

Original Research Article

Genetic Structure of Algerian Populations

PHILIPPE LEFEVRE-WITIER,^{1*} HADI AIRECHE,² MOHAMED BENABADJI,² PIERRE DARLU,³
KRISTIN MELVIN,⁴ ANDRE SEVIN,⁵ AND MICHAEL H. CRAWFORD⁴

¹Centre d'Immunologie et Génétique Humaine, CIGH-CNRS, 31300 Toulouse, France

²Centre National Algérien de Transfusion Sanguine, Hospital Mustapha, 16000, Alger, Algeria

³Institut National de la Santé et de la Recherche Médicale, U535, B.P. 1000, 94817 Villejuif, France

⁴Laboratory of Biological Anthropology, University of Kansas, Lawrence, Kansas 66045

⁵Centre d'Anthropologie, 39 Allées J. Guesde, 3100 Toulouse, France

ABSTRACT Blood samples were collected in Algeria from 4,444 army recruits and tested for 10 genetic polymorphic systems. These samples were collected from territorial Wilayas (administrative units of Algeria) from which the young soldiers had originated. Based on similar geography and economic and political history, these Wilayas were clustered into 10 regions. These regions, not part of the governmental administrative units, were characterized by allelic frequencies, and analyzed using R-matrix principal components, Wright's F_{ST} , spatial autocorrelation, and Mantel tests. Hierarchical relationships between the culturally defined regions were examined using two different analytical methods of phylogenetic tree constructions: neighbor-joining, and unweighted pair group average arithmetic (UPGMA). These results indicated the predominance of genetic homogeneity due to the gene flow between regions, but with some migration emanating from sub-Saharan Africa and Mediterranean Europe. Wright's F_{ST} value of 0.0063, based on 16 alleles, suggested a relatively small genetic microdifferentiation of the regions. In Algeria, gene flow apparently swamped most of the effects of stochastic processes and disrupted the relationship between geography and genetics, as characterized by the isolation-by-distance model. Some genetic differences and similarities were observed between regions or clusters of regions. The resulting genetic structure of the Algerian populations is best explained by a combination of gene flow, ecology, and history. *Am. J. Hum. Biol.* 18:492–501, 2006. © 2006 Wiley-Liss, Inc.

The Maghrib populations of the northern part of the African continent are currently located in three modern nations: Tunisia, Algeria, and Morocco (Fig. 1). The people of this large sub-arid or arid territory are either Berber- or Arabic-speakers, following their religious and cultural conversion to Islam after 670 AD.

CLASSICAL MARKERS

During the last 60 years, numerous studies were conducted on the genetic markers, their frequencies, and their distributions in Tunisia, Algeria, and Morocco. All studies focused on small geographic isolates or tribal groups. Only Kossovitch (1953), during the 1930–1938 time frame, attempted to construct gene maps based on ABO and rhesus blood group polymorphic distributions for Morocco. However, field investigations started in the late 1930s and early 1940s could not be completed because of the disruption associated with World War II. However, Benabadji and Chamla (1971) reinitiated the genetic studies of this region by examining a large sample of Algerian blood donors for the ABO and D/d rhesus

alleles. As part of a large collaborative project involving different laboratories from the Paul Sabatier Medical University of Toulouse, the French National Center for Scientific Research (CNRS), and the National Blood Center of Algeria, P. Lefevre-Witier and M. Benabadji conducted a series of projects on Tuareg tribes, mostly from the Algerian Sahara. Since 1964, multidisciplinary studies have been conducted on sedentary and nomadic populations of the Algerian Sahara, mostly on Tuareg tribes of the central region. Over a period of 30 years, more than 10,000 samples were examined for an ever-increasing number of red cell and serum genetic polymorphisms (Aireche et al., 1982; Aireche and Benabadji, 1986; Benabadji et al., 1969; Benabadji, 1970; Constans et al.,

This study is dedicated to the memory of the late Mohamed Benabadji.

*Correspondence to: Dr. Philippe Lefevre-Witier, Le Pied du Bois, BP 28, Entrepierres 04201, Sisteron Cedex, France. E-mail: plw@cict.fr

