Original Research Article

Genetic Position of Valencia (Spain) in the Mediterranean Basin According to *Alu* Insertions

S. GARCÍA-OBREGÓN,[1] M.A. ALFONSO-SÁNCHEZ,[1] A.M. PÉREZ-MIRANDA,[1] C. VIDALES,[2] D. ARROYO,[3] AND J.A. PEÑA[1]*
[1]*Departamento de Genética y Antropología Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco, Bilbao, Spain*
[2]*Unidad de Genética Molecular, Policlínica Gipuzkoa, San Sebastián, Spain*
[3]*Progenie Molecular, Valencia, Spain*

ABSTRACT   In this work, eight human-specific *Alu* insertion polymorphisms (ACE, TPA25, PV92, APO, FXIIIB, D1, A25, and B65) were typed in 106 unrelated healthy individuals born in the province of Valencia (Spain), with the aim of analyzing the genetic relationships between this region of the Iberian Peninsula and other Mediterranean populations. To that end, *Alu* data on Eastern European, Western European, and North African populations were compiled from previous studies. The genetic information was stressed by means of genetic distances (R matrix method), nonmetric multidimensional scaling (MDS) and analyses of molecular variance (AMOVA). In Valencia, the most common *Alu* insertion was APO (0.940), and the least frequent was A25 (0.104). The average gene diversity (GD) computed for the sample examined was comparatively high (0.382). The insertion frequencies estimated for the eight *Alu* markers were very similar to the mean frequencies calculated for the whole set of populations included in the study, suggesting the hybrid nature of the Valencia's gene pool. MDS and AMOVA results generated from *Alu* data reveal that the Mediterranean has acted as a strong genetic boundary between the north (Europe) and the south (Northern Africa), resulting in significant gene diversity between the populations of the two regions. Restricted exclusively to the European scope, we suggest the possibility that the Mediterranean could have also acted as a migratory passageway, propitiating the dissemination of cultures and genes between the east and west of Europe and giving rise to some homogenization of gene frequencies among coastal dwelling populations. Am. J. Hum. Biol. 18:187–195, 2006.   © 2006 Wiley-Liss, Inc.

Among the more useful genetic markers in molecular studies from an evolutionary viewpoint are polymorphic *Alu* insertions, which are sequences of approximately 300 base pairs (bp) in length ancestrally originated from the 7SL RNA gene by retrotransposition. *Alu* elements are the most abundant short interspersed elements (SINEs) in the human genome: there are more than one million copies, which represents around 10% of the human genome. These *Alu* elements are recent in origin and propagation in evolutionary terms, coinciding with the radiation of the primates 65 million years ago (reviewed in Batzer and Deininger, 2002).

Alu sequences have been classed into subfamilies of different genetic ages based on hierarchical series of mutations shared by subfamily members (Batzer et al., 1996). Most of the *Alu* elements that have inserted more recently into the human genome belongs to the subfamilies known as "young" *Alu* (Y, Yc1, Yc2, Ya5, Ya5a2, Ya8, Yb8, and Yb9), and several of them are exclusive to our species. The distribution of these elements varies in different human population groups (Stoneking et al., 1997).

Several characteristics of polymorphic *Alu* insertions make them valuable markers in

phylogenetic analyses. Thus, for instance, the absence of the insertion makes possible the knowledge of the ancestral state, which constitutes an advantageous attribute in investigations aimed at unraveling the demographic, genetic, and evolutionary history of the human species. Furthermore, alleles are identical by descent, so it is highly unlikely that the same *Alu* insertion could occur more than once independently at the same locus. This means that polymorphic *Alu* insertions reflect unique evolutionary events (Batzer et al., 1994).

The aim of the present study is to analyze the genetic variability of a Mediterranean coastal population from Spain (Valencia) based on the polymorphism of eight *Alu* insertions (ACE, TPA25, PV92, APO, FXIIIB, D1, A25, and B65). Because studies on DNA polymorphisms in Valencia population have been rather scanty, this Iberian region remains relatively unknown from the genetic point of view. In addition, several of the investigations carried out to date do not go beyond showing the allelic frequencies of the molecular markers analyzed. Thus, frequencies of both autosomal short tandem repeats, STRs (Aler et al., 2001a; Tomas et al., 2001a,b) and of Y-chromosomal STRs (Aler et al., 2001b) have been published in the last years. More recently, the findings of a study centered on mitochondrial DNA diversity (Picornell et al., 2005) revealed a lack of genetic affinity among maternal lineages of Valencia and some African populations.

Historically, the Mediterranean Sea has played a key role as a trade route for western civilization, and it has been the setting for numerous conquests, colonizations, and demographic expansions. The strategic position of Valencia could have helped make this region an important pole of attraction for the first bearers of Neolithic culture who crossed the Mediterranean and reached the Iberian Peninsula by sea around 5600 BC (Zilhao, 2001). Subsequently the region was inhabited by various peoples, some of them originally from elsewhere in the Mediterranean. Therefore, a high genetic diversity can be expected among the current inhabitants of the Valencia region. Our findings on *Alu* diversity in Valencia are further examined within the context of other Eastern European, Western European, and North African *Alu* data. This integrative approach provides a thorough comparison of *Alu* affinity levels and genetic associations from the Valencian collection with those of other populations in the Mediterranean basin, with a view to providing additional genetic information that could help to understand the populating process of the Mediterranean region.

