Human Alu insertion polymorphisms in North African populations

Lotfi Cherni*[1, 2], Sabeh Frigi*[1], Hajer Ennafaa[1], Nabil Mtiraoui[3], Touhami Mahjoub[3], Amel Benammar-Elgaaied[1]

***Equally contribution**

Address:[1]: Laboratory of Molecular genetics, Immunology and Human Pathology at the Faculty of Sciences of Tunis, University El Manar, 2092 Tunis, Tunisia.

[2]: High institute of biotechnology, university of Monastir, Tunisia.

[3]: Laboratory of Biology and Genetics of Haematological and auto-immune disease, Faculty of Pharmacy of Monastir university of Monastir.

Correspondence to: Dr. Sabeh Frigi, Laboratory of Molecular genetics, Immunology and Human Pathology at the Faculty of Sciences of Tunis, University El Manar, 2092 Tunis, Tunisia. E-mail: sabehfrigi@yahoo.fr

Key words: North Africa; Tunisia; Libya; Alu insertion polymorphisms

Abstract

Several features make Alu insertions a powerful tool used in population genetic studies: the polymorphic nature of many Alu insertions, the stability of an Alu insertion event and, furthermore, the ancestral state of an Alu insertion is known to be the absence of the Alu element at a particular locus and the presence of an Alu insertion at the site that forward mutational change. This study analyses seven Alu insertion polymorphisms in a sample of 297 individuals from the autochthonous population of Tunisia (Thala, Smar, Zarzis and Bou Salem) and Libya with the aim of studying their genetic structure with respect to the populations of North Africa, Western, Eastern and Central Europe. The comparative analyses carried out using the MDS and AMOVA methods reveal the existence of spatial heterogeneity, and identify four population groups. Study populations (Libya, Smar, Zarzis

and Bou Salem) are closest to North African populations whereas Thala is isolated and is closest to Western European populations.

In conclusion, Results of the present study support the important role that migratory movements have played in the North African gene pool, at least since the Neolithic period.

Introduction:

The Maghreb is separated from the Sahel and sub-Saharan Africa by the vast stretches of the Saharan Desert, and from Europe by the Mediterranean Sea. Our research group (2010) has made an important contribution to studies of African human biology and culture in suggesting the complexity of Maghreban population history (Keita, 2010). Unveiling the history of human settlement in NW Africa is a complex task. It is the result of continuous complex network of migrations, invasions and admixture of people from different origins.

During prehistory, in the coastal Maghreb various Neolithic inter-regional interactions are in evidence based on archaeology. Indeed, the Maghreb has hosted several Neolithic traditions (Camps, 1982, Phillipson 2005) among which the Capsian tradition showed continuity with previous cultures, with evidence of these accepting domesticated sheep and goat into a local subsistence pattern, thus becoming neolithicised with a pastoralist economy (Sheppard et al., 1990, Rahmani, 2003, 2004; Keita, 2010). This might indicate that different people and heterogeneous populations could become unified under singular cultural practices and one language family by simply cultural adoption or adaptation (Keita, 2010). This would have been the case for Berber, the most ancient known language in North Africa.

During history, settlements of the Phoenicians, Romans, Vandals and others (Camps, 1982), along with the later importation of Europeans (Bennett 1960, Davis, 2004) would have

seemingly contributed to the current biological picture. The Muslim conquest of North Africa linked the Maghreb, the core of which comprises Morocco, Algeria and Tunisia, with the Middle East. Bedouin emigrated from Arabia to North Africa, followed by the Banu Hilal and Banu Sulaym who travelled through the Egyptian desert on camelback. The complete "Arabisation" of North Africa occurred after the arrival of the Hilalians who brought Arabic to what are now Tunisia, Algeria, Morocco and Western Sahara. Nonetheless many Berbers still speak Tamazight or Berber today. The level of arabisation at genetic level is less obvious, since North African populations are genetically distinct from Middle East population, though sharing with them language and religion (Ennafaa et al. 2009).

