
HLA Genes in the Chuvashian Population from European Russia: Admixture of Central European and Mediterranean Populations

ANTONIO ARNAIZ-VILLENA,^{1*} JORGE MARTINEZ-LASO,^{1*} JUAN MOSCOSO,¹
GREGORY LIVSHITS,² JORGE ZAMORA,¹ EDUARDO GOMEZ-CASADO,¹
CARLOS SILVERA-REDONDO,¹ KRISTIN MELVIN,³ AND MICHAEL H. CRAWFORD³

Abstract HLA alleles have been determined for the first time in individuals from the Chuvashian population by DNA typing and sequencing. *HLA-A*, *-B*, *-DR*, and *-DQ* allele frequencies and extended haplotypes have also been determined, and the results compared to those for Central Europeans, Siberians and other Asians, Caucasians, Middle Easterners, and Mediterranean peoples. Genetic distances, neighbor-joining dendrograms, and correspondence analysis have been performed. Present-day Chuvash speak an Altaic-Turkic language and are genetically related to Caucasians (Georgians), Mediterraneans, and Middle Easterners, and not only to Central or Northern Europeans; Chuvash contain little indications of Central Asian-Altaic gene flow. Thus, present-day Chuvash who speak an Altaic-Turkic language are probably more closely related to ancient Mesopotamian-Hittites and northern European populations than to central Asia-Altaic people.

The HLA system is the most polymorphic genetic system described in humans and consists of several closely linked loci. The strength of the multiallelic HLA system to single out individuals for paternity testing has been shown to be greater than the additive strength of many other polymorphic enzymes and blood group systems. The discovery of new loci and the presently available DNA typing and sequencing of new alleles have dramatically increased the variety of HLA allelism (Marsh et al. 2001). Certain alleles are frequent only in specific populations (i.e., *A36*, *A43* in Negroids); the strong linkage disequilibrium between

¹Department of Microbiology I-Immunology, Faculty of Medicine, Universidad Complutense, 28040 Madrid, Spain.

²Department of Anatomy, Sackler Faculty of Medicine, Tel-Aviv, Israel.

³Laboratory of Biological Anthropology, Department of Anthropology, University of Kansas, Lawrence, Kansas 66045.

*The contribution by A. Arnaiz-Villena and J. Martinez-Laso is equal and the order of authorship is arbitrary.

Human Biology, June 2003, v. 75, no. 3, pp. 375–392.

Copyright © 2003 Wayne State University Press, Detroit, Michigan 48201-1309

KEY WORDS: HLA, CHUVASH, CENTRAL EUROPE, BULGARS, BULGARIANS, MACEDONIANS, MEDITERRANEANS, RUSSIANS, SIBERIANS, NA-DENE

HLA neighboring loci demonstrates that certain combinations of contiguous alleles (HLA haplotypes) show a characteristic frequency or are distinctive in certain living populations (Imanishi et al. 1992a). Thus, the analysis of the HLA allelism represents a valuable tool to trace migration of ancient human populations and also the ethnical composition of the present day populations (Tokunaga et al. 1996) together with other currently used genetic markers. This system is a useful tool for studying the origins of determined groups like the Chuvashian people (see below).

The Chuvash are a European ethnic minority residing in western Russia, who constitute the majority of the population of the Republic of Chuvashia (67.8% are Chuvash and 26.7% are of Russian origin). The Chuvash Republic, with its capital Shupashkar (in Russian, Cheboksary), is located in the middle flow of the Volga River, in the center of the European part of the Russian Federation. It is bordered by the Mordova Republic, the Mari Republic, the Republic of Tatarstan, and the Nizhni-Novgorod and Ulianovsk districts of the Russian Federation. The Chuvash language belongs to the Volga-Bulgar subgroup of the Turkic branch of the Altaic language family (Figure 1).

The Chuvash are believed to have originated from Turkic-Altaic Bulgar tribes who migrated in the 4th century A.D. from Central Asia together with the Huns to the western region of the Volga River. It is not unlikely that they represent an amalgamation of Bulgars with the Finno-Ugric tribes already living in the area (Livshits et al. 1999a, 1999b). The ancestors of the Chuvash were also found as seminomadic tribes of ancient Bulgars who lived in the North Caucasus steppes in the 5th to 8th centuries. In the 7th to 8th centuries a portion of the Bulgars left for the Balkans, while another subdivision moved to the mid-Volga region and made up the ethnic base of the Chuvash and Kazan Tatars. In the 10th century, a Volga-Bulgar state was founded. In 1236 the Volga-Bulgar Empire was destroyed by the Mongol Golden Horde. After the death of Genghis Khan (A.D. 1227), the Bulgars came under the control of Kazan Khanate, which was founded in 1437. During the 15th and 16th centuries, the Chuvash formed a single nation, the base of which was a rural Bulgarian population that had not adopted Islam. Side-by-side with the Tatars and Maris, the Chuvash fought fiercely against Russian invasion, until their defeat and annexation by Russia in 1551. The Chuvash Autonomous Region was formed on June 24, 1920. On April 21, 1925, it was reformed into the Chuvash Autonomous Soviet Socialist Republic, and in 1992 into the Chuvash Republic.

This article provides for the first time an HLA gene profile and the characteristic HLA haplotypes among the Chuvash. We examine the relative contribution of ancient northern and central Europeans and the influence of western and eastern Mediterranean populations, Caucasians, Asians, and Siberians to the present-day Chuvash gene pool, including the possible relationship between this population and the Bulgarians with whom they share an historical common origin.