MATERIALS AND METHODS

The Autonomous Community of Valencia is located on the Mediterranean coast (Fig. 1) in the east of the Iberian Peninsula. It comprises the provinces of Alicante, Castellón, and Valencia, stretching over 474 km of coastline and occupying a total surface area of 23,305 km^2. The current population density is more than twice the average for Spain at 178.6 inhabitants/km^2.

With a view to characterizing the genetic variability of the targeted region, eight *Alu* insertions (ACE, TPA25, PV92, APO, FXIIIB, D1, A25, and B65) were typed in a set of 106 unrelated and healthy individuals from the province of Valencia (Spain). The birthplace of an individual, surveyed back to the third generation (the grandparents), was the criterion to decide whether a person was native or not from Valencia. In our study, the term "Valencia" includes not only the provincial capital (Valencia city) but also other surrounding areas belonging to Valencia province. Ethical guidelines for research with human beings were adhered to as stipulated by each of the institutions involved in the study.

Genomic DNA was extracted from bloodstains using a Qiagen kit (QIAmp DNA Micro Kit) and stored at $-20°$C. PCR reactions were performed in a final volume of 10 μl using 50 ng of DNA, PCR buffer (50 mM of KCl, 1.5 mM of MgCl$_2$, 10 mM of Tris-HCl, and 0.01% gelatin), 1.5 mM of MgCl$_2$, 0.1 mM of each dNTP, 1 μM of each primer and 0.04 U *Taq* polymerase in a Thermal Cycler Gene Amp PCR System 9700 (PerkinElmer, Norwalk, CT). Each sample was subjected to the following amplification conditions: a preliminary step of 94°C for 10 min, 30 subsequent cycles that consisted of denaturing at 94°C for 1 min, followed by annealing at different temperatures depending on the marker (Table 1) for 1 min 30 s and an extension of 1 min 30 s at 72°C, with a final elongation step of 72°C for 10 min. All the PCR reactions were directly electrophoresed in 1.5% agarose gels stained with ethidium bromide (0.5 μl/ml), viewed under UV light and documented using digital photography.

Allelic frequencies for the eight *Alu* loci typed in the Valencian collection, gene diversity, and polymorphism information content (Botstein et al., 1980) were calculated using version 3.0 of the Power Marker program (Liu
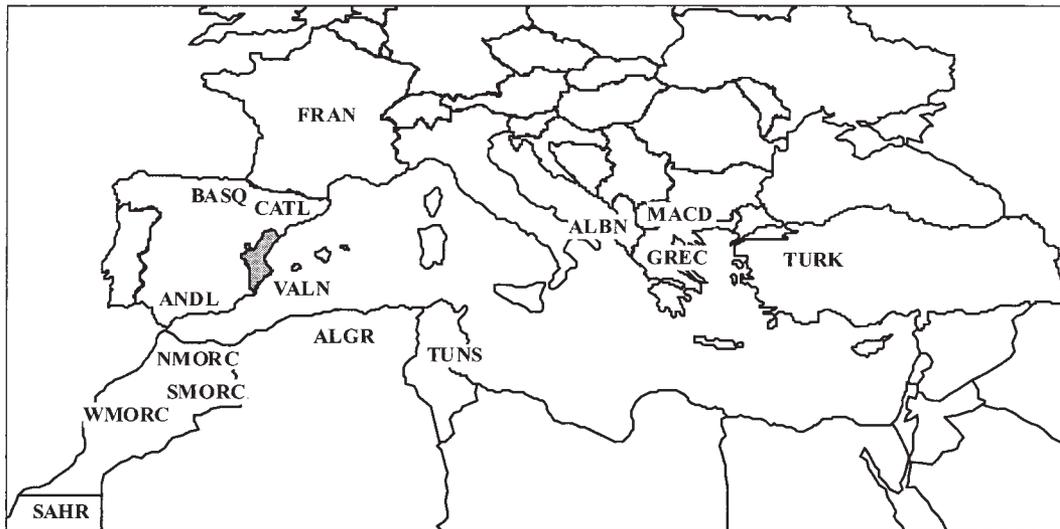
Fig. 1. Map showing the geographic location of Valencia (shaded area) and other populations included in the present study (Comas et al., 2000, 2004; Stoneking et al., 1997). Key: ANDL, Andalusia; ALBN, Albania; ALGR, Algeria; BASQ, Basque Country; CATL, Catalonia; FRAN, France; MACD, Former Yugoslav Republic of Macedonia; GREC, Greece; NMORC, North Morocco; SAHR, Sahara; SMORC, Southeast Morocco; TUNS, Tunisia; TURK, Turkey; VALN, Valencia; WMORC, West Morocco.