Received 6 October 2005; Accepted 16 January 2006

Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ajhb.20511

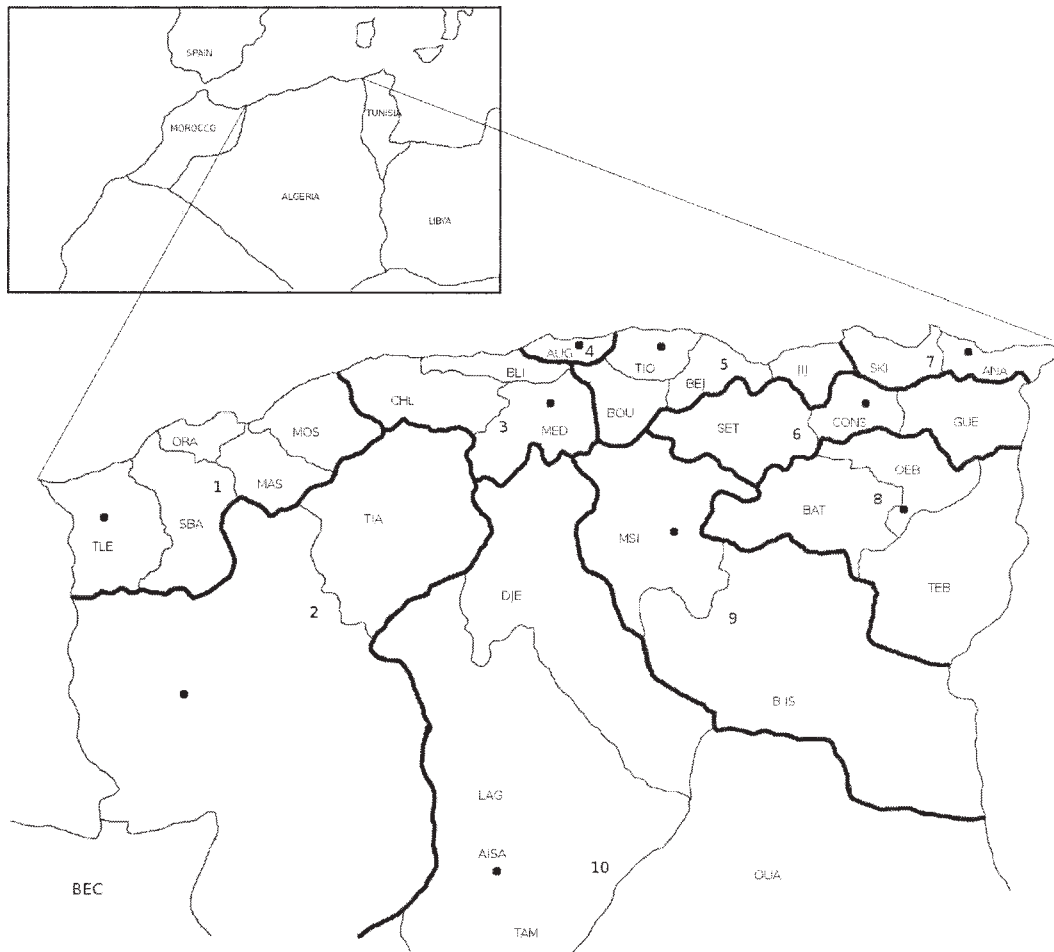


Fig. 1. The geographical locations of the Regions and Wilayas of Algeria. Region 1: Sidi Bel Abbes (SIBA), Region 2: Saida (SAID), Region 3: Miliana (MILI), Region 4: Algier (ALGE), Region 5: Bejaia (BEJA), Region 6: Constantine (CONS), Region 7: Annaba (ANNA), Region 8: Ain Beida (AIBE), Region 9: Msila (MSIL), Region 10: Ain Salah (AISA).

1980; Lefevre-Witier, 1972, 1974, 1982, 1996; Lefevre-Witier and Vergnes, 1977; Lefevre-Witier et al., 1968; Ruffle and Lefevre-Witier, 1972). Bosch et al. (1997) synthesized the available data on classic markers from North African populations by using principal-components analyses (PCAs) and synthetic gene maps. In the first PCA, they detected an east-west gradient, with the highest factor loadings in Egypt, and the lowest in southern Algeria.

MOLECULAR SURVEYS

A number of recent studies surveyed North African populations based on Nonrecombining Y, single-nucleotide polymorphisms (STRs),

and mitochondrial DNA (mtDNA). The mtDNA analysis was of a west Algerian student sample ($N = 50$) from the Oran region, and concluded that Algerians are closest genetically to an Israeli-Arab population, and show less affinity to Mediterranean populations, i.e., Sicily, Majorca, Calabria, and Sardinia (Ivanova et al., 1999). Similarly, Bosch et al. (2000) sampled northwest Africans from Morocco, north central and southern Berbers, Saharawis, and Mozabites. They tested DNA samples with two panels of autosomal STRs and compared them to reference populations. The North African groups clustered with an Arab sample. A set of 119 binary markers and 15 microsatellites was used by Arredi et al.

(2004) to characterize the Y-chromosome haplogroups of 275 men from Algeria, Tunisia, and Egypt. They interpreted the results of Y-chromosome typing as indicative of an expansion of an early pastoral and agricultural population with considerable gene flow. The Y-chromosome diversity is geographically structured, and fits the isolation-by-distance model.

The present study reports on: 1) red cell blood group variation in a large number of samples from populations of Algeria; 2) geographical patterns of gene frequency distributions which were used to characterize the genetic structure of Algeria; and 3) tests of the isolation-by-distance model through the examination of genetic diversity in the Maghrib regions, utilizing a spatial autocorrelation analysis.

BRIEF HISTORY OF ALGERIA

For the last 100,000 years, *Homo sapiens* populations have inhabited the Maghrib, the northwestern region of the African continent, corresponding to an area defined by 30° longitude East and West of the Greenwich meridian and approximately 15° latitude North, between 20–35°. The actual climatic conditions in this area range from Mediterranean semiaridity in the northern part, to an abiotic desert in the south (Fig. 1).

Some researchers postulated that the Sahara region offered Paleolithic humans slightly better living conditions than what was found in the colder and mountainous Maghrib. However, this northern part of Africa (Sahara, Mahgrib, and Mahchrek) apparently provided early humans with a relatively homogenous environment during the Paleolithic. A more striking contrast appeared during the Neolithic period, when aridity became severe in the Sahara, thereby limiting further demographic expansion to a greater degree than in the Maghrib.

The increasing archeological knowledge of the pre- and protohistoric periods indicates a major occupation of the Maghrib by Berber-speaking and so-called “Berber” populations. Wherever their exact place of origin, the Berbers were most likely autochthonous to the region (Camps, 1974). Dispersed within a zonal area of more than 4,000 km from the Red Sea to the Atlantic Ocean, these Berber populations have experienced severe geographical and ecological pressures, but had limited contacts with foreigners since the beginning of the historical era.