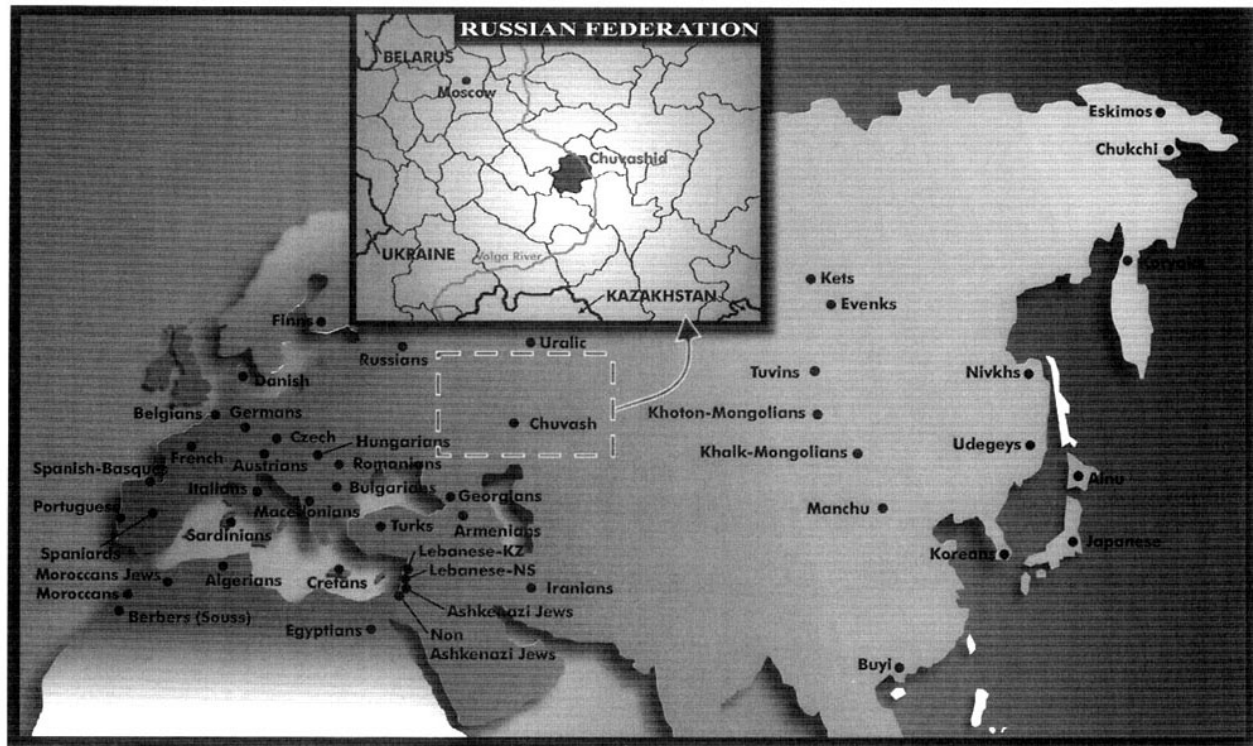


Figure 1. Map showing the location of the Chuvashian Republic and populations tested in the present study. Green and orange colors indicate the north European and Mediterranean genetic backgrounds, respectively, both present in the Chuvash population.

Materials and Methods

Population Samples. Eighty-two unrelated individuals from the Chuvash ethnic group, living in the Republic of Chuvashia (Turkic speakers) and with at least two recorded generations of Chuvash ascent, were used for HLA genotyping and phylogenetic calculations. The original sample was collected by Livshits and coworkers (Livshits et al. 1999a, 1999b) as described (see also Figure 1), and the appropriate informed consent by the individuals was obtained.

The origin of all other populations used for comparisons is detailed in Table 1.

HLA Genotyping and DNA Sequencing. Generic HLA class I (*A* and *B*) genotyping (resolution equivalent to the standard serology techniques) and high-resolution HLA class II (*DRB1* and *DQB1*) was performed using a reverse dot-blot technique with the Automated Innolipa system (Innogenetics N.V., Zwijndrecht, Belgium). *HLA-A*, *-B*, *-DRB1*, and *-DQB1* allele DNA sequencing was done only when indirect DNA typing yielded ambiguous results (Arnaiz-Villena et al. 1992).

Statistical Analysis. Statistical analyses were performed using Arlequin v2.000 software (Schneider et al. 2000). In summary, this program calculated *HLA-A*, *-B*, *-DRB1*, *-DQA1*, and *-DQB1* allele frequencies, Hardy-Weinberg equilibrium, and the linkage disequilibrium between two alleles at two different loci. Linkage disequilibrium and the level of significance (*p*) for 2×2 comparisons were determined using the formulae of Mattiuz and coworkers (Mattiuz et al. 1970) and the 11th International Histocompatibility Workshop methodology (Imanishi et al. 1992b). In addition, the most frequent complete haplotypes were tentatively deduced using a methodology developed in the 11th International Histocompatibility Workshop. The criteria used were: (1) the two-, three-, and four-*HLA*-loci haplotype frequencies (Imanishi et al. 1992c; Clayton and Lonjou 1997); (2) the previously described haplotypes in other populations (Imanishi et al. 1992c; Clayton and Lonjou 1997; Arnaiz-Villena et al. 1995, 1997, 1999, 2001a, 2001b; Martínez-Laso et al. 1995, 2001; Gomez-Casado et al. 2000); and (3) haplotypes if they appeared in two or more individuals and the alternative haplotype was well defined (Imanishi et al. 1992c; Clayton and Lonjou 1997; Arnaiz-Villena et al. 1995, 1997, 1999, 2001a, 2001b; Martínez-Laso et al. 1995, 2001; Gomez-Casado et al. 2000). In order to compare allelic and haplotype *HLA* frequencies with other populations, the reference tables of the 11th and 12th International *HLA* Workshops were used (Imanishi et al. 1992c; Clayton and Lonjou 1997) (also see Table 1). The software DISPAN containing the programs GNKDST and TREEVIEW (Nei 1973, 1983) was used to construct phylogenetic trees (dendrograms) with the allelic frequencies by using the neighbor-joining (NJ) method (Saitou and Nei 1987) with genetic distances between populations (DA) (Nei 1972). Correspondence analysis in a bidimensional representation was

Table 1. Populations Used in the Present Study

<i>Identification No.</i>	<i>Population</i>	<i>n</i>	<i>Reference</i>
1	Chuvash	82	Present study
2	Finns	157	Clayton and Lonjou 1997
3	Germans	256	Imanishi et al. 1992c
4	Russians	200	Kapustin et al. 1999
5	Belgians	55	Imanishi et al. 1992c
6	Austrians	94	Imanishi et al. 1992c
7	Czech	79	Imanishi et al. 1992c
8	Bulgarians	116	Clayton and Lonjou 1997
9	Danish	124	Imanishi et al. 1992c
10	Hungarians	98	Imanishi et al. 1992c
11	Romanians	327	Martinez-Laso et al. (unpublished results)
12	Uralic (from the Caucasus)	48	Imanishi et al. 1992c
13	Italians	284	Imanishi et al. 1992c
14	French	179	Imanishi et al. 1992c
15	Spaniards	176	Martinez-Laso et al. 1995
16	Portuguese	228	Arnaiz-Villena et al. 1997
17	Spanish-Basques	80	Martinez-Laso et al. 1995
18	Sardinians	91	Imanishi et al. 1992c
19	Cretans	135	Arnaiz-Villena et al. 1999
20	Macedonians	172	Arnaiz-Villena et al. 2001a
21	Armenians	105	Imanishi et al. 1992c
22	Turks	228	Arnaiz-Villena et al. 2001b
23	Algerians	102	Arnaiz-Villena et al. 1995
24	Moroccans	98	Gomez-Casado et al. 2000
25	Georgians	80	Sanchez-Velasco and Leyva-Cobian 2001
26	Berbers (Souss)	98	Izaabel et al. 1998
27	Moroccan Jews	94	Roitburg-Tambur et al. 1995
28	Ashkenazi Jews	80	Martinez-Laso et al. 1996
29	Non-Ashkenazi Jews	80	Martinez-Laso et al. 1996
30	Lebanese KZ	93	Clayton and Lonjou 1997 (and -NS)
31	Egyptians (Siwa)	101	Clayton and Lonjou 1997
32	Koreans	100	Imanishi et al. 1992c
33	Khoton-Mongolians	85	Munkhbat et al. 1997
34	Khalk-Mongolians	202	Munkhbat et al. 1997
35	Tuvinians	197	Martinez-Laso et al. 2001
36	Singapore-Chinese	71	Imanishi et al. 1992c
37	Manchu	50	Geng et al. 1995
38	Buyi	70	Imanishi et al. 1992c
39	Japanese	493	Imanishi et al. 1992c (and Ainu)
40	Udegeys	23	Grahovac et al. 1998
41	Koryaks	92	Grahovac et al. 1998
42	Chukchi	59	Grahovac et al. 1998
43	Kets	22	Grahovac et al. 1998
44	Evenks	35	Grahovac et al. 1998
45	Nivkhs	32	Grahovac et al. 1998
46	Eskimos	35	Grahovac et al. 1998
47	Athabascans (Na-Dene)	124	Monsalve et al. 1998
48	Tlingit (Na-Dene)	53	Imanishi et al. 1992c

n = number of individuals analyzed for each population.

a. KZ = Kafar Zubian (town).

obtained by using the VISTA v5.02 computer program (Young and Bann 1996) (<http://forrest.psych.unc.edu>). Correspondence analysis consists of a geometric technique that may be used for displaying a global view of the relationships among populations according to HLA (or other) allele frequencies. This methodology is based on the allelic frequency variance among populations (similar to the classical components methodology) and on the topological display of their statistical differences.