TABLE 1. *Primers, annealing temperature (AT), and chromosomal location for eight Alu markers*

Locus	5′-End primer	3′-End primer	AT (°C)	Chromosome
ACE	CTGGAGACCACTCCCATCCTTTCT	GATGTGGCCATCACATTCGTCAGAT	56	17
TPA25	GTAAGAGTTCCGTAACAGGACAGCT	CCCCACCCTAGGAGAACTTCTCTTT	63	8
PV92	AACTGGGAAAATTTGAAGAGAAAGT	TGAGTTCTCAACTCCTGTGTGTTAG	58	16
APO	AAGTGCTGTAGGCCATTTAGATTAG	AGTCTTCGATGACAGCGTATACAGA	47	11
FXIIIB	TCAACTCCATGAGATTTTCAGAAGT	CTGGAAAAAATGTATTCAGGTGAGT	58	1
D1	TGCTGATGCCCAGGGTTAGTAAA	TTTCTGCTATGCTCTTCCCTCTC	57	3
A25	CCACAAATAGGCTCATGTAGAAC	TATAATATGGCCTGGATTATACC	58	8
B65	ATATCCTAAAAGGGACACCA	AAAATTTATGGCATGCGTAT	58	11

and Muse, 2004). To test for Hardy–Weinberg equilibrium (HWE), a Fisher's exact probability test was conducted to estimate P values (Guo and Thompson, 1992) using Arlequin version 2.000 (Schneider et al., 2000).

The genetic associations and population affinities of the Valencian sample in a broader geographic context were analyzed by means of the R matrix method (Harpending and Jenkins, 1973). To that end, *Alu* data on Eastern European, Western European and North African populations were compiled from previously published works (Comas et al., 2000, 2004; Stoneking et al., 1997) (see Fig. 1). To represent genetic distances between pairs of populations computed through the R matrix in a two-dimensional genetic map, nonmetric multidimensional scaling (MDS) analysis (Kruskal,

1964) was performed using the SPSS (version 13.0, SPSS, Inc., Chicago, IL) statistical package. In order to ascertain the proportions of the genetic variance due to differences within and between populations, genetic variance was hierarchically apportioned according to geographic criteria through the analysis of molecular variance (AMOVA) (Excoffier et al., 1992) using the Arlequin program. In this statistical analysis, a permutation procedure allows testing the significance of the fixation indices F_{SC} and F_{CT}, which measure the relative contribution of the genetic variation among populations within groups, and among groups, respectively. Later, we established an overall test in order to check the statistical significance of F_{CT} values by combining the separate probability values for each locus through

TABLE 2. *Alu insertion frequencies with their standard errors (±SE), gene diversity (GD), and polymorphism information content (PIC) in a sample from Valencia province (Spain)*

Alu locus	$2N$[a]	Frequency	SE	GD	PIC
ACE	204	0.3873	±0.0342	0.4746	0.3620
TPA25	212	0.5566	±0.0308	0.4936	0.3718
PV92	194	0.2320	±0.0277	0.3563	0.2928
APO	200	0.9400	±0.0156	0.1128	0.1064
FXIIIB	208	0.4760	±0.0345	0.4988	0.3744
D1	208	0.3221	±0.0316	0.4367	0.3413
A25	202	0.1040	±0.0204	0.1863	0.1689
B65	204	0.5294	±0.0342	0.4983	0.3741

[a]$2N$ sample size in number of chromosomes typed.

the equation $\chi_{[2k]}^2 = -2\sum \ln p_i$, where k is the number of loci and p_i the separate probability value associated with the F_{CT} values for each i locus (Sokal and Rohlf, 1997). Finally, we used two methods to assess the congruence between geographic and genetic distances between populations. First, to obtain a consensus topogenetic map, the first two eigenvectors of the MDS were extracted and rotated to maximum congruence with geographic coordinates using methods described by Lalouel (1973). The second method employed was the matrix comparison test devised by Mantel (1967) and modified by Smouse et al. (1986), which can be used to compare distance and similarity or dissimilarity matrices, provided that the matrices are calculated from independent data sets (Dietz, 1983). This procedure generates a null distribution of correlation coefficients by randomly permuting the rows and columns of one of the compared matrices. Thus, empirical significance levels can be obtained for the correlation coefficient.

RESULTS

The *Alu* insertion frequencies for the eight loci typed in the sampled population are listed in Table 2. As can be noted, neither fixed insertions nor insertions absent from the population analyzed were found, i.e., all the *Alu* loci were polymorphic. Although this work is not primarily intended to produce a detailed account of the insertion frequencies, some salient features are noteworthy. Thus, for instance, APO was by far the most common element, with an insertion frequency of over 90% (0.940). Conversely, A25 was the least-frequent element at just over 10% (0.104). For the remaining *Alu* markers, intermediate insertion frequencies were observed, ranging from 0.232 (PV92) to 0.557 (TPA25). HWE was assessed by an exact test to calculate the

P value using the Markov-chain Monte Carlo method (Guo and Thompson, 1992). No significant departure from HWE expectations was detected in any of the eight *Alu* insertions analyzed.

