): Genetic Diversity of HLA. Functional and Medical Implications. Paris: EDK, 1997.
11. Schneider S, Kueffer JM, Roessli D, Excoffier L: Arlequin ver.1.1: A Software for Population Genetic Data Analysis. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva, 1997.
12. Stanton AG: Primer of Biostatistics. New York: MacGraw-Hill, Inc., 1992.
13. Guo S, Thomson E: Performing the exact test of Hardy Weinberg proportion for multiple alleles. Biometrics 48: 361, 1992.
14. Ewens WJ: The sampling theory of selectively neutral alleles. Theor Pop Biol 3:87, 1972.
15. Watterson GA: The homozygosity test of neutrality. Genetics 88: 405, 1978.
16. Apple RJ, Erlich HA: HLA class II genes: structure and diversity. In Browning M, McMichael A (eds): HLA and MHC Genes and Molecule and Function. Oxford: BIOS Scientific Publishers Ltd, 1996.
17. Slatkin M: An exact test for neutrality based on the Ewens sampling distribution. Genet Res 64:71, 1994.
18. Slatkin M: A correction to the exact test based on the Ewens sampling distribution. Genet Res 68:259, 1996.
19. Anderson R: Some stochastic models in population genetics. MSc. Thesis. Victoria, Australia: Monash University, 1979.
20. Stewart FM: Computer algorithm for obtaining a random set of allele frequencies for a locus in an equilibrium population. Genetics 86:482, 1977.
21. Nei M: Molecular Evolutionary Genetics. New York: Columbia University Press, 1987.
22. Nei M: Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321, 1973.
23. Nei M, Tajjma F, Tateno Y: Accuracy of estimated phylogenetic trees from molecular data. J Mol Evol 19:153, 1983.
24. Marsh SGE, Bodmer JG, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich IIA, Hansen JA, Mach B, Mayr WR, Parham P, Petersdorf EW, Sasazuki T, Schreuder GMTh, Strominger JL, Svejgaard A, Terasaki PI: Nomenclature for factors of the HLA system, 2000. Tissue Antigens 57:236, 2001.
25. Ginsberg C, Quakyi I, Mshana R, Ngongan J, Johnson A: HLA-DR profiles in Cameroon and Ivory Coast. Hum Immunol 49(Suppl.1):48, 1996.
26. Benmamar D, Martinez-Laso J, Varela P, Bekhoucha F, Morales P, Arnaiz-Villena A: Evolutionary relationships of HLA DR-8 alleles and description of a new subtype (DRB1*0806) in the Algerian population. Hum Immunol 36:172, 1993.
27. Documents et renseignements publiés sous les auspices de la résidence générale – Villes et Tribus du Maroc. Casablanca et les chaouia. Tome I. Paris: Ernest Leroux, 1915.
28. McEvedy C, Jones R: Atlas of World Population History. Harmondsworth: Penguin Books, 1978.
29. Arnaiz-Villena A, Martinez-Laso J, Alonso-Garcia J: Iberia: population genetics, anthropology and linguistics. Hum Biol 71:725, 1999.
30. Ruhlen M: The origin of the Na-Dene. Proc Natl Acad Sci USA 95:13994, 1998.
31. Arnaiz-Villena A, Martinez-Laso J, Gomez-Casado E, Diaz-Campos N, Santos P, Martinho A, Breda-Coimbra H: Relatedness among Basques, Portuguese, Spaniards and Algerians studied by HLA allelic frequencies and haplotypes. Immunogenetics 47:37, 1997.
32. Lee J: Chinese normal: 204. In Gjertson DW, Terasaki PI (eds): HLA 1998. Lenexa, Kansas: American Society for Histocompatibility and Immunogenetics, 1998.
33. Geng L, Imanishi T, Azaka T, Tokunaga K, Zhu D, Mizuki N, Xu S, Geng Z, Gojobori T, Tsuji K, Inoko H: Determination of HLA class II alleles by genotyping in a Manchu population in the northern part of China and its relationship with Han and Japanese populations. Tissue Antigens 46:111, 1995.