For a period of 12 centuries (11th century BC to 146 AD), the Phoenicians maintained contact through trade in the harbors of eastern Algeria, but established few permanent settlements. Eastern Algeria experienced less contact and fewer invasions than in northern Tunisia. For the next five centuries, the Romans appeared in a large part of the eastern area (Lfriquya), but were limited to farm management, administration, military control, and partial Christianization of some Berber groups. In the 5th century AD, the Romans were replaced by Vandals and Byzantines, few in number, but in the same eastern coastal areas and towns. The severe poverty that the invaders caused in such rich regions of Tunisia and Algeria reinforced the strong familial and tribal structure of the Berbers, and precipitated numerous revolts. In the 7th century, the Arabs (bringing their Islamic religion, language, and prodigious culture) expanded to most of the populations of Maghrib, and drove the remaining Berber-speaking populations (their opposition) into mountainous retreats. There is no precise information on the demographic and genetic impact of the newcomers from the Arabian peninsula, but given the persistence of their political influence in the Maghrib region, it must have been considerable, at least in the largest cities of that time. This region experienced one century of peace (from end of the 11th century to the end of the 12th century) when the Empires of Almoravids and Almohads merged. After 1159 AD, both Berber dynasties, through their conquests, consolidated the first autochthonous power in the entire Maghrib region. The Arab empire generated in Maghrib a constant partition between eastern political dynasties (Aghlabids, Fatimids, Zirids, Hammadids, and, after the 13th century, Hafsids) and western dynasties (Idrissids, Kharidjits, and later Marinids). Such a partition was accompanied by permanent conflicts and wars that could have played a significant role in the patterns of gene flow among the Maghribin populations distributed over 3,000 km. As stated by Larroui (1970), “Even if homogeneity (political and cultural) has been accomplished in the whole of Maghreb by the Arabic presence and power, a permanent tendency to decentralization and conflicts is also operating and should be explained by the social structure of historical Maghreb.” Such population dynamics were witnessed during the Ottoman period in Algeria from 1515–1830 AD under the administration and military control of a series of 63 suc-

TABLE 1. Distribution of population sample by region and by red cell polymorphism tested

Region	ABO	RH	Kell	Duffy	Kidd	MNSs	Lu	P	Xg
1	563	560	423	561	390	559	547	570	64
2	182	176	142	180	136	179	171	182	23
3	507	475	394	507	389	506	498	514	82
4	347	315	302	295	304	341	353	351	78
5	999	973	880	993	879	986	971	1,009	178
6	825	815	665	832	642	826	801	825	139
7	275	283	200	283	188	282	268	284	35
8	469	464	372	470	354	442	432	469	56
9	183	169	138	182	139	180	171	181	19
10	94	91	74	94	74	92	91	91	14
Total	4,444	4,321	3,590	4,397	3,495	4,393	4,303	4,484	688

cessive Beyler-Beys, Pashas, Aghas, and Deys up to the time of French colonization.

MATERIALS AND METHODS

In 1980, the Blood Transfusion Center of the Algerian Army took part in this Franco-Algerian scientific collaboration in which blood samples were collected from 4,444 young male recruits from all geographical regions of Algeria. Aireche (1987), using blood group frequency data, reported on an initial analysis that failed to demonstrate clear genetic patterning in Algeria. In this study, we reanalyzed the data by compiling the original sample into regions (based on ecology, history, and geography), and then searched for patterns of observable genetic and geographic variation.

Population samples

Aireche (1987) subdivided the original sample of volunteers by Wilayas of origin (see Fig. 1; thin lines denote the boundaries of Wilayas), the basic administrative units of Algeria. We combined the contiguous Wilayas into clusters that shared a common history and ecology, in order to avoid fragmentation of the total sample into a series of minuscule administrative groupings. In addition, a more comprehensive historical and ecological representation of Algeria could be achieved on a regional level.

This subdivision of Algeria into 10 regions is shown in Figure 1 (bold lines indicate boundaries of regions, while thin lines indicate the Wilayas of origin). The following is a listing of regions distributed from west to east:

- Region 1 (Sidi bel Abbes; Siba) was constructed by the compilation of gene frequencies from five coastal Wilayas situated around the historical town of Tlemcen and the economically important harbor of Oran.

- Region 2 (Saida; Said) is a western high plain consisting of two large Wilayas, where the market towns of Saida and Mecheria play a central role in pastoral production.
- Region 3 (Miliana; Mili) includes three Wilayas of the coastal Mitidja rich plain, developed by French colonization around the towns of Blida, Medea and Miliana.
- Region 4 (Algiers; Alge) is the central main harbor and capital of the country; it has become a large melting pot of diverse populations.
- Region 5 (Bejaia; Beja) includes the Berber mountain Kabylia and the sizable towns of Tizi-Ouzou and the harbor of Bejaia. This region contains four Wilayas.
- Region 6 (Constantine; Cons) is an eastern rich, high plain adjacent to the Tunisian border, whose main towns are Setif, Constantine, and Guelma (three Wilayas are included).
- Region 7 (Annaba; Anna) contains two Wilayas on the extreme eastern coast of the Mediterranean Sea, including the harbors of Skikda and Annaba.
- Region 8 (Ain Beida; Aibe) includes three Wilayas. The geographic center of this region is Ain Beida, surrounded by plateaus and mountains (Aures) to the south of Constantine, and is geographically proximal to Tunisia.
- Region 9 (Msila; Msil) includes the central continuation of region 8 up to the fringe of the Sahara with the Biskra oases (two Wilayas).
- Region 10 (Ain Salah; Aisa) is a compilation of all six arid Wilayas of the Atlas plateau around the town of Djelfa, to the Sahara up to Tamanrasset.

Table 1 provides the distribution of our population samples (by region and blood markers),

given this new territorial pattern. The maximum number of recruits tested was for the P blood group, at 4,484, followed by the ABO blood group system, with a total of 4,444. Smaller subsamples were phenotyped for systems such as Xg.