The degree of genetic variability in the Valencian collection was assessed by computing gene diversity (GD) and polymorphism information content (PIC) for each locus (see Table 2). Bearing in mind that biallelic markers such as *Alu* insertion polymorphisms show diversity with a maximum of 0.5, both of these parameters presented relatively high values in five out of the eight loci typed in our study (FXIIIB, B65, TPA25, ACE, and D1), all of them characterized by insertion frequencies close to 50%. In these *Alu* markers, gene diversity values were all above 0.40, with a range of variation of between 0.437 (D1) and 0.499 (FXIIIB), while the extreme values of PIC ranged from 0.341 to 0.374 in the same *Alu* loci. On the other hand, the lowest genetic variability was observed in APO (GD, 0.113; PIC, 0.106) and A25 (GD, 0.186; PIC, 0.169), which exhibited the highest and the lowest insertion frequencies, respectively. As might be expected, the average gene diversity in the Valencian sample (considering the eight *Alu* typed) was 0.382, a relatively high value when compared with other Mediterranean populations analyzed previously (see Comas et al., 2004).

In order to assess the genetic position of Valencia within the Mediterranean context and to ascertain population affinities based on *Alu* diversity, we compiled *Alu* data from previously published papers (Comas et al., 2000, 2004; Stoneking et al., 1997). First, the R matrix method (Harpending and Jenkins, 1973) was applied to compute genetic distances between all pairs of populations (data not shown). Then a nonmetric multidimensional scaling
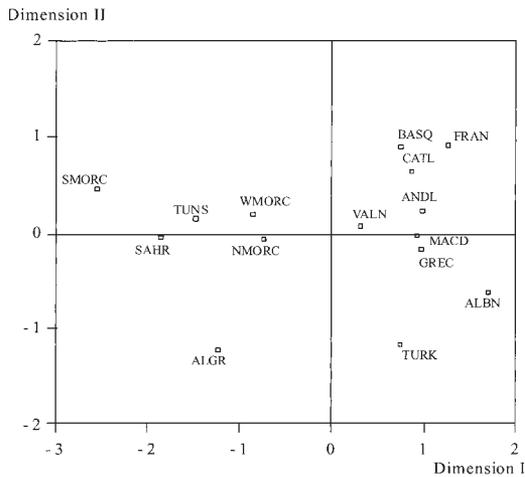
Fig. 2. Nonmetric multidimensional scaling (MDS) applied on *R* matrix based on eight *Alu* insertions to analyze genetic relationships among 15 Mediterranean populations. Key: ANDL, Andalusia; ALBN, Albania; ALGR, Algeria; BASQ, Basque Country; CATL, Catalonia; FRAN, France; MACD, Macedonia; GREC, Greece; NMORC, North Morocco; SAHR, Sahara; SMORC, Southeast Morocco; TUNS, Tunisia; TURK, Turkey; VALN, Valencia; WMORC, West Morocco.

(MDS) analysis was used to represent the data generated by the *R* matrix on a two-dimensional genetic map.

Figure 2 illustrates the results of nonmetrical multidimensional scaling (MDS) applied to the *R* matrix. Populations were grouped according to geography and this topology has proved highly robust from the statistical viewpoint, since the two-dimensional genetic map accounted for 95.03% of the total variance. In agreement with previous studies on human diversity using polymorphic *Alu* insertions (Stoneking et al., 1997; Watkins et al., 2001), African and non-African populations clearly segregated along dimension I (with a 76.9% of the total variance accounted for) in the MDS representation. Thus, all North African populations (Southeast Morocco, West Morocco, North Morocco, Sahara, Tunisia, and Algeria) were plotted in the negative segment of the first axis (dimension I). Among these populations, Southeast Morocco, Sahara, and Algeria seem to be the most genetically dissimilar because they occupy the most distant positions from the centroid. These North African populations are characterized by having the broadest ranges of insertion frequencies of all the populations included in this study. Specifi-

cally, the Southeast Morocco population has the highest insertion frequencies for A25 (0.235) and PV92 (0.398), and the lowest for B65 (0.510), while the Sahara population has the lowest frequencies for APO (0.836) and TPA25 (0.397). Likewise, the Algerian population has both the highest frequency for B65 (0.734) and the lowest for D1 (0.149) (see Comas et al., 2000).

On the other hand, the European populations are spread along dimension II: Western European populations (Andalusia, Basque Country, Catalonia, Valencia, and France) are clustered in the quadrant formed by the positive segments of dimensions I and II, whereas Eastern European populations (Albania, Macedonia, Greece, and Turkey) are concentrated in the negative segment of dimension II. This second axis accounted for 18.1% of the variance.

The central position of the study population (Valencia) on the two-dimensional genetic map is also worthy of attention; indeed, it is the closest sample to the centroid of the distribution. Consistent with this finding, the insertion frequencies estimated for the eight *Alu* typed in the Valencian sample are very similar to the mean frequencies calculated for the whole set of populations included in the study (see Tables 2 and 3). These results suggest that the Valencian gene pool is hybrid in nature, most probably because of a relatively high rate of mixture with other populations that have occupied the Mediterranean area over different periods of history.