Among genetic markers used, uniparental markers have been very informative for demographic studies. However, their sexual inheritance and higher susceptibility to drift and selection could lead to biased information (Jobling & Tyler-Smith, 2003). Therefore, biparental markers are necessary to obtain a complete picture of the history and dispersals of human populations. Scattered throughout the human genome, occured as unique events in our evolution and apparently selectively neutral (Batzer & Deininger, 2002), polymorphic Alu insertions (PAI) have been intensively used. Alu elements are the predominant type of short interspersed elements (SINEs) in the human genome, with over 1 million copies comprising 10% of the total genome (Houck et al., 1979; Venter et al., 2001). The amplification of Alu elements through evolutionarily recent events coincided with the radiation of primates (Batzer and Deininger, 2002; Kapitonov and Jurka, 1996). Alu elements increase in number by retrotransposition, a process involving the insertion of reverse transcribed DNAs of Alu-derived transcripts, back into the genome, apparently by hijacking the L1 retrotransposition machinery (Boeke, 1997; Cost and Boeke, 1998; Dewannieux et al., 2003). Based on a hierarchical series of sequence mutations, Alu elements are classified into three major families designated as J, S, and Y, representing the oldest, intermediate, and youngest Alu

sequences, respectively. Each of these families is further divided into one or several subfamilies based on subfamily-specific diagnostic mutations (Batzer et al., 1990, 1996; Jurka and Smith, 1988). It is estimated that approximately 5000 young Alu elements are specific to humans (Batzer and Deininger, 1991). Among these young Alu elements, approximately 25% have inserted so recently that they are polymorphic among different human population groups, families, or even individuals with respect to their presence or absence in the genome (Batzer and Deininger, 2002). Because Alu insertions are unique events that are identical by-descent, they have been useful in genetic mapping and population genetics studies (Batzer et al., 1994; Batzer and Deininger, 1991; Perna et al., 1992; Roy- Engel et al., 2001; Salem et al., 2003; Stoneking et al., 1997; Tishkoff et al., 2000)

Alu markers have been used extensively for population structure and evolution, both at global (Bazter *et al.* 1996; Romualdi *et al.* 2002; Bamshad *et al.* 2003) and regional levels (Majumder *et al.* 1999; Comas *et al.* 2000, 2004; de Pancorbo *et al.* 2001; Nasidze *et al.* 2001; Ennafaa et al. 2006; Bahri et al. 2008; Frigi et al. 2010a). The objective of this study is to compare the genetic structure of some North African populations with respect to the other populations of North Africa, Western, Eastern and Central Europe, by using MDS and AMOVA methods. Based on Y chromosome and mtDNA studies, Berbers seem to have been issued from admixture of North African men and Iberian women, with a variable sub-Saharian female contribution (Cherni et al. 2009; Ennafaa et al. 2009; Frigi et al. 2010; Keita et al. 2010). Their characterization through the study of seven *Alu* polymorphisms, and thereby establishing their position with respect to other North African and European populations, offers new genetic data that contributes towards clarifying how the North Africa was populated within the framework of population movements in the Mediterranean area.

Material and Methods

Studied populations

A sample from Tunisia formed by 245 unrelated healthy blood donors has been analyzed (figure 1):

- 48 from Thala which is a town in western Tunisia located in the mountainous backbone of Tunisia. It is a municipality of 13,968 inhabitants. The traces of human presence in the region date back to the Paleolithic. Several cemeteries have been saying Capsian discoveries. Thala had a rich and eventful history, the human presence dating back at least 50,000 BC marked the Berber identity throughout the region. Its growth was partly caused by its industry based on pottery.

- 47 from Bou Salem which is a town in northwestern Tunisia, between Beja and Jendouba. Attached to the governorate of Jendouba, it is a municipality with 20,098 inhabitants. In Roman times, Bou Salem was the site of an imperial domain.

- 64 from Smar (government of Tataouine) which is a Beber region. It's a municipality of 13900 inhabitants.

- 86 from Zarzis which is inhabited by Accaras tribe. These current inhabitants settled in the region recently (in the late sixteenth century). According to historians they are native of Western Sahara; they have left for unknown reasons. The tribes Accara installed initially in southern Algeria and the Sahel in Tunisia and then resumed the road to the south with the intention to return to Mecca for pilgrimage.