Results

Characteristic HLA Allele Frequencies of the Chuvashian Population Compared to Other Populations. The expected and observed allelic frequencies for *HLA-A*, *-B*, *-DRB1*, and *-DQB1* loci do not differ significantly. There is no deviation from Hardy-Weinberg equilibrium. Table 2 shows the HLA allele frequencies found in the Chuvash population. Thirteen different *HLA-A* and 25 different *HLA-B* alleles were observed among the Chuvash. Six *HLA-A* alleles and four *HLA-B* alleles had frequencies higher than 5% (*A*01*, *A*02*, *A*03*, *A*11*, *A*24*, *A*68*, *B*07*, *B*18*, *B*27*, and *B*35*), and these are characteristic of Central European and Mediterranean populations (Imanishi et al. 1992c; Clayton and Lonjou 1997; Arnaiz-Villena et al. 1995, 1999). Alleles with frequencies higher than 3.5% may overlap the ones having a 5% frequency. With regard to HLA class II alleles, 18 different *DRB1* alleles were found and only six had frequencies higher than 5%; *DQ* allele frequencies reflect the *DRB1* locus allele distribution due to the strong linkage disequilibrium between these two loci.

Two types of analyses were carried out in order to compare Chuvashian HLA frequencies with other Central European, Mediterranean, Siberian, and Na-Dene population frequencies: (1) with *DRB1-DQB1* data, which is probably a more informative and discriminating methodology; and (2) with generic (low-resolution) *DR-DQ* data. These two types of analyses were both performed because some of the populations used for comparison lacked *HLA-DRB1* and *-DQB1* high-resolution typing, and only generic *HLA-DR* and *-DQ* data were available (Portuguese, Turks, Iranians, Armenians, Egyptians, Bulgarians, Czech, Austrian, Belgians, Finns, Romanians, Hungarians, and Uralic [from the Caucasus]; see Table 1). These partially HLA-typed populations should have been ignored, but they could be analyzed conjointly taking into account *DRB1* and *DQB1* or generic *DR* and *DQ* frequencies (Tables 3 and 4; Figures 2, 3, 4, and 5) (Imanishi et al. 1992c). Finally, it should be pointed out that class I generic typing tends to homogenize the comparisons based on *DRB1-DQB1* high resolution typing (Gomez-Casado et al. 2000); one class I allele obtained by generic DNA typing may contain several class I alleles, and many populations lacked *HLA-A* and *-B* data (see references from Table 1).

Figure 2 depicts an HLA class II (*DRB1-DQB1*) neighbor-joining tree. Populations are grouped into two main branches. The first includes Siberians and

Table 2. HLA-A, -B, -DRB1, and -DQB1 Allele Frequencies in the Chuvashian Population

Alleles	Allele Frequencies %	Alleles	Allele Frequencies %
<i>HLA-A</i>		<i>HLA-DRB1</i> ²	
A*01	7.3	0101	18.3
A*02	25.0	0102	1.9
A*03	15.8	0301	7.9
A*11	6.1	0401	1.2
A*23	3.7	0403	1.8
A*24	18.9	0405	1.2
A*25	3.7	0701	14.6
A*26	1.2	0801	7.9
A*30	1.2	0901	4.9
A*31	4.9	1101	8.5
A*32	3.7	1103	0.6
A*33	1.8	1104	3.0
A*68	6.7	1201	1.8
		1301	4.3
		1302	0.6
<i>HLA-B</i>		1404	0.6
B*07	18.3	1501	19.5
B*08	3.7	1601	1.2
B*13	2.4		
B*14	1.8	<i>HLA-DQB1</i> ^b	
B*15	1.2	02	18.9
B*18	6.7	0301	14.0
B*27	7.9	0302	3.0
B*35	18.9	0303	8.5
B*37	1.9	0305	0.6
B*38	1.8	0402	7.9
B*39	3.7	0501	21.3
B*40	0.6	0502	1.8
B*41	0.6	0601	0.6
B*44	4.9	0602	18.3
B*47	1.2	0603	4.3
B*48	1.2	0605	0.6
B*49	3.0		
B*50	4.3		
B*51	4.3		
B*54	0.6		
B*56	0.6		
B*57	2.4		
B*40(B60) ^a	2.4		
B*40(B61)	4.9		
B*15(B62)	0.6		

a. Numbers in brackets indicate the serologic antigen most probably corresponding to the genetic allele obtained.

b. Alleles *DQB1**0201 and 0202 were all assigned as *DQB1**02.

Table 3. Genetic Distances (DA) between Chuvash and Other Populations ($\times 10^2$) Obtained by Using *HLA-DRB1-DQB1* Allele Frequencies (See Table 1 for Identification of Populations)

<i>Population</i>	<i>DRB1-DQB1</i>	<i>Population</i>	<i>DRB1-DQB1</i>
French	9.67**	Khoton-Mongolians	27.23**
Russians	10.52*	Lebanese-KZ	27.36***
Spaniards	11.93*	Manchu	27.53**
Germans	14.42**	Sardinians	28.94**
Danish	15.51*	Koreans	30.78*
Georgians	16.21**	Kets	31.49***
Tuvinians	16.67**	Singaporean-Chinese	36.11*
Italians	17.51**	Evenks	37.96***
Algerians	19.27**	Japanese	39.36*
Moroccans	20.20**	Nivkhs	44.79*
Berbers (Souss)	20.42**	Tlingit	46.31***
Cretans	20.50**	Buyi	46.33***
Khalk-Mongolians	21.81**	Athabascans	46.53**
Ashkenazi Jews	22.40**	Chukchi	48.75***
Non Ashkenazi Jews	26.32**	Koryaks	52.15***
Macedonians	26.75*	Udegeys	54.82*
Moroccan Jews	26.81**	Eskimos	60.32***

SE = Standard Error. Values are between 0–1. * = SE between 0.01– 0.03; ** = SE between 0.03–0.05; *** = SE between 0.05 –0.16.

Na-Dene Native Americans, and the second groups Asians and Caucasoids. The latter group is subdivided into three groups comprised of central Europeans (including Russians, Danish, and Chuvash), western Mediterraneans (Spaniards and North Africans), and eastern Mediterraneans (including Macedonians, Cretans, Jews, and Lebanese). This distribution is also confirmed in the correspondence analysis (Figure 3): the three groups are clearly delimited and an east-to-west Mediterranean and central European gradient is evident. The Chuvash population shows the closest genetic distance with the French followed by Russians, Spaniards, Germans, and Danes (Table 3). These results are confirmed using *DR* and *DQ* generic typings (see Figures 4 and 5), which were used in order to include other central and southern Europeans (Finns, Austrians, Belgians, Czech, Hungarians, and Romanians) and other Mediterranean populations (Portuguese, Iranians, Armenians, Egyptians, and Turks; see Table 1). A *DR-DQ* neighbor-joining tree (Figure 4) maintains the central European and the east-to-west Mediterranean distribution.