Based on the MDS results, which indicate an overlapping of *Alu* diversity with geography, we analyze how the observed genetic heterogeneity is spatially structured by hierarchical analysis of molecular variance (AMOVA). AMOVA analyses were performed to ascertain maximum genetic variance between groups (F_{CT}) and minimum genetic variance between populations within groups (F_{SC}), which guarantees the genetic consistency of a given classification. According to MDS results, of all the possible combinations the hierarchical classification that best fits this criterion was that dividing the whole set of populations into three groups: Western Europe, Eastern Europe, and North Africa. Upon assignment of the populations within these three broad geographic regions, AMOVA analyses were performed for each of the *Alu* insertions examined (Table 3). No statistically significant differences were found between populations within groups (F_{SC}) for the frequencies of any of the *Alu* elements

TABLE 3. *Results of the analyses of molecular variance* (F_{SC} *and* F_{CT}) *for eight polymorphic* Alu *insertions, considering population groups classed according to geography**

Insertion	Frequency range	Mean frequency	F_{SC} (%)[a]	F_{CT} (%)[b]
ACE	0.240–0.482	0.362	0.44 ns[c]	1.95[d]
TPA25	0.397–0.617	0.546	−0.15 ns	0.08 ns
PV92	0.175–0.398	0.257	1.01 ns	3.19[e]
APO	0.836–1.000	0.938	0.09 ns	3.21[e]
FXIIIB	0.293–0.600	0.425	0.86 ns	4.24[e]
D1	0.149–0.460	0.302	0.01 ns	1.08[d]
A25	0.048–0.235	0.134	−0.17 ns	0.65 ns
B65	0.510–0.734	0.599	−0.01 ns	0.48 ns

*Mean and range of variation of the allelic frequencies were calculated for the whole set of populations considered, including Valencia: Eastern Europe (Albania, Greece, Macedonia, and Turkey); Western Europe, (Andalusia, Basques, Catalonia, France, and Valencia); and North Africa, (Algeria, Sahara, Tunisia, North, Southeast, and West Morocco). Sources: Stoneking et al. (1997) and Comas et al. (2000, 2004).
[a]F_{SC}, genetic variation among populations within groups.
[b]F_{CT}, genetic variation among groups.
[c]ns, nonsignificant.
[d]Statistical significance for $P < 0.05$.
[e]Statistical significance for $P < 0.01$.

analyzed. By contrast, five out of the eight *Alu* loci contribute significantly to the geographic patterning of the genetic heterogeneity. The overall test for the significance of F_{CT}, which combines the separate probability values for each locus, was statistically significant ($P < 0.01$), indicating substantial genetic substructuring between the three geographical areas. The PV92, APO, and FXIIIB loci stand out for their relatively high contributions (3.19%, 3.21%, and 4.24%, respectively, $P < 0.01$) to the genetic variation between groups (F_{CT}). The overall test to verify the significance of F_{ST} was also statistically significant ($P < 0.01$). Similar results were obtained when the populations were distributed into just two geographical regions—Africa and Europe—and the F_{CT} levels for the same five insertions (FXIIIB, APO, PV92, ACE, and D1) were significant. Finally, when European populations were analyzed and distributed in two groups (Western and Eastern Europe), the statistical significance of the F_{CT} values was obtained for only two insertions (TPA25, 0.48%, $P < 0.01$; B65, 1.70%, $P < 0.05$).

The topology of the genetic structuring, i.e., the concordance between the genetic distances generated from the R matrix and the geographic distance matrix, was assessed using the Mantel test of matrix correspondence. Results revealed that both matrices were significantly correlated ($R = 0.283$, $P < 0.05$). This means that genetic relationships among populations might have been modeled mainly by isolation by distance. Nevertheless, when genetic and geographic topologies were ad-

justed (Fig. 3) the European populations tended to show a degree of kinship higher than expected according to the model of isolation by distance. In effect, from the thorough observation of Figure 3 it can be noted a clear approximation of the genetic coordinates of Eastern and Western European populations. Consistent with results previously described, the sample from Valencia occupied an intermediate position between the African and non-African populations. Finally, it is worth noting that Turkey remained separated from the rest of the populations considered and approximately equidistant of the two principal continental clusters.