A sample of 52 individuals was collected from different regions of Libya, another North African country. 'Libya', which initially indicated all Africa before corresponding to the current country, is a Greek derivation from "Libou", the name of an ancient native Amazigh (i.e. Berber) tribe. Libya today is bordered by the Mediterranean Sea to the north, by Algeria and Tunisia to the west, by Niger and Chad to the south and by Sudan and Egypt to the east. Arabic is the only official language, in spite of the numerous amazighophone areas in Nafousa, Zouara, Sokna, Aoujila, Ghadames, Awbarai, Ghat, Fugha, Oulad Itba, and many

nomadic Sahara tribes. For more than 20% of the population, the Amazigh language (i.e. Berber with its Libyan varieties) is the mother tongue.

Analyzed samples

Sampling was conducted in the hospital of each region under the supervision of healthcare staff and was preceded by a survey for every individual including the origin of the four grand parents and the families' names. We sampled only unrelated individuals with a level of Tunisian and Libyan autochthony up to three generations.

The study was approved by the local health authorities. The principles of confidentiality were strictly applied during the sampling process and informed consent was obtained from all individuals participating in the study. The 297 blood samples were used to purify genomic DNA using the conventional phenol/chloroform extraction method.

Alu insertion polymorphism analysis

Seven human-specific Alu insertion polymorphisms: ACE, APO-A1, F13B, TPA-25, PV92, B65 and D1 respectively located on chromosomes 17, 11, 1, 8, 16, 11and 3, were typed in each sample, using the previously described primers (Acrot et al. 1995a; b; Batzer et al. 1996;Garcia-Obregon et al. 2006). The PCR amplification conditions were performed as already described (Stoneking et al. 1997). Allele frequencies were calculated by direct counting. Hardy-Weinberg equilibrium was assessed by an exact test (Guo and Thompson 1992) provided by the Arlequin program v 2.0 package (Schneider et al. 2000).

Genetic relationships among populations were also depicted by a non-metric multidimensional scaling (MDS) analysis (Kruskal, 1964) based on analysis of the R-matrix (Harpending and Jenkins 1973). In order to ascertain the proportion of genetic variance attributable to differences within or between populations, genetic variance was hierarchically apportioned through the analysis of molecular variance (AMOVA; Excoffier et al. 1992)

performed with the Arlequin program v 2.0 package (Schneider et al. 2000). In this statistical analysis, a permutation procedure allows testing of significance of the fixation indices, which measure the relative contribution of genetic variation among populations within groups and among groups, respectively.

In order to study the geographic model of the Mediterranean populations genetic structure, the model of Harpending and Ward (1982) was used, which graphically represents the expected frequency of heterozygotes, according to the Hardy-Weinberg law, for each population with respect to the distance to the centroid, *ri* which is given by the formula $ri = (pi - P)2/[P(1-P)]$ where *pi* and *P* are respectively, the frequency of the *Alu* insertion in the population *i*, and in all of the populations included in the analysis as a whole. In turn, Harpending and Ward propose an island model of population structure in which there is a lineal relationship between heterozygosity and the distance to the centroid: $hi = H(1-ri)$ where *hi* and *H* correspond respectively to the heterozygosity value in the population *i* and in all of the populations as a whole.

Results

Alu insertion polymorphism

Alu insertions were polymorphic in all populations studied; Alu insertion frequencies and heterozygosity of the seven loci typed are shown in Table 1. Allele frequencies varied in Tunisian populations from 0.062 for PV92 in Smar to 0.833 for APO-A1 in Thala and in Libya from 0.240 for ACE to 0.615 for APO-A1. Hardy-Weinberg equilibrium was assessed by an exact test to calculate the p-value using the Markov-chain Monte Carlo method (Guo and Thompson 1992). The Hardy-Weinberg test appeared significant for the case of APO-A1 in Smar, ACE in Bou Salem and D1 in Libya, but after correcting for multiple tests applying a rough Bonferroni correction, the statistical test does not reach significant value.

In Tunisian populations, the average heterozygosity for each population (table 1) was lower in Zarzis (0.322) than in the other populations (range 0.373-0.444). Heterozygosity in Libyan population is closer to the Tunisian population of Thala.