Statistical analysis

Most samples collected from the Algerian soldiers/volunteers were tested for nine red blood cell polymorphic systems: ABO, rhesus, Kell, Duffy, Kidd, Lutheran, MNSs, P, and Xg. Allelic frequencies (first computed by individual Wilayas) were combined and recompiled into “regions” clusters. Allelic frequencies for Wilayas were computed either by the maximum likelihood method or by direct gene counting for the 31 alleles and nine systems (Aireche et al., 1982; Aireche and Benabadji, 1986; Aireche, 1987). The following loci and alleles were tested on the samples:

- ABO: O, A, and B;
- Rhesus: RH*RI, R2, RO, Rz, r, r', r'', and ry;
- Kell: K and k;
- Duffy: FY*A, B, and Y;
- Sutter: JK*A and B;
- Lutheran: LU*A, B, and LU;
- MNS: MS, Ms, NS, Ns, MSu, and NSu;
- P system: P1 and P2; and
- Xg system: XG*1 and XG*0.

However, given the interdependence of alleles in each polymorphic system, only 16 alleles were used for one of the analyses (Table 2), and 10 for all others (Tables 2 and 3).

Analytical methods

Based on the polymorphic blood group data, the following population structural analyses were performed on the compiled regional data:

1. In the F_{ST} of Wright (1931), the average F_{ST} value was computed, based on 16 allelic frequencies from 10 regions (see Table 2).
2. R-matrix (Harpending and Jenkins, 1973), a principal-component analysis of squared Euclidian distances (Harpending and Rogers, 1984), was based on normalized gene frequencies (Fig. 2).
3. Heterozygosity vs. distance from the centroid (Harpending and Ward, 1982) involved a regression analysis of mean per locus heterozygosity (het) on the distance from the centroid of distribution of gene frequencies per region (r_{ij}). Populations that fall below

TABLE 2. Algerian gene frequencies ($\times 1,000$) of 16 alleles for 10 regions

Region	RH*RI	RH*R2	RH*RO	RH*Y	ABO*A	ABO*O	MNS*MS	MNS*Ms	MNS*NS	K	JK*B	Fy*a	Fy*b	LA	Xg0	P*2
1	3814	0912	2331	2784	1883	6895	2135	2855	0799	0461	4260	2790	4748	0138	3906	4168
2	3874	0836	2377	2688	1960	6611	2156	2856	1015	0528	4919	2726	4251	0149	6087	4258
3	3933	0917	2533	2397	2510	6314	2132	2820	0636	0482	3895	2613	4565	0159	3415	4138
4	4385	0740	1975	2594	2165	6561	1985	2825	0875	0414	4391	2562	4319	0217	3205	4134
5	4373	0746	1964	2768	2222	6509	1894	2753	0975	0523	4499	2966	5299	0185	3820	4396
6	4251	0963	2256	2200	1957	6800	2040	3277	0698	0566	4354	2749	4948	0156	3885	4341
7	4522	0832	1962	2504	2276	6398	2115	2910	0559	0475	4202	2171	4185	0180	5429	4279
8	4378	1064	2375	1758	2098	6737	1759	2969	0671	0376	4138	2927	5031	0213	2857	4381
9	4115	1057	1908	2386	1984	6905	1690	3544	0622	0616	4101	2579	4209	0149	3684	4701
10	3592	1499	2828	2082	1373	7255	1817	3671	0513	0338	3446	2361	3004	0138	500	3780

TABLE 3. Algerian gene frequencies of 10 more alleles for 10 regions

Region	RH*RZ	RH*r'	RH*r''	ABO*B	MNS*Ns	MSu	Nsu	Fy	Lu*b	Lu
1	0.0026	0.0132	0.0000	0.1222	0.3899	0.0321	0.0000	0.2461	0.9367	0.0495
2	0.0073	0.0076	0.0074	0.1429	0.3698	0.0175	0.0000	0.3022	0.9831	0.0020
3	0.0007	0.0137	0.0077	0.1176	0.4179	0.0233	0.0000	0.2821	0.9468	0.0374
4	0.0113	0.0000	0.0192	0.1273	0.3759	0.0000	0.0556	0.3119	0.9783	0.0000
5	0.0028	0.0071	0.0050	0.1269	0.4264	0.0000	0.0114	0.1735	0.9461	0.0354
6	0.0000	0.0305	0.0024	0.1243	0.3984	0.0000	0.0000	0.2303	0.9606	0.0239
7	0.0009	0.0153	0.0018	0.1326	0.4130	0.0000	0.0286	0.3644	0.9280	0.0540
8	0.0000	0.0426	0.0000	0.1166	0.4492	0.0110	0.0000	0.2042	0.9787	0.0000
9	0.0074	0.0334	0.0125	0.1111	0.4145	0.0000	0.0000	0.3212	0.9831	0.0020
10	0.0000	0.0000	0.0000	0.1373	0.3439	0.0322	0.0238	0.4635	0.8773	0.1089

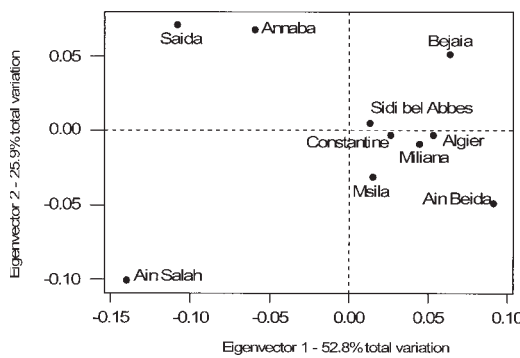


Fig. 2. Principal-components analysis of an R-matrix, based on gene frequencies of the 10 Regions of Algeria.