DISCUSSION

Alu elements are among the best identical-by-descent markers for studying human evolution (Deininger and Batzer, 1999). This analysis of eight human-specific *Alu* polymorphisms in Valencia (Spain) shows that the allelic frequencies observed in the study population are very similar to the averages calculated over all the populations in the study. Consequently, Valencia stands at the centre of the whole group of Mediterranean populations, as shown by the MDS results. Likewise, the comparatively high average gene diversity in the Valencian sample (considering the eight *Alu* typed) evidences the potentially high rate of mixture of the Valencian gene pool. Such findings seem to be strongly conditioned by the historical past of the Valencia region, which is in turn largely determined by its geo-
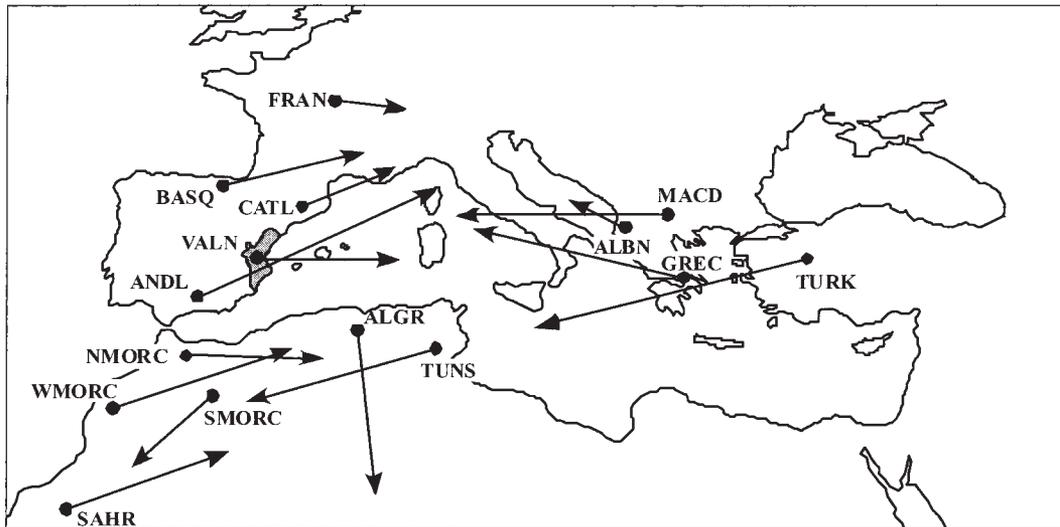
Fig. 3. Matrix fitting of geographic and genetic coordinates for 15 Mediterranean populations. Full circles represent geographic locations of the populations analyzed. Arrows indicate the location predicted by genetic kinship. Key: ANDL, Andalusia; ALBN, Albania; ALGR, Algeria; BASQ, Basque Country; CATL, Catalonia; FRAN, France; MACD, Macedonia; GREC, Greece; NMORC, North Morocco; SAHR, Sahara; SMORC, Southeast Morocco; TUNS, Tunisia; TURK, Turkey; VALN, Valencia; WMORC, West Morocco.

graphical location in the Western Mediterranean.

The Mediterranean has been important as a trade route throughout history, and has witnessed the flourishing and disappearance of numerous civilizations, some of which had great military, economic, and cultural influence on the lands along its coasts. Specifically, the Western Mediterranean was for a long time the setting for fierce economic and political struggles. As a result, the Iberian Peninsula (including Valencia) was subject to numerous external demographic influences. We have already mentioned the arrival of Neolithic culture in Valencia at a relatively early date, deduced from the dating of remains at Cova de les Cendres and Cova de l'Or (Zilhao, 2001). According to that study, this region of the Peninsula first met the Neolithic culture in the 7th millennium BP, probably by sea.

Other peoples of Mediterranean origin, such as Phoenicians, Greeks, Carthaginians, Romans, Arabs, and Berbers, later colonized the region. The Phoenician colonization took place around 800 BC. Their main settlements were in Africa Minor, on the islands of the central and western Mediterranean and in southern Spain (the Tartesians). Interestingly, the influence of these colonizations is still evident in the genetic background of the current population of

Valencia, e.g., in the high frequency of the cystic fibrosis mutation G542X. In order to explain the high frequency of G542X in the Valencian population, some authors have argued that this mutation was introduced into Spain via the Mediterranean, probably by the Phoenicians, between 2500 and 3000 years ago (Casals et al., 1993).

Another interesting finding of this work is the genetic distinction observed between the Mediterranean populations of North Africa and Europe. These results are in agreement with those reported in previous studies using classical genetic markers (Bosch et al., 1997; Simoni et al., 1999), short tandem repeats (Bosch et al., 2000), Y-chromosome (Bosch et al., 1999), polymorphic *Alu* insertions (Comas et al., 2000), and HLA-class II loci (Pérez-Miranda et al., 2003, 2004). Recent investigations have postulated that, even though the populations of the Iberian Peninsula are separated from North Africa by just 15 km of sea, genetic evidence indicates that the Straits of Gibraltar have acted as a strong barrier to gene flow between Africa and Europe (Comas et al., 2000; Simoni et al., 1999). The genetic heterogeneity of the two groups of Mediterranean populations could have been reinforced by the low population density of North Africa throughout history, which could have facilitated the effects of genetic drift

among these populations (Simoni et al., 1999). It has also been argued that this genetic dissimilarity could originate from a Mesolithic (or older) in situ differentiation of the populations in the northwestern regions of Africa (Bosch et al., 1997). On the other hand, it has been pointed out that a certain degree of gene flow from sub-Saharan Africa could have existed in the southern part of North Africa and in Saharans and Southeastern Moroccans because of a continuous gene flow across the Sahara desert (Bosch et al., 2000; Comas et al., 2000; Esteban et al., 2004). This gene flow could have helped maintain the differences between the population groups on the two shores of the Mediterranean. The causes listed here could in turn explain the relative genetic heterogeneity observed within the cluster of North African populations, among which the Algerian, Sahara, and Southeast Morocco populations stand out (see Fig. 2), with extreme frequencies for several *Alu* elements. On the contrary, some authors have found genetic affinities between Andalusians and North Africans based on classical markers, arguing that such similarities might be associated with both an ancient common genetic background and/or a special historical influence during the Muslim occupation of the Iberian Peninsula (Kandil et al., 1999). The findings of the present study do not seem to support such a hypothesis.