MDS analysis

In order to ascertain population affinities based on Alu diversity, we compiled data obtained for the seven polymorphic loci from previously published papers (table 2 and figure 2) (Stoneking et al. 1997; Comas et al. 2000; 2004; Romualdi et al. 2002; Garcia-Obregon et al. 2006; Santovito et al. 2007; Bahri et al. 2008; Frigi et al. 2010a). A non-metric multidimensional scaling (MDS) analysis was carried out (figure 3). Populations were grouped according to geography. In agreement with previous studies on human diversity using polymorphic Alu insertions (Stoneking et al. 1997; Watkins et al. 2001; Garcia-Obregon et al. 2006) African and non-African populations clearly segregated along dimension I in the MDS representation. Thus, North African populations (North Morocco; West Morocco; Southeast Morocco; Algeria; Tunisia; North, Centre-South Tunisia; Sahara; Sejnane; Takrouna; Libyan; Smar, Bousalem and Zarzis) were plotted in the positive segment of the first axis (dimension I). The exception is for the Tunisian population of Thala which is isolated and seems to be the most genetically dissimilar. This North African population is characterized by having the highest insertion frequencies for B56, ACE, PV92, APO-A1, D1 and F13B (table 1).

On the other hand, the European populations are spread along dimension II: Western and Central European populations (Andalusia; Basque Country; Catalonia; Valencia; Canaries; France ; Swiss; Germany and Genova) are concentrated in the negative segment of dimension II, whereas Eastern European populations (the Aromuns of Albania; Albania;

Greek Cypriots; Turk Cypriots) are situated in the positive segment of dimension II. For Tunisian populations, the two Berber communities Sejnane and Takrouna are separated from North African populations and are closest to the Syrians and some Central European or Mediterranean populations (Germany and Genova) as already described by Frigi et al. 2010a.

Study populations (Libya, Smar, Zarzis and Bou Salem) are closest to North African populations whereas Thala is isolated and is closest to Western European populations. The different positions of North African populations suggest that the North African 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. It indicates also the complex history of human settlement in NW Africa.

AMOVA Analysis

When considering all the North African populations (our samples of Libya, Smar, Thala, Bou Salem and Zarzis, and those studied by Comas et al. (2000), Bahri et al. (2008) and Frigi et al. (2010a): Northern Moroccans; Western Moroccans; South Eastern Moroccans; Saharawis; Algerians; Tunisians, North-Centre-South Tunisians, Sejnane and Takrouna) as a single group in order to establish apportionment of genetic variance, the fraction of the genetic variance resulting from differences between populations was 2.01% ($p < 0.001$), whereas the remaining variance was found within populations (97.99%) (table 3).

When considering only Tunisian populations, AMOVA analysis reveals absence of geographic, linguistic and ethnic effect (table 3) along with significant genetic distances. This indicates the existence of limited gene flow between Tunisian populations.

Taking into account the fact that the previous results indicate that the populations included in the analyses may be differentiated into four groups (that correspond to the geographic areas of Western-Central Europe, Eastern Mediterranean Europe, North Africa and an admixture fourth group (Germany, Genova, Syria, Sejnane and Takrouna) (figure 3)), an analysis of

molecular variance (AMOVA) was carried out, in order to evaluate the genetic structure of the populations with more precision, along with degree of subdivision or geographic distribution of the genetic diversity. The reliability of the four groups is revealed by the results of the AMOVA analysis, in the case of the genetic between-group variation when the seven markers are considered simultaneously. Hence, the fraction of genetic variance resulting from differences among groups was 3.60% ($p < 0.001$), whereas the remaining variance was found within populations (95.30%). This data shows a remarkable heterogeneity among the considered groups and a weak variability between populations of the same group.

Gene flow

To gain more information about the internal genetic structure of our samples we have tested the Harpending and Ward (1982) model (Figure 4) for the populations in Figure 1 and 2. Our Tunisian samples: Thala, Smar ,Bou Salem; the Libyans and other samples are above the regression line, with Smar at a very short distance from the centroid, while Thala and Libyan populations are the most distant from the theoretical prediction. These populations have probably received considerable levels of gene flow from other populations. On the other hand, the position of Zarzis, Sejnane, Takrouna and the remaining samples under the line suggests the influence of isolation events in their genetic background.