On the other hand, genetic distances obtained by using *DR-DQ* generic typing allele frequencies (Table 4) show that Russians, Finns, Austrians, Belgians, and Czechs are the populations closest to the Chuvash, followed by the other western and eastern Mediterranean populations.

Table 4. Genetic Distances (DA) between Chuvash and Other Populations ($\times 10^2$) Obtained by Using Generic *HLA-DR-DQ* Allele Frequencies (See Table 1 for Identification of Populations)

<i>Population</i>	<i>DR-DQ</i>	<i>Population</i>	<i>DR-DQ</i>
Russians	2.97**	Italians	9.23*
Finns	2.99**	Armenians	9.39**
Austrians	3.65*	Khalk-Mongolians	9.48*
Belgians	3.73**	Moroccan Jews	9.74**
Czech	3.74*	Turks	9.85*
Spaniards	3.97**	Manchu	9.95**
Portuguese	4.83**	Lebanese-KZ	10.59**
French	5.05**	Berbers (Souss)	10.67*
Bulgarians	5.39**	Egyptian	11.48**
Algerians	5.21**	Khoton-Mongolians	12.26*
Uralic	5.47*	Non Ashkenazi Jews	12.34**
Danish	5.59**	Singaporean-Chinese	14.14*
Macedonians	5.69**	Japanese	14.66*
Georgians	5.89*	Buyi	15.43*
Iranians	6.18**	Hungarians	16.74*
Sardinians	6.51*	Kets	17.33**
Romanians	7.01*	Athabascans	21.60**
Tuvinians	7.03*	Evenks	24.51***
Cretans	7.10**	Tlingit	27.63**
Moroccans	7.36*	Nivkhs	28.82**
Ashkenazi Jews	7.75**	Eskimos	30.53***
Germans	7.79*	Chukchi	31.21***
Spanish-Basques	8.24**	Udegeys	37.15***
Koreans	8.70*	Koryaks	44.28***

SE = Standard Error. Values are between 0–1. * = SE between 0.05–0.09; ** = SE between 0.09–0.15; *** = SE between 0.15–0.25.

HLA-A, -B, -DRB1, and -DQB1 Linkage Disequilibria among the Chuvash. Extended HLA haplotypes were defined in the Chuvash and compared with those previously reported in other populations (Table 5). *HLA-A-B* and *B-DRB1* two-loci linkage disequilibrium data (not shown) reveal that the most frequent combinations are characteristic of European and Mediterranean (western and eastern) populations (*B*07-DRB1*1501*, haplotype frequency [HF] = 12.6; *B*35-DRB1*0101*, HF = 7.2; *A*03-B*35*, HF = 7.2; *B*50-DRB1*0701*, HF = 4.3, *A*33-B*14*, HF = 1.8) The most frequent *HLA-A-B-DR-DQ* extended haplotypes found in the Chuvash population (Table 5) reflect mainly an admixture of central European (*A*03-B*35-DRB1*0101-DQB1*0501*) and Mediterranean (*A11*-B*35-DRB1*0701-DQB1*02*, *A*24-B*44-DRB1*0701-DQB1*02*, *A*33-B*14-DRB1*0102-DQB1*0501*) contributions (see Table 5 and footnote). In addition, several haplotypes have been found in the Chuvash not previously de-

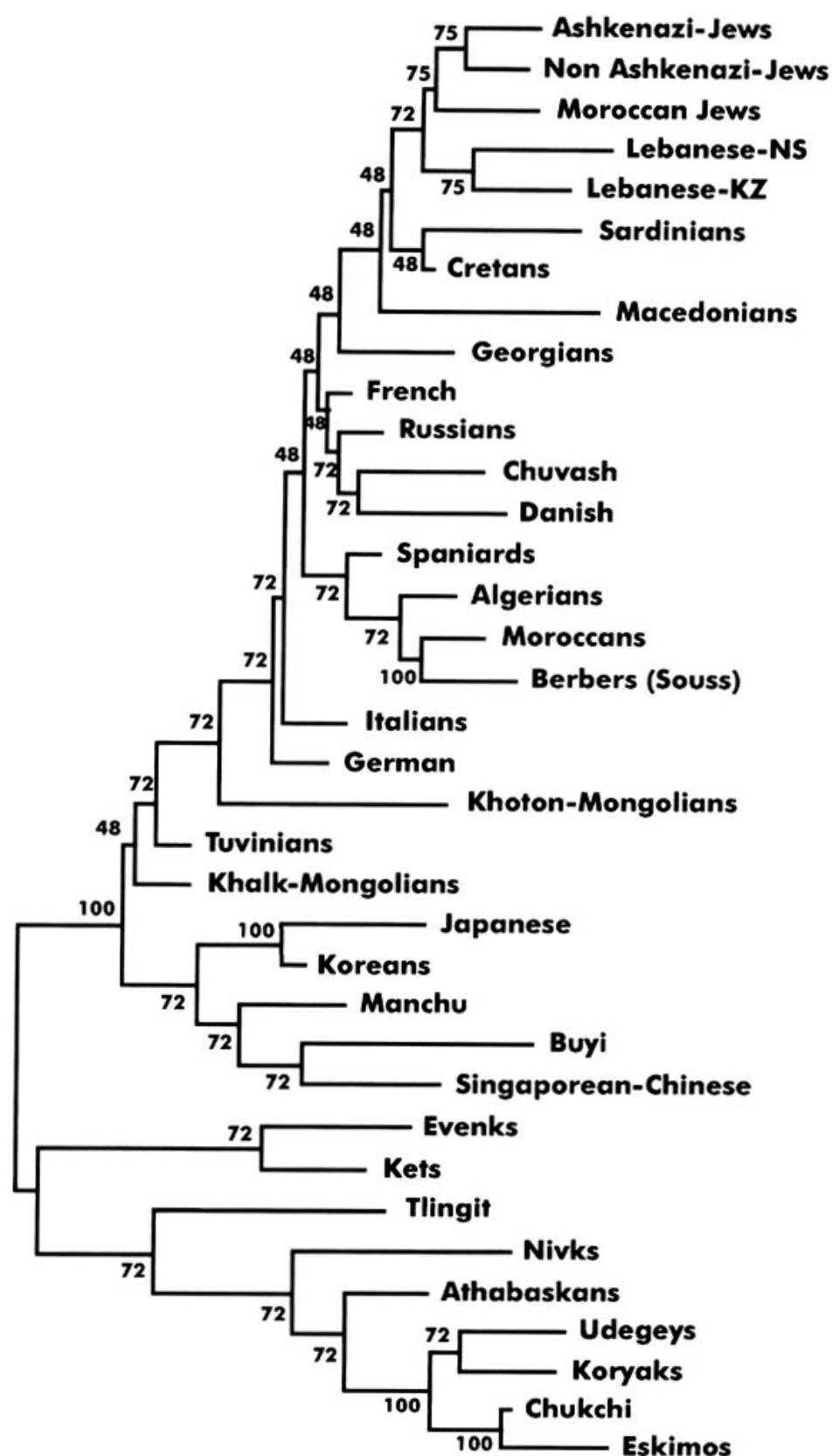


Figure 2. Neighbor-joining dendrogram showing relatedness between Chuvash and other populations. Genetic distances between populations (DA) were calculated by using *HLA-DRB1* and *-DQB1* (high resolution). Data from other populations were from references detailed in Table 1.