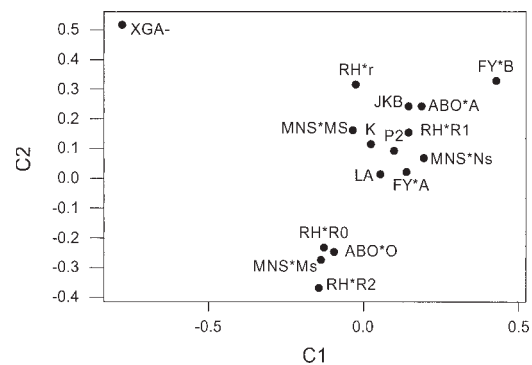


Fig. 3. Plot of the alleles contributing to the genetic relationships shown in Figure 2 of the 10 Regions of Algeria.

the theoretical regression line [mean het (1 - r_{ii})] experience significantly more drift than those proximal to the line. Populations that experienced significant gene flow tend to exhibit high levels of heterozygosity, and low r_{ii} are found well above the theoretical regression (Fig. 4).

4. Phylogenetic trees: We constructed trees based on two different types of hierarchical classifications (Fig. 5): Neighbor-joining (Saitou and Nei, 1987) trees were constructed from a matrix consisting of Reynolds distances (Reynolds et al., 1983) computed on 16 allelic frequencies, with corrections for differences in population sample sizes. Unweighted pair-group average arithmetic (UPGMA) trees were constructed from a matrix of chi-square distances (Jacquard, 1974), and computed using 16 allelic frequencies.
5. Mantel tests were used to compare the genetic distance matrix with a great-circle geographic distance matrix (distances based on the geographical center of each region).
6. Spatial autocorrelation analysis was performed for the purpose of evaluating pat-

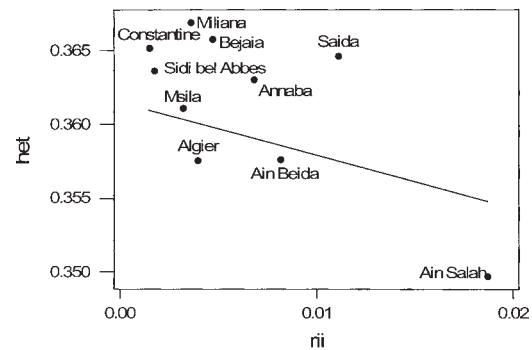


Fig. 4. Plot of a theoretical regression line between mean per locus heterozygosity (het) and distance from the centroid of distribution (r_{ii}) in 10 Regions of Algeria.

terns of biological variation, and to test the isolation-by-distance model in the large geographical Algerian area (Sokal, 1978). In this study, 16 allelic frequencies were used to elucidate the patterning of genes in geographic and cultural regions of Algeria. Latitude and longitude were estimated from the centroid of each region. Spatial

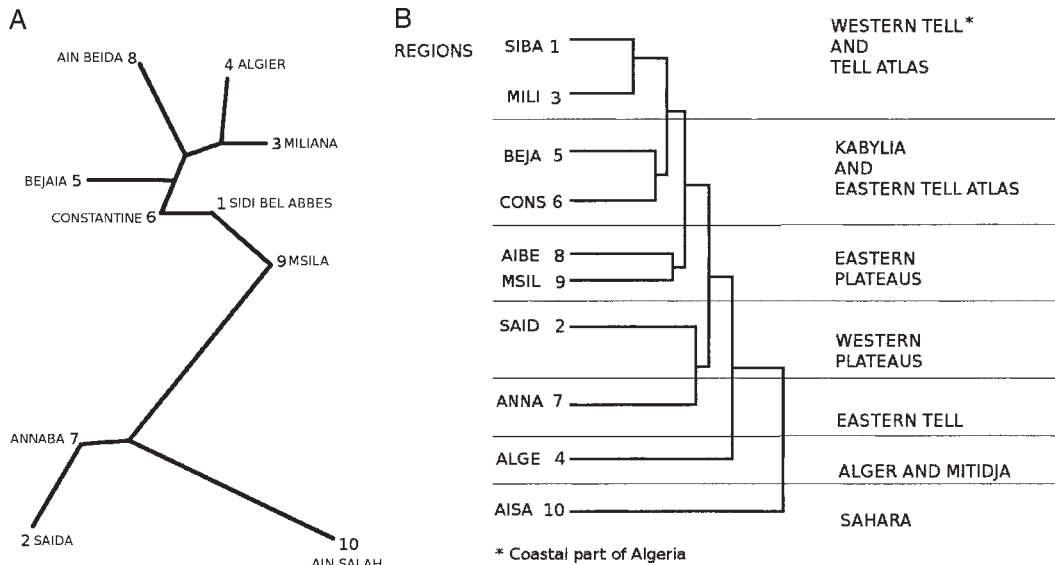


Fig. 5. **a:** Unrooted neighbor-joining tree of the hierarchical relationships between the 10 Regions of Algeria. Only bootstraps greater than 50% are incorporated in the plot. **b:** Unweighted average phylogenetic distance tree (UPGMA) for 10 Regions of Algeria.

autocorrelation analysis and correlogram construction were based on the procedures of Cliff and Ord (1981). Moran's I correlograms were constructed using standardized z-scores and spatial lags.

RESULTS AND DISCUSSION

Wright's F_{ST} values by locus indicate the weak level of heterogeneity of the total Algerian population for each genetic polymorphism tested (Table 4). The overall F_{ST} value, 0.0063 for the 10 regions, indicates a low level of genetic differentiation among regions. Only the distribution of alleles of the XG erythrocytic blood group ($F_{ST} = 0.04$) exhibits a significant difference for eastern region 2 (Said) and western region 7 (Anna). The reason for this statistically significant difference in the two regions is unclear, given that there is no known contact, historic or economic, between these areas. However, both regions exhibit relatively high heterozygosities and medium levels of r_{ii} , suggesting some gene flow (possibly from a common source) and the action of stochastic processes (Fig. 4).