34. Tiercy JM, Sanchez-Mazas A, Excoffier L, Shi-Isaac X, Jeannet M, Mach B, Langaney A: HLA-DR polymorphism in a Senegalese Mandenka population: DNA oligotyping and population genetics of DRB1 specificities. Am J Hum Genet 51:592, 1992.
35. Arnaiz-Villena A, Benmamar D, Alvarez M, Diaz-Campos N, Varela P, Gomez-Casado E, Martinez-Laso J: HLA allele and haplotype frequencies in Algerians: relatedness to Spaniard and Basques. Hum Immunol 43:259, 1995.
36. Lester S, Dowse T, Frossard P, McCluskey J: Arab Emirate Bedouin Normal. In Terasaki PI, Gjertson DW (eds.): HLA 1997. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1997.
37. Singal DP, D'Souza M, McEnemy L, Joseph S: Asian Indian Normal. In Terasaki PI, Gjertson DW (eds.): HLA 1997. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1997.
38. Arnaiz-Villena A, Iliakis P, Gonzalez-Hevilla M, Longas J, Gomez-Casado E, Sfyridaki K, Trapaga J, Silvera-Redondo C, Matsouka C, Martinez-Laso J: The origin of

Cretan populations as determined by characterization of HLA alleles. Tissue Antigens 53:213, 1999.

39. Fort M, de Stefano GF, Cambon-Thomsen A, Giraldo-Alvarez P, Dugoujon JM, Ohayon E, Scano G, Abbal M: HLA class II allele and haplotype frequencies in Ethiopian Amhara and Oromo populations. Tissue Antigens 51:327, 1998.
40. Gebuhrer L, Betuel H: Caucasian French normal. In Gjertson DW, Terasaki PI (eds.): HLA 1998. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1998.
41. Lazidou P, Adam K, Polymenidis Z: Caucasian Greek normal. In Terasaki PI, Gjertson DW (eds.): HLA 1997. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1997.
42. Adorno D: Caucasian Italian normal. In Gjertson DW, Terasaki PI (eds.): HLA 1998. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1998.
43. Lakshmi G, Montero R: Nairobi normal. In Terasaki PI, Gjertson DW (eds.): HLA 1997. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1997.
44. Bias W, Gazit E: 12th International Histocompatibility Workshop: anthropology SWAS regional report. In Charron D (ed.): Genetic Diversity of HLA: Functional and Medical Implications. Paris: EDK, 1997.
45. Hristova-Dimceva A, Verduijn W, Schipper RF, Schreuder GM: HLA-DRB and -DQB1 polymorphisms in the Macedonian population. Tissue Antigens 55:53, 2000.
46. Amar A, Roitberg-Tambur A, Witt CS, Safirman C, Kwon OJ, Brautbar C: Jewish Non-Ashkenazi normal. In Terasaki PI, Gjertson DW (eds.): HLA 1997. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1997.
47. Contu L, Carcassi C: Sardinian normal. In Terasaki PI, Gjertson DW (eds.): HLA 1997. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1997.
48. Tiercy JM, Sanchez-Mazas A, Excoffier L, Langaney A: Senegalese normal. In Terasaki PI, Gjertson DW (eds.): HLA 1997. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1997.
49. Muro M: Caucasian Spanish normal. In Gjertson DW, Terasaki PI (eds.): HLA 1998. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1998.
50. Sanchez-Velasco P, Escribano de Diego J, Martinez-de-Pancorbo M, Leyva-Cobian F: Caucasian Spanish-Basque normal (province of Vizcaya, Arratia Cluster). In Gjertson DW, Terasaki PI (eds.): HLA 1998. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1998.
51. Hmida S, Gauthier A, Dridi A, Quillivic F, Genetet B, Boukef K, Semana G: HLA class II gene polymorphism in Tunisians. Tissue Antigens 45:63, 1995.
52. Saruhan-Direskeneli G, Uyar FA, Bakar S, Eraksoy M: Molecular analysis of HLA-DRB1, -DQA1, and -DQB1 polymorphism in Turkey. Tissue Antigens 55:171, 2000.