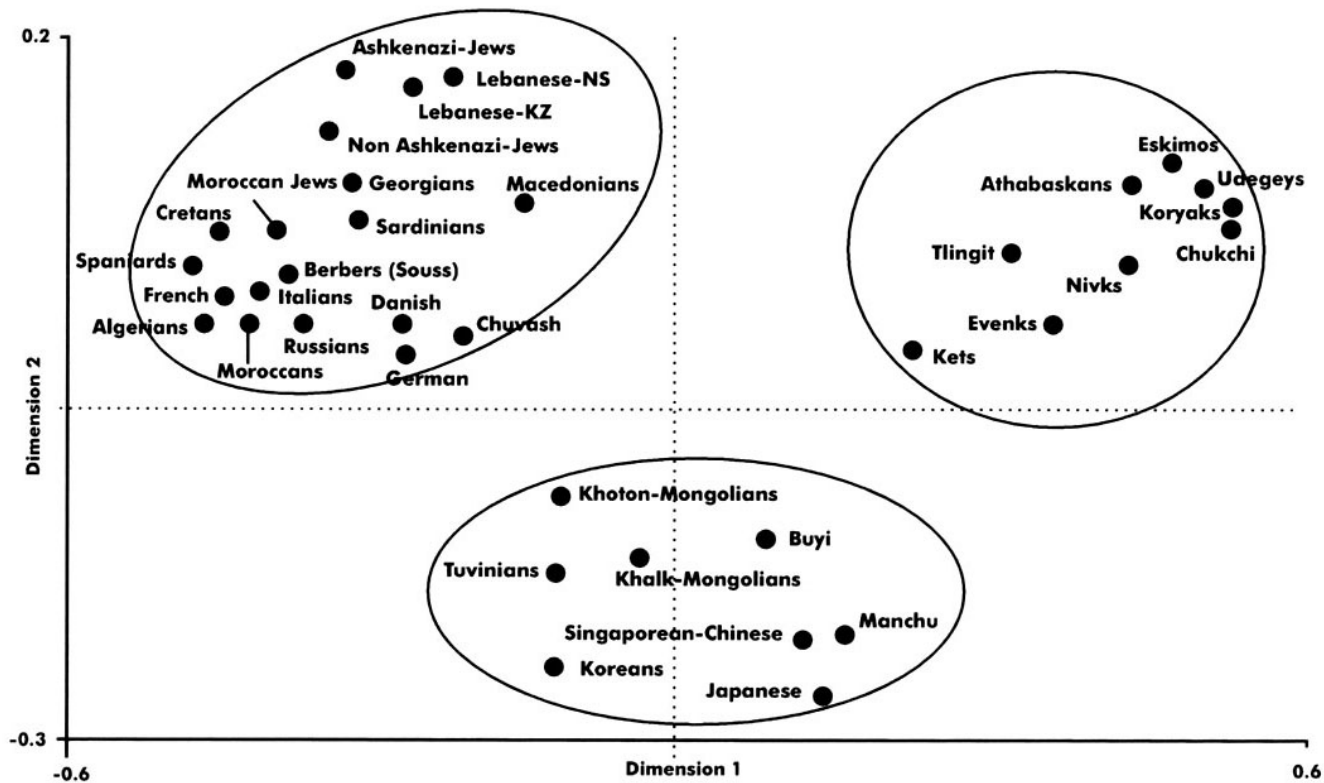


Figure 3. Correspondence analysis showing a global view of the relationship between Chuvash and other populations according to HLA allele frequencies in n dimensions (bidimensional representation). *HLA-DRB1* and *-DQB1* allele frequencies data.

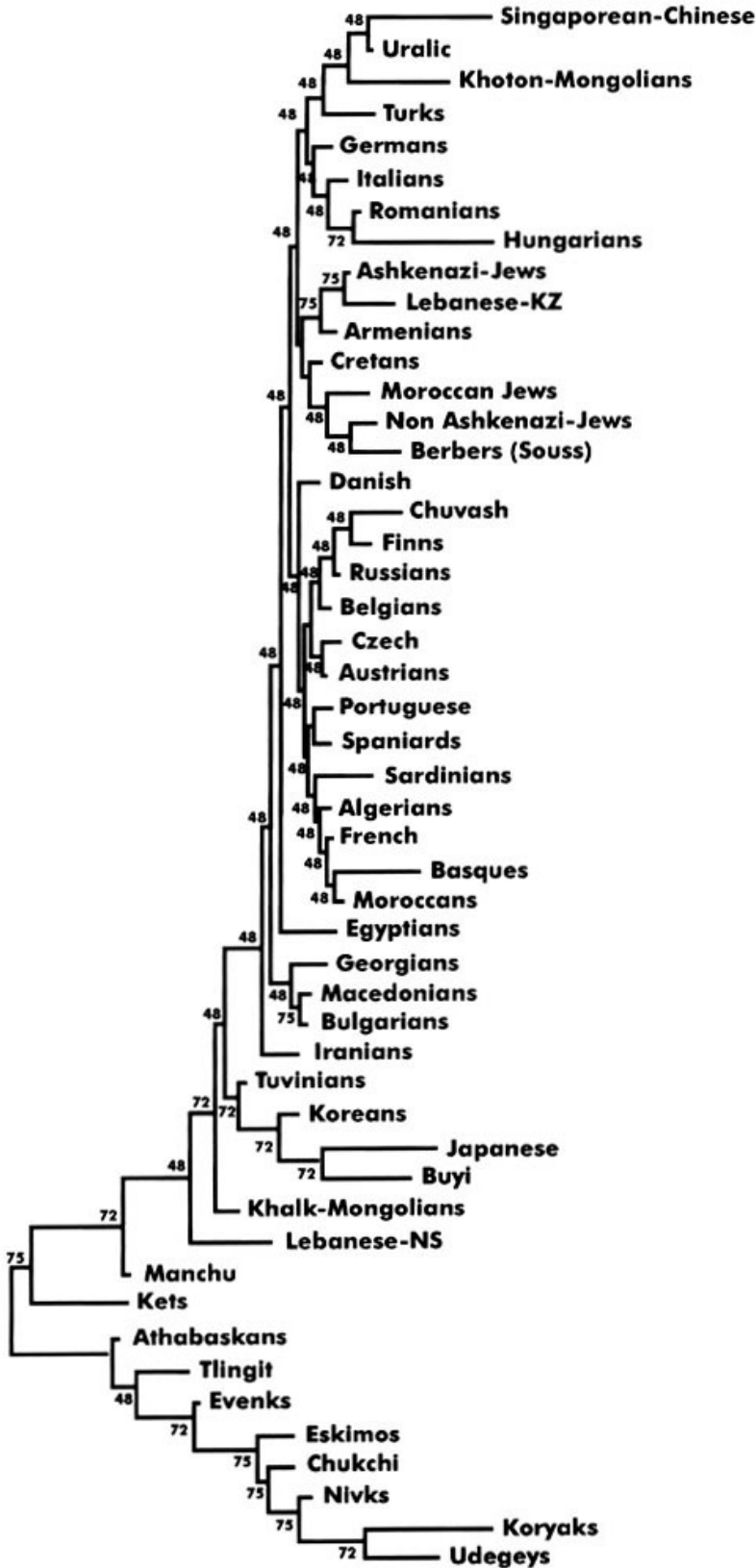


Figure 4. Neighbor-joining dendrogram showing relatedness between Chuvash and other populations. Genetic distances between populations (DA) were calculated by using *HLA-DR* and *-DQ* (generic typing). Data from other populations were from references detailed in Table 1.