The R-matrix PCA reveals some genetic differences among Algerian regions (Fig. 2). The first eigenvector, accounting for 52.8% of the total variation in the sample, confirms a clus-

ter of seven regions, while the second eigenvector (25.9% of the variation) separates region 2 (Said) and region 7 (Anna) from the others. Region 10, the Ain Salah grouping, consists of the Algerian Saharan Wilayas. This cluster of one region is clearly separated by its high frequencies of RH*RO, RH*R2, MNS*Ms, and ABO*O, as indicated by the PCA of the S-matrix dispersal of alleles used in the analysis (Fig. 3). Saida and Annaba are distinguished by their high frequency of XGA-

Figure 4 provides a plot of heterozygosity vs. distance from the centroid of distribution. Several regions fall proximal to the theoretical (prediction) regression line. The difference in gene frequencies for region 10 (Ain Salah-Aisa) revealed by the PCA can be accounted for by the differentiation from sub-Saharan Africa to the Mediterranean Sea and the observed morphological presence of black African admixture, still evident in the Sahara but diminishing toward coastal areas. This observed dispersion of regions can also be explained by genetic drift, given the limited contacts of southern with northern populations. Region 2 (Saida) appears to have experienced considerable gene flow from the south, perhaps through trade and slavery relationships with sub-Saharan populations, the importance of which, especially for slavery, must be measured and confirmed.

TABLE 4. F_{ST} values for 16 alleles of the red cell polymorphisms

Alleles	F_{ST} by allele
RH*R1	0.003497
RH*R2	0.005095
RH*R0	0.004643
RH*r	0.005267
ABO*A	0.005011
ABO*B	0.003180
XG*0	0.039880
LA	0.000460
MNS*MS	0.001682
MNS*Ms	0.004453
MNS*NS	0.003568
K	0.001410
JK*B	0.003297
FY*A	0.002765
FY*B	0.001491
P2	0.002062
Overall	0.006324

Out of these three analyses, no single genetic pattern emerges. However, the trees obtained by hierarchical classification bring more interpretable results by demonstrating some slight contrasts between regions. The neighbor-joining tree (Fig. 5a) reveals an evident genetic structure in Algerian populations, even if differences from region to region are limited. We can observe in the R-matrix plot that region 4 (Alge) and region 10 (Aisa) cluster far apart. Saharan populations are genetically closer to their sub-Saharan black neighbors, but Algiers is a large melting pot of local Afro-Mediterranean and Euro-Mediterranean invaders. In the two trees, western region 2 (Said) clusters closely to the Saharan region, given the historical and economical factors we mentioned above. Eastern region 7 (Anna) again displays a genetic similarity with western region 2. This could be explained by the high XG0 gene frequency in both samples, but it could be purely stochastic, as mentioned in the comment on F_{ST} results. Regions 5 (Beja) and 6 (Cons) cluster together, given the relationships they have maintained for centuries or millennia in their mountains and plateaus as a population of Berber-speaking tribes. Both regions, and especially region 6, appear close to region 8 (Aibe) and region 9 (Msil) as a common part of a prehistoric and historic eastern settlement. The western populations of regions 1 (Siba) and 3 (Mili) cluster together with a homogeneity due probably to their particular coastal environment and political history, especially in the Ottoman period.

The UPGMA tree (Fig. 5) clearly confirms this difference in the genetic structure of the eastern vs. western gene pool on the one hand,

and the Mediterranean coastal areas vs. the Atlas high plains and plateaus on the other, and the particular positions of Algiers in the north and the Sahara in the south. A geographic projection of the UPGMA tree depicts the proposed genetic structure of Algeria. (Fig. 6).

The Mantel test, a comparison of matrices based on geographic distances (measured from the center of each region) and genetic distances, indicates no modification of gene frequencies as a result of geographic separation: $R = 0.4659$, $R^2 = 0.2171$, $P = 0.1880$ (NS).

However, since there are too few data points (10 regions) and insufficient statistical power, the results prove not to be statistically significant. Figure 7 is a plot of the z-scores of three alleles in relation to the lag distances. All three alleles, while showing apparent patterns of spatial autocorrelation and significant lag values, fail to exhibit overall statistical significance. Although allelic frequency clines were interpreted by Bosch et al. (1997) for North Africa from synthetic gene maps (i.e., maps of principal-component scores based on correlations of interpolated surfaces), caution must be exercised (Sokal et al., 1999). In criticism of the use of synthetic gene maps, Sokal et al. (1999, p. 1) wrote, "These maps are subject to large errors and apparent geographic trends may be detected in spatially random data."

CONCLUSIONS

The genetic homogeneity of this Algerian population, as demonstrated by statistical analyses computed on this sample and set of markers, is suggestive of the long history of autochthonous populations spreading into the Maghrib region at the end of Paleolithic (Epi-paleolithic) and early Neolithic times, and experiencing gene flow from invading armies and the sub-Saharan slave trade. Apparently, considerable gene flow seems to have resulted from the movement of traders or invaders into this large territory. The role of gene flow, rather than genetic drift, is dramatically shown in the heterozygosity vs. r_{ii} plot that exhibits high levels of heterozygosity and low r_{ii} in 7 of 10 regions. Only Ain Salah appears to have undergone genetic microdifferentiation due to stochastic processes.