At this point, it must be also stressed the relatively remote position of Turkey from the rest of the populations in the two-dimensional genetic maps. In previous studies aimed at analyzing nucleotide substitutions and sequence diversity of the mitochondrial DNA (mtDNA) control region (D-loop), some authors have found a certain geographic patterning of the mtDNA polymorphism and suggested a stepping-stone position of the Anatolian Peninsula (Turkey) and the Caucasus region between the Middle East and Europe (Calafell et al., 1996; Comas et al., 1996). This suggestion is based on the hypothesis of the replacement of Neanderthals in Europe by the arrival of anatomically modern humans during the cultural expansion of the Upper Paleolithic age, between 50,000 and 100,000 years ago.

The intensity of the gene flow between the various Mediterranean regions and the balance between gene flow and genetic drift in each region would seem to be the main causes of the genetic similarities and differences observed between Valencia and the remaining populations analyzed. The gene pool of a population can also be modeled by mutation and selective pressure. The incidence of mutation on the variation of *Alu* insertion frequencies is practically negligible since they are identical-by-descent markers. Similarly, the role of natural selection on *Alu* elements is a bone of contention for scientists, since these markers are noncoding DNA fragments. It has been suggested that only 0.1% of human genetic disorders can be attributed to the insertion of *Alu* elements and 0.3% to the recombination of those elements (Deininger and Batzer, 1999). So although the existence of selective pressures cannot be completely ruled out, it does not seem to have played any major role in the configuration of the frequencies of the *Alu* insertions analyzed in this study.

In conclusion, the analysis of eight human-specific *Alu* insertions has revealed that the Mediterranean Sea has played a highly important role in the genetic make-up of the populations surrounding it: it has been determinant in the genetic distinctiveness between the peoples of the North and South (where it has acted as a strong genetic boundary). However, based mainly on results of the adjustment between geographic and genetic coordinates (see Fig. 3), we suggest the possibility that the Mediterranean Sea could also have propitiated the dissemination of cultures and genes between Eastern and Western Europe. In this last case, the Mediterranean seems to have acted as a migratory corridor, attenuating the differences in gene frequencies between the coastal populations. Results of the present study indicate that data on polymorphic *Alu* insertions may provide significant information about the evolutionary history of human populations. Nevertheless, it would be advisable that in the future other genetic markers are studied in the population of Valencia to be compared with the results obtained in neighboring populations to elucidate more precisely the role of this region of the Iberian Peninsula in the successive peopling processes of the Mediterranean area.

ACKNOWLEDGMENTS

LITERATURE CITED

Aler M, Salas A, Gisbert M, Carracedo A. 2001a. Data for nine autosomal STRs markers from Valencia (East Mediterranean coast of the Iberian Peninsula). Forensic Sci Int 116:37–39.

Aler M, Salas A, Sánchez-Diz P, Murcia E, Carracedo A, Gisbert M. 2001b. Y-chromosome STR haplotypes from a Western Mediterranean population sample. Forensic Sci Int 119:254–257.

Batzer MA, Deininger PL. 2002. Alu repeats and human genomic diversity. Nature 3:370–379.

Batzer MA, Deininger PL, Hellmann-Blumberg U, Jurka J, Labuda D, Rubin CM, Schmid CW, Zietkiecz E, Zuckerkandl E. 1996. Standardized nomenclature for Alu repeats. J Mol Evol 42:3–6.

Batzer MA, Stoneking M, Alegria-Hartman M, Bazan H, Kass D, Shaikh T, Novick G, Ioannou P, Scheer D, Herrera R, Deininger P. 1994. African origin of human-specific polymorphic alu insertions. Proc Natl Acad Sci U S A 91:12288–12292.

Bosch E, Calafell F, Pérez-Lezaun A, Clarimón J, Comas D, Mateu E, Martínez-Arias R, Morera B, Brakez Z, Akhayat O, Sefiani A, Hariti G, Cambon-Thomsen A, Bertranpetit J. 2000. Genetic structure of north-west Africa revealed by STR analysis. Eur J Hum Genet 8:360–366.

Bosch E, Calafell F, Pérez-Lezaun A, Comas D, Mateu E, Bertranpetit J. 1997. A population history of Northern Africa: evidence from classical genetic markers. Hum Biol 69:295–311.

Bosch E, Calafell F, Santos FR, Pérez-Lezaun A, Comas D, Benchemsi N, Tyler-Smith C, Bertranpetit J. 1999. Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. Am J Hum Genet 65:1623–1638.

Botstein D, White RI, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331.

Calafell F, Underhill P, Tolun A, Angelicheva D, Kalaydjieva L. 1996. From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks. Ann Hum Genet 60:35–49.