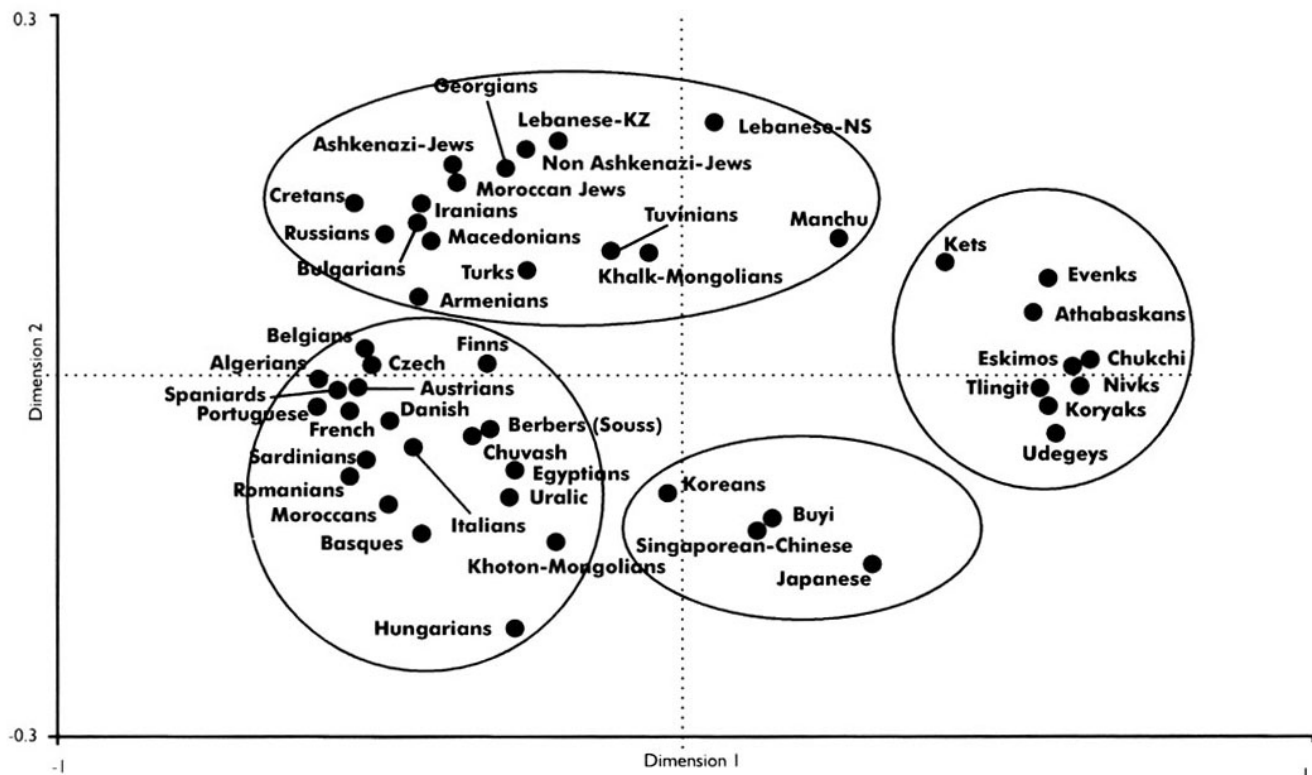


Figure 5. Correspondence analysis showing a global view of the relationship between Chuvash and other populations according to HLA allele frequencies in n dimensions (bidimensional representation). *HLA-DR* and *-DQ* generic typing allele frequencies data.

Table 5. Most Frequent *HLA-A*, *-B*, *-DRB1*, *-DQA1*, and *-DQB1* Extended Haplotypes in the Chuvashian and Other Populations (See References in Table 1)

Haplotypes	Haplotype Frequency (%)	Possible Origin
<i>A*03-B*35-DRB1*0101-DQB1*0501</i> ^a	5.5	Central European
<i>A*03-B*07-DRB1*1501-DQB1*0602</i> ^b	4.9	North African/Western European
<i>A*24-B*07-DRB1*1501-DQB1*0602</i>	4.3	Autochthonous
<i>A*68-B*35-DRB1*0801-DQB1*0402</i> ^c	3.7	?
<i>A*11-B*35-DRB1*0701-DQB1*0303</i> ^d	3.0	Chuvashian-Italian
<i>A*02-B*27-DRB1*0801-DQB1*0402</i> ^e	2.4	Uralic
<i>A*23-B*49-DRB1*1101-DQB1*0301</i> ^f	2.4	Chuvashian-Italian
<i>A*02-B*50-DRB1*0701-DQB1*0201</i> ^g	1.8	Eurasiatic
<i>A*01-B*50-DRB1*0701-DQB1*0201</i>	1.8	?
<i>A*24-B*44-DRB1*0701-DQB1*0201</i> ^h	1.8	Ibero-Basque
<i>A*25-B*18-DRB1*1501-DQB1*0602</i> ⁱ	1.8	Ibero-Macedonian
<i>A*33-B*14-DRB1*0102-DQB1*0501</i> ^j	1.8	Mediterranean

- a. Present also in Czech (haplotype frequency [HF] = 2.8), Italians (HF = 1.7), Romanians (HF = 2.4), and Hungarians (HF = 3.1).
 b. Haplotype described as North African/Western European.
 c. Also present in Tlingit (HF = 2.0).
 d. Present in low frequency in Italians (HF = 0.4).
 e. Partially found in Uralic with A3 allele (HF = 2.1).
 f. Present in low frequency in Italians (HF = 0.4).
 g. Haplotype found in Spaniards (HF = 1.2), Italians (HF = 0.7), Germans (HF = 1.0), and Manchu (HF = 2.2).
 h. Only found in the Basque population (HF = 2.8).
 i. Haplotype present in Spaniards (HF = 0.3), Portuguese (HF = 1.5), and Macedonians (HF = 1.2).
 j. Described as a Mediterranean haplotype.