Discussion:

In general, reconstructions of the recent history of human populations based on genetic data, whether classic markers, nuclear DNA polymorphisms, or mitochondrial DNA, reveal heterogeneity in North-African populations, perhaps greater than in other continents (Varela et al. 2008).

With the aim of adding information to the proposed existence of heterogeneity amongst North-African populations, Libya and some Tunisian populations were analyzed and have been compared with other North African and European populations, as explained in the

objectives of the study. *Alu* insertion polymorphisms were selected as they represent an important source of information about genetic diversity, both from the point of view of current diversity, and variability throughout the process of human evolution.

Starting from the fact that the markers analyzed are biallelic and that each allele is observed with a frequency greater than 0,1, to enhance the size of the samples would change anything to the results. This would have been different if we have analyzed STR that are multiallelic markers and enhancing the samples could reveal rare alleles. With a sample of 50 individuals you reveal alleles with frequencies upper or equal to 1%. Alleles with lesser frequencies could not be determinant for the structure of a population. Moreover, the Tunisian population analyzed are small with about 10 000 to 20 000 inhabitants and the samples collected from each community could be considered as representative. Starting from the size of the populations studied, the number of family names should not exceed 100. The family names are used to define if the subject belonging to the sample are not related.

The heterozygosity by population averaged across the seven Alu was the lowest in Zarzis (0.322) and the highest in Thala (0.444) and Libya. The high heterzygosity in the later populations evidences the potentially high rate of mixture of their gene pool. Such findings seem to be strongly conditioned by the historical past of the Maghreban region.

Using the MDS analysis, the presence of four clusters was observed, which are coherent with the genetic diversity that would be expected as a result of historical dynamics and the exchange of populations. The results of the AMOVA analysis support the reliability of the four clusters defined by the MDS, given statistical significance of the between-group variation for the seven markers. The important finding of this study is the proximity of Libya to North African populations. Regarding Tunisian populations, we remark that Bou Salem and Smar have affinities to North-African populations. Zarzis which is relatively isolated is also closer to these populations. The closeness between Libya and some Tunisian populations

indicates the gene flow between Tunisia and Libya across the Sahara. The position of the population of Thala is worthy of attention. It's isolated and separated from North African populations but is closer to Western European populations. This finding reflects a high Eurasian genetic component for some North African populations and confirms the heterogeneity of North-African populations.

The sub-Saharan origin of North African populations have been proposed on the basis of results indicating local evolution of Y chromosome and mtDNA African haplogroups (Ennafaa et al. 2009; Frigi et al. 2010). The much later transaharan trade in enslaved persons no doubt played a role in genetic contributions, but the egress from a dessicating Sahara with subsequent population formations would explain some of the "sub-Saharan" variation be it from western or eastern Africa. However, from our results, the North African populations appear more related to European than to sub-Saharan populations. The "Eurasian" component seems to have come in over a longer period of time (Keita, 2010). A small amount of gene flow per generation into a population/geographical region can drastically change its original gene frequencies in only a few thousand years as noted by Cavalli Sforza (1991). This genetic flow from Europe seems have happened since Neolithic period. Despite the fact that Neolithic expansion had the same effect in Northern Africa as in Europe, the Straits of Gibraltar acted as a barrier between the two continents, limiting gene flow between North-western Africa and Western Europe through the Iberian Peninsula (Comas et al. 2000;Garcia-Obregon et al. 2006; Varela et al. 2008; Frigi et al. 2010). This justifies the fact that the majority of North African populations appear in the MDS analysis as a separate group from European populations.

The method of Harpending and Ward (1982), which uses the theoretic regression of population heterozygosity and distance to the centroid as a valid indicator to analyze the geographic model of genetic structure in populations, has been applied to the populations of

North Africa in order to evaluate the relationship of these populations with other Mediterranean populations in this area. In this case, the Smar population is very close to the theoretical prediction line, as