Casals T, Nunes V, Palacio A, Gimenez J, Gaona A, Ibanez N, Morral N, Estivill X. 1993. Cystic fibrosis in Spain: high frequency of mutation G542X in the Mediterranean coastal area. Hum Genet 91:66–70.

Comas D, Calafell F, Mateu E, Perez-Lezaun A, Bertranpetit J. 1996. Geographic variation in human mitochondrial DNA control region sequence: the population history of Turkey and its relationship to the European populations. Mol Biol Evol 13:1067–1077.

Comas D, Calafell F, Benchemsi N, Helal A, Lefranc G, Stoneking M, Batzer M, Bertranpetit J, Sajantila A. 2000. Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary through the Gibraltar Straits. Hum Genet 107:312–319.

Comas D, Schmid H, Braeuer S, Flaiz C, Busquets A, Calafell F, Bertranpetit J, Scheil HG, Huckenbeck W, Efremovska L, Schmidt H. 2004. Alu insertion polymorphisms in the Balkans and the origins of the Aromuns. Ann Hum Genet 68:120–127.

Deininger PL, Batzer MA. 1999. Alu repeats and human disease. Mol Genet Metab 67:183–193.

Dietz EJ. 1983. Permutation tests for association between two distance matrices. Syst Zool 32:21–26.

Esteban E, González-Pérez E, Harich N, López-Alomar A, Vía M, Luna F, Moral P. 2004. Genetic relationship among Berbers and South Spaniards based on CD4 microsatellite/*Alu* haplotypes. Ann Hum Biol 31:202–212.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491.

Guo SW, Thompson EA. 1992. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48:361–372.

Harpending H, Jenkins T. 1973. Genetic distance among southern African populations. In: Crawford M, Workman P, editors. Methods and theories of anthropological genetics. Albuquerque, NM: University of New Mexico Press. p 177–99.

Kandil M, Moral P, Esteban E, Autori L, Mameli GE, Zaoui D, Calo C, Luna F, Vacca L, Vona G. 1999. Red cell enzyme polymorphisms in Moroccans and southern Spaniards: new data for the genetic history of the western Mediterranean. Hum Biol 71:791–802.

Kruskal JB. 1964. Nonmetric multidimensional scaling: a numerical method. Psychometrika 29:115–129.

Lalouel JM. 1973. Topology of population structure. In: Morton NE, editor. Genetic structure of populations. Honolulu, HI: University of Hawaii Press. p 139–149.

Liu K, Muse S. 2004. PowerMarker: new genetic data analysis software. Version 3.0. Free program distributed by the author over the internet from http://www.powermarker.net.

Mantel NA. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220.

Pérez-Miranda AM, Alfonso-Sánchez MA, Peña JA, Calderón R. 2003. HLA-DQA1 polymorphism in autochthonous Basques from Navarre (Spain): genetic position within European and Mediterranean scopes. Tissue Antigens 61:465–474.

Pérez-Miranda AM, Alfonso-Sánchez MA, Vidales MC, Calderón R, Peña JA. 2004. Genetic polymorphism and linkage disequilibrium of the HLA-DP region in Basques from Navarre (Spain). Tissue Antigens 64:264–275.

Picornell A, Gomez-Barbeito L, Tomas C, Castro JA, Ramon MM. 2005. Mitochondrial DNA HVRI variation in Balearic populations. Am J Phys Anthropol 128:119–130.

Schneider S, Roessli D, Excoffier L. 2000. Arlequin: a software for population genetics data analysis. Geneva: Genetics and Biometry Laboratory, University of Geneva.

Simoni L, Gueresi P, Pettener D, Barbujani G. 1999. Patterns of gene flow inferred from genetic distances in the Mediterranean region. Hum Biol 71:399–415.

Smouse PE, Long JC, Sokal RR. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Syst Zool 35:627–632.

Sokal RR, Rohlf FJ. 1997. Biometry. The principles and practice of statistics in biological research,3rd ed. New York: WH Freeman and Company.

Stoneking M, Fontius J, Clifford S, Soodyall H, Arcot S, Saha N, Jenkins T, Tahir M, Deininger P, Batzer MA. 1997. Alu insertion polymorphisms and human evolution: evidence for a larger population size in Africa. Genome Res 7:1061–1071.

Tomas C, Picornell A, Castro JA, Ramon MM. 2001a. STR data for the AmpFlSTR profiler plus loci from Majorcan, Minorcan and Valencian populations (Eastern Spain). Forensic Sci Int 121:201–204.

Tomas C, Picornell A, Castro JA, Ramon MM. 2001b. Genetic analyses in five western Mediterranean populations: variation at five tetrameric short tandem repeat loci. Hum Biol 73:349–363.

Watkins WS, Ricker CE, Bamshad MJ, Carroll ML, Nguyen SV, Batzer MA, Harpending HC, Rogers AR, Jorde LB. 2001. Patterns of ancestral human diversity: an analysis of *Alu*-insertion and restriction-site polymorphisms. Am J Hum Genet 68:738–752.

Zilhao J. 2001. Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe. Proc Natl Acad Sci U S A 98:14180–14185.