Other low frequency haplotypes found in the Chuvash: *A*24-B*39-DRB1*1301-DQB1*0603* (HF = 1.8; autochthonous), *A*24-B*51-DRB1*0101-DQB1*0501* (HF = 1.8; autochthonous), *A*01-B*08-DRB1*0301-DQB1*0201* (HF = 1.2; defined as of pan-European origin), *A*11-B*18-DRB1*0101-DQB1*0501* (HF = 1.2; autochthonous), *A*02-B*07-DRB1*0301-DQB1*0201* (HF = 1.2; autochthonous), *A*02-B*07-DRB1*1501-DQB1*0102* (HF = 1.2; described as North African/Western European origin), *A*02-B*18-DRB1*0301-DQB1*0201* (HF = 1.2; found in Basques [HF = 2.1], Sardinians [HF = 3.1], Italians [HF = 0.8]), *A*02-B*27-DRB1*0101-DQB1*0501* (HF = 1.2; present only in Koreans [HF = 1.9]), *A*02-B*37-DRB1*0901-DQB1*0303* (HF = 1.2; autochthonous), *A*02-B*44-DRB1*0101-DQB1*0501* (HF = 1.2; found in Greeks [HF = 1.7] and Italians [HF = 0.7]), *A*02-B*47-DRB1*1501-DQB1*0602* (HF = 1.2; autochthonous), *A*02-B*40(B60)-DRB1*1101-DQB1*0301* (HF = 1.2; found in Czech [HF = 3.8]), *A*03-B*07-DRB1*0101-DQB1*0501* (HF = 1.2; found in Uralic population [HF = 3.1]), *A*23-B*44-DRB1*1104-DQB1*0301* (HF = 1.2; present in Sardinians [HF = 2.0]), *A*24-B*35-DRB1*1101-DQB1*0301* (HF = 1.2; found in Austrians [HF = 2.7], Italians [HF = 0.7], Germans [HF = 1.4], Romanians [HF = 0.7], and Georgians [HF = 1.2]), *A*30-B*13-DRB1*0701-DQB1*02* (HF = 1.2; described of Eurasiatic origin) and *A*32-B*08-DRB1*0301-DQB1*0501* (HF = 1.2; autochthonous). See Imanishi et al. 1992c; Clayton and Lonjou 1997; Arnaiz-Villena et al. 1995; Arnaiz-Villena et al. 1997; Arnaiz-Villena et al. 1999; Arnaiz-Villena et al. 2001a; Arnaiz-Villena et al. 2001b; Martinez-Laso et al. 1995; Martinez-Laso et al. 2001; Gomez-Casado et al. 2000.

scribed in any other population, reflecting an autochthonous HLA genetic background with an uncertain origin.

On the other hand, low frequency haplotypes defined in this population (see footnote to Table 5) confirm the various contributions of central Europeans and Caucasians (Georgians) (*A*24-B*35-DRB1*1101-DQB1*0301*), Euroasiatics (*A*30-B*13-DRB1*0701-DQB1*02*), and Mediterraneans (*A*02-B*18-DRB1*0301-DQB1*02*, *A*23-B*44-DRB1*1104-DQB1*0301*). These haplotype results are concordant with those obtained by the allele frequency analyses (genetic distances, neighbor-joining trees, and correspondence analysis; see above). Otherwise, a migration from Chuvashian area to central Europe and the Mediterranean area cannot be discarded.

Historical Background and Discussion

The Chuvash and Their Relationship with Other Populations. The Chuvash population is most closely related to central European populations (Czech, Austrians, Belgians, Finns, Russians), Mediterraneans, and Georgians, as demonstrated by the neighbor-joining trees and correspondence analyses (Figures 2, 3, 4, and 5) and with the closest genetic distance values (Tables 3 and 4). These relationships reflect the European, Middle Eastern (Iranians, Jews, Lebanese, Armenians), and Mediterranean (Portuguese, Spaniards, Algerians, French, Macedonians) origins of this population, in contrast with the postulated central Asian origin (see introductory paragraphs). The Chuvash have been described as one of the largest non-Slav communities inhabiting the Volga River in the southwestern part of Russia.

On the other hand, specific Turkish alleles or haplotypes (comparing the Chuvash with the present-day population of Turkey) (Arnaiz-Villena et al. 2001b) have not been found in the Altai-Turkic-speaking Chuvash population. The possibilities that the actual Turkish HLA background in Turkey is different from that of the ancestral Turkic population in Central Asia or that only a large cultural and low genetic Turkish influence was brought into the region without modifying the genetic background of the local population may not be discarded (Arnaiz-Villena et al. 2001b). These results are also confirmed in neighbor-joining, correspondence, and genetic distance analyses using *HLA-DR-DQ* markers. The Turks cluster together with other Mediterranean populations and far from the Chuvash.

Figures 2, 3, 4, and 5 show that Siberians are outside the Oriental cluster and do not indicate any close genetic affinities to the Chuvash. All Siberian populations cluster together and also cluster with the Na-Dene speakers (Athabascans and Tlingit) and Eskimos, both in the dendrograms and in the correspondence analyses (see also Martinez-Laso et al. 2001; Arnaiz-Villena et al. 2000). These results are confirmed with the DA genetic distances data (Tables 3 and 4), where the Chuvash-Siberians distance is 30%–50%.

However, some Uralic haplotypes have been found in the Chuvashian pop-

ulation: (1) a part of *A*02-B*27-DRB1*0801-DQB1*0402* and (2) *A*03-B*07-DRB1*0101-DQB1*0501* in low frequency (see Table 5 and footnote), reflecting a certain influence of the populations living in the region close to the Chuvash.

Chuvash, Bulgarians, Caucasians and Bulgars. The Chuvash are believed to originate from the ancient Bulgars that inhabited the western region of the Volga River and came from Central Asia in the 4th century A.D. The Great Bulgaria was divided into five different hordes, the fourth corresponding to the Chuvash and the fifth representing present-day Bulgarians (see introductory paragraphs). From the data obtained in the present work, the genetic backgrounds of both populations are clearly different. The Chuvash have a central European and some Mediterranean genetic background (probably coming from the Caucasus), while the Bulgarians have a classical eastern Mediterranean composition, grouping with Macedonians and Iranians in the neighbor-joining trees obtained by using *DR* and *DQ* genetic distances (Figure 4) and confirmed by correspondence analysis (Figure 5). It is possible that only a cultural and low genetic Bulgar influence was brought into the region without modifying the genetic background of the local population.

The Chuvash relatedness to Caucasian (Georgians) and Middle Eastern people suggests that the Chuvash genetic pool has characteristics from ancient Middle Eastern, Caucasian, and Mesopotamian people. A south-to-north gene flow may then be postulated for the Chuvash while only northern European or Asian relatedness was previously postulated (Livshits et al. 1999a). Middle East, Anatolian, and Caucasian people are likely to genetically represent the ancient population stock of the area, that is, Hittites and Mesopotamians. Other invasions have had a noticeable cultural but not a genetic impact (Arnaiz-Villena et al. 2001c, 2002). On the other hand, whether the Chuvash are a remnant of the Bulgar Hordes cannot be resolved with the data presented in this paper, and further studies are necessary.

Acknowledgments This work was supported in part by the Spanish Ministry of Education through grants PM95-57 and PM96-21, and the Madrid Regional Government through grants 06/70/97 and 8.3/14/98. We are grateful to Alberto Garcia for his help with art design work on the computer.

Received February 2002; revision received 7 November 2002.

Literature Cited

- Arnaiz-Villena, A., M. Timón, A. Corell et al. 1992. Primary immunodeficiency caused by mutations in the gene encoding the CD3- γ subunit of the T-lymphocyte receptor. *N. Eng. J. Med.* 327:529.