The slight differences demonstrated by the tree reconstructions can be observed on the two main axes of the observed structure. The first axis is oriented east-west, and the most likely clustering factor is the territorial organization and numerous conflicts of local politi-

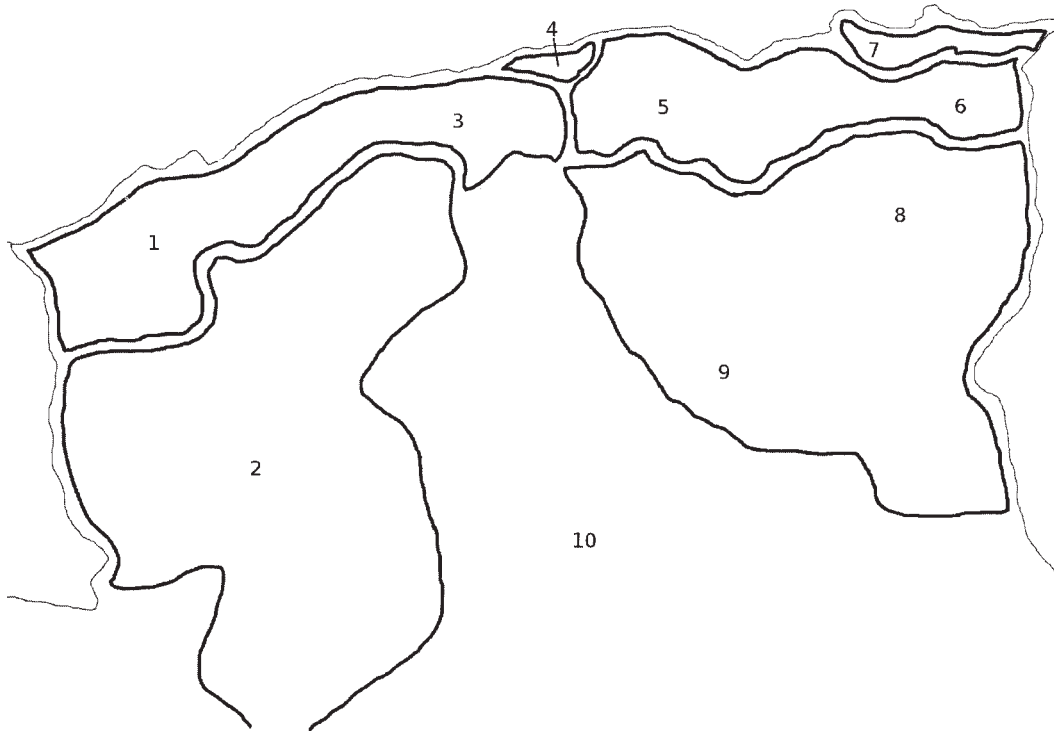


Fig. 6. Map projection of hierachical tree for Regions of Algeria.

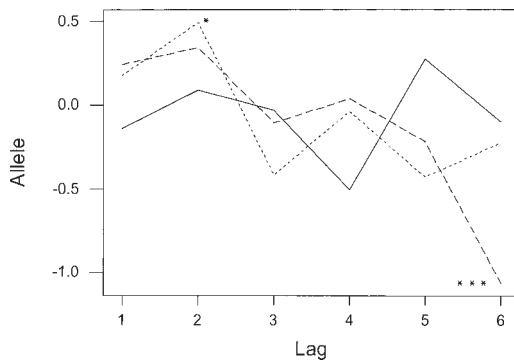


Fig. 7. Spatial autocorrelation of three alleles in Algeria. ABO*B, solid line ($P = 0.166-0.9569$); RH*R1, large-dashed line ($P = 0.00006^{***}$ (lag 6) -0.9847); P2, small-dashed line ($P = 0.03082^*$ (lag 2) -0.6392).

cal and military powers (see above); two more (minor) factors can be seen for the eastern part of Algeria. On the one hand, a genetic contribution (gene flow) from some other Mediterranean areas, such as Italy, Greece, or the Near East, is perhaps related to the allelic frequency cline described by Bosch et al. (1997),

but on the other hand, genetic drift operated in the isolated and highly endogamous Berber-speaking groups of the mountains of Kabylia (region 5) and Aures (region 8) (Smith, 1974). However, caution must be exercised in the interpretation of apparent clines that are based on synthetic gene maps, which are more impressionistic and fail to statistically test hypotheses (Sokal et al., 1999).

On the north-south axis, the observed variation in latitude suggests a genetic cline of over 3,000 km from sub-Saharan black Africans to Afro-Mediterraneans. This apparent cline places the coastal Berber people closer to Europeans of the northern bank of the Mediterranean, while on the basis of observations of morphological variation, the southern Berbers more closely resemble black Africans. This north-south genetic cline parallels the extreme changes of climate and environment, stretching from humid tropical Africa, to the searing heat and aridity of the Sahara, to the dry plateaus and plains of the Atlas region, to the temperate Mediterranean seashore.

The relatively small size of our sample may not represent the total Algerian population. A

limited number of genetic parameters was used in our study, but due to the compilation of allelic frequencies of Algeria by virtual regions, we are able to propose an approach to the global genetic structure of this country, as pictured in a provisional map (Fig. 6).

This preliminary analysis should be considered complementary to the synthetic studies of Mourant et al. (1975), Cavalli-Sforza et al. (1994), and Bosch et al. (1997), as a guide for further research on the genetics of Algeria, its North African neighbors and “cousins,” and surrounding areas.

LITERATURE CITED

- Aireche D. 1987. Polymorphisme erythrocytaire dans la population algérienne. Thèse de Doctorat en Sciences Médicales (Pharmacie), INESM (Institut National de l'Éducation en Sciences Médicales), Alger. 204 p.
- Aireche D, Gueguen A, Golmard JM, Benabadi M. 1982. Détermination des fréquences géniques dans le système rhesus en Algérie. *Rev Fr Trans Immunol Hematol* 25: 383.
- Aireche H, Benabadi M. 1986. Le système Duffy en Algérie. *Rev Fr Trans Immunol Hematol* 29:121.
- Arredi B, Poloni ES, Paracchini S, Zerjal T, Fathallah DM, Pascali VL, Novelletto A, Tyler-Smith C. 2004. A predominantly Neolithic origin for Y-chromosomal DNA variation in North Africa. *Am J Hum Genet* 75:338–345.
- Benabadi M, Benlatrache C, Merad F, Suaudeau C, Benmoussa M, Kornprobst G, Cladel G. 1977. Le déficit en glucose-6-phosphate déshydrogénase en Algérie. In *Semaine de Hôpitaux*, Paris 16(53):899–904.
- Benabadi M, Ruffie J, Lefevre-Witier P. 1969. Les grandes données de l'hémotypologie Sahariennes. In: M. Benabadi and A Boujnah. *Biologie des populations Sahariennes*. Actes du Colloque International. Alger: Institut National de la Santé Publique, p 51–74.
- Benabadi M, Chamla MC. 1971. Les groupes sanguins ABO et rhesus des Algériens. *Anthropologie* 75:427–442.
- Benabadi M, Ruffie J, Lefevre-Witier P. 1969. Les grandes données de l'hémotypologie Saharienne. In: *Biologie des populations Sahariennes—actes du colloque international*. Alger: Institut National de la Santé Publique. p 51–74.
- Bosch E, Calafell F, Perez-Lezaun A, Comas D, Mateu F, Bertranpetit J. 1997. Population history of North Africa: evidence from classical genetic markers. *Hum Biol* 69: 295–311.
- Bosch E, Calafell F, Perez-Lezaun A, Comas D, Izaabel H, Akhayat O, Sefiani A, Hariti GG, Dugoujon JM, Bertranpetit J. 2000. Y chromosome STR haplotypes in four populations from northwest Africa. *Int J Legal Med* 114: 36–40.
- Camps G. 1974. Les civilisations préhistoriques de l'Afrique du Nord et du Sahara. Paris: Doin.
- Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. *The history and geography of human genes*. Princeton: Princeton University Press.
- Constans J, Lefevre-Witier P, Richard P, Jaejer G. 1980. Gc (vitamin D binding protein) subtype polymorphism and variant distribution among Sahara, Middle East, and black African populations. *Am J Phys Anthropol* 52: 435–441.
- Felsenstein J. 1973. Maximum likelihood estimation of evolutionary trees from continuous characters. *Am J Hum Genet* 25:471–492.
- Harpending H, Jenkins T. 1973. Genetic distance among southern African populations. In: Crawford MH, Workman P, editors. *Methods and theories of anthropological genetics*. Albuquerque: University of New Mexico Press.
- Harpending H, Rogers A. 1984. *ANTANA: a package for multivariate data analysis*. Distributed by the authors.
- Harpending H, Ward RH. 1982. Chemical systematics and human populations. In: MH Nitecki, editor. *Biological Aspects of Evolutionary Biology*. Chicago: University of Chicago Press, p. 177–199.
- Ivanova R, Astrinidis A, Djoulah S, Lepage V, Wijnen E, Hors J, Charron D. 1999. Mitochondrial DNA polymorphisms of a west Algerian population (Oran region). *Biomed Pharmacother* 53:386–392.
- Jacquard A. 1974. *The Genetic Structure of Populations*. NY: Springer-Verlag, p 420–492.
- Kossovitch N. 1953. *Antropologie et groupes sanguins des populations du Maroc*. Paris: Masson.
- Larroui A. 1970. *L'histoire du Maghreb: un essai de synthèse—petite collection*. Paris: Maspero.
- Lefevre-Witier P. 1972. Note sur l'hétérogénéité biologique des Touaregs In: Actes du colloque “Étude des Populations Sahariennes”—Abbaye de Senanque (France)—15–18 Octobre 1971.
- Lefevre-Witier P. 1974. Un isolat du sud-Sahara: les Touaregs Kel Kummer: structures génétiques des systèmes sanguins erythrocytaires et sériques. *Population*. Paris. INED (Institute of Demographic Studies). 3:518–527.
- Lefevre-Witier P. 1982. Ecological and biological structure of pastoral Isseqqamaren Tuareg. In: Crawford MH, Mielke J, editors. *Current development of anthropological genetics. Volume 2: population structure and ecology*. New York. Plenum Press. p 94–125.
- Lefevre-Witier P. 1996. Ideles du Hoggar—biologie et écologie d'une communauté Saharienne. Paris: CNRS-Éditions “Anthropologie.” p 272.
- Lefevre-Witier P, Vergnes H. 1977. Genetic structure of Ideles (Ahaggar - Algeria). In: Meier RJ, Otten C, Abdel-Hameed F, editors. *Evolutionary models and studies in human diversity*. Chicago: Mouton-Aldine. p 255–277.
- Lefevre-Witier P, Vergnes H, Clerc A. 1968. Study of certain enzyme characters of populations living in North Sahara. In: *Proceedings of the 8th LCAES Science Council of Japan—Tokyo*. p 188–191.
- Mourant AE, Kopec A, Sobczak K. 1975. *The distribution of the human blood groups and other polymorphisms*. Oxford monographs on medical genetics. Oxford: Oxford University Press. p 84–88.
- Reynolds J, Weir B, Cockerham CC. 1983. Estimation of the coancestry coefficient: basis for a short term genetic distance. *Genetics* 105:767–779.
- Ruffie J, Lefevre-Witier P. 1972. *Ergebnisse serologischer Fungen in der Sahara*. In: *Die Sahara und ihre Ralldgebiete—Band II: human Geographie*. Munich: Weltforum Verlag. p 255–263.
- Saitou N, Nei M. 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425.
- Smith CAB. 1974. Measures of homozygosity and inbreeding in populations. *Ann Hum Genet* 37:377.
- Sokal RR, Jacquez GM, Wooten MC. 1989. Spatial autocorrelation analysis of migration and selection. *Genetics* 121(4): 845–855.
- Sokal RR, Oden NL, Thomson B. 1999. A problem with synthetic maps. *Hum Biol* 71:1–13.
- Thompson EA. 1975. *Human evolutionary trees*. Cambridge. Cambridge University Press.
- Wright S. 1931. Evolution of Mendelian populations. *Genetics* 16:97–101.