- Arnaiz-Villena, A., D. Benmamar, M. Alvarez et al. 1995. HLA allele and haplotype frequencies in Algerians. Relatedness to Spaniards and Basques. *Hum. Immunol.* 43:259.
- Arnaiz-Villena, A., J. Martinez-Laso, E. Gomez-Casado et al. 1997. Relatedness among Basques, Portuguese, Spaniards, and Algerians studied by HLA allelic frequencies and haplotypes. *Immunogenetics* 47:37.
- Arnaiz-Villena, A., P. Iliakis, M. Gonzalez-Hevilla et al. 1999. The origin of the Cretan population as determined by characterization of HLA alleles. *Tissue Antigens* 53:213.
- Arnaiz-Villena, A., G. Vargas-Alarcon, J. Granados et al. 2000. HLA in Mexican Mazatecos, the peopling of the Americas and the uniqueness of Amerindians. *Tissue Antigens* 56:405.
- Arnaiz-Villena, A., K. Dimitroski, A. Pacho et al. 2001a. HLA genes in Macedonians and the sub-Saharan relatedness of the Greeks. *Tissue Antigens* 57:118.
- Arnaiz-Villena, A., M. Karin, N. Bendikuzé et al. 2001b. HLA alleles and haplotypes in the Turkish population: Relatedness to Kurds, Armenians and other Mediterraneans. *Tissue Antigens* 57:308.
- Arnaiz-Villena, A., J. Martinez-Laso, and J. Alonso-Garcia. 2001c. The correlation between languages and genes: The Usko-Mediterranean peoples. *Hum. Immunol.* 62:1051.
- Arnaiz-Villena, A., E. Gomez-Casado, and J. Martinez-Laso. 2002. Population genetic relationships between Mediterranean populations determined by HLA allele distribution and a historic perspective. *Tissue Antigens* 60:111.
- Clayton, J., and Lonjou, C. 1997. Allele and Haplotype frequencies for HLA loci in various ethnic groups. In *Genetic Diversity of HLA, Functional and Medical Implications*, D. Charron, ed. Vol. 1. Paris, France: EDK, 665–820.
- Geng, L., Imanishi, T., K. Tokunaga, D. Zhu et al. 1995. 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.
- Gomez-Casado, E., P. del Moral, J. Martinez-Laso et al. 2000. HLA genes in Arabic-speaking Moroccans: Close relatedness to Berbers and Iberians. *Tissue Antigens* 55:239.
- Grahovac, B., R-I. Sukernik, C. O'hUigin et al. 1998. Polymorphism of the HLA class II loci in Siberian populations. *Hum. Genet.* 102:27.
- Imanishi, T., A. Wakisaka, and T. Gojorobi. 1992a. Genetic relationships among various human populations indicated by MHC polymorphisms. In *HLA 1991*, K. Tsuji, M. Aizawa, and T. Sasazuki, eds. Vol. 1. Oxford, UK: Oxford University Press, 627–632.
- Imanishi, T., T. Akaza, A. Kimura et al. 1992b. Estimation of allele and haplotype frequencies for HLA and complement loci. In *HLA 1991*, K. Tsuji, M. Aizawa, and T. Sasazuki, eds. Vol. 1. Oxford, UK: Oxford University Press, 76–79.
- Imanishi, T., Akaza, T., A. Kimura et al. 1992c. Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In *HLA 1991*, K. Tsuji, M. Aizawa, and T. Sasazuki, eds. Vol. 1. Oxford, UK: Oxford University Press, 1065–1220.
- Izaabel, H., H.J. Garchon, S. Caillat-Zucman et al. 1998. HLA class II DNA polymorphism in a Moroccan population from the Souss, Agadir area. *Tissue Antigens* 51:106.
- Kapustin, S., A. Lyschov, J. Alexandrova et al. 1999. HLA class II molecular polymorphism in Healthy Slavic individuals from western Russia. *Tissue Antigens* 54:517.
- Livshits, G., E. Ginsburg, and E. Kobiliansky. 1999a. Heterogeneity of genetic control of blood pressure in ethnically different populations. *Hum. Biol.* 71:685.
- Livshits, G., D. Karasik, O. Paulovsky et al. 1999b. Segregation analysis reveals a major gene effect in compact and cancellous bone mineral density in 2 populations. *Hum. Biol.* 71:156.
- Marsh, S.G.E., J.G. Bodmer, E.D. Albert et al. 2001. Nomenclature for factors of the HLA System, 2000. *Hum. Immunol.* 62:419.
- Martínez-Laso, J., D. De Juan, N. Martínez-Quiles et al. 1995. The contribution of the HLA-A, -B, -C and -DR, -DQ DNA typing to the study of the origins of Spaniards and Basques. *Tissue Antigens* 45:237.
- Martínez-Laso, J., E. Gazit, E. Gómez-Casado et al. 1996. HLA DR and DQ polymorphism in Ashkenazi and non-Ashkenazi Jews: Comparison with other Mediterraneans. *Tissue Antigens* 47:63.

- Martinez-Laso, J., M. Sartakova, L. Allende et al. 2001. HLA molecular markers in Tuvinians: A population with both Oriental and Caucasoid characteristics. *Ann. Hum. Genet.* 65:245.
- Mattiuz, P.L., D. Ihde, A. Piazza et al. 1970. New approaches to the population genetics and segregation analysis of the HLA system. In *Histocompatibility Testing*, W.F. Wodmer, ed. Copenhagen, Denmark: Munksgaard, 193–206.
- Monsalve, M.V., G. Edin, and D.V. Devine. 1998. Analysis of HLA class I and class II in Na-Dene and Amerindian populations from British Columbia, Canada. *Hum. Immunol.* 59:48.
- Munkhbat, B., T. Sato, M. Hagihara et al. 1997. Molecular analysis of HLA polymorphism in Khoton-Mongolians. *Tissue Antigens* 50:124.
- Nei, M. 1972. Genetic distances between populations. *Am. Nat.* 106:283.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70:3321.
- Nei, M., F. Tajima, and Y. Tateno. 1983. Accuracy of estimated phylogenetic trees from molecular data II. Gene frequency data. *J. Mol. Evol.* 19:153.
- Roitberg-Tambur, A., C.S. Witt, A. Friedmann et al. 1995. Comparative analysis of HLA polymorphism at the serologic and molecular level in Moroccan and Ashkenazi Jews. *Tissue Antigens* 46:104.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406.
- Sanchez-Velasco, P., and F. Leyva-Cobián. 2001. The HLA class I and class II allele frequencies studied at the DNA level in the Upper Caucasus and their relationships to western European populations. *Tissue Antigens* (in press).
- Schneider, S., D. Roessli, and L. Excoffier. 2000. *Arlequin: A software for population data analysis*. Geneva: Genetics and Biometry Lab. Dept. of Anthropology and Ecology. University of Geneva. <http://lgb.unige.ch/arlequin/software/>
- Tokunaga, K., T. Imanishi, K. Takahashi et al. 1996. On the origin and dispersal of East Asian populations as viewed from HLA haplotypes. In *Prehistoric Mongoloid Dispersals*, T. Akazawa and E.J.E. Szatmary, eds. Oxford, UK: Oxford University Press, 187–197.
- Young, F.W., and C.M. Bann. 1996. A Visual Statistics system. In *Statistical Computing Environments for Social Researchers*, R.A. Stine and J. Fox, eds. New York, NY: Sage Publications, 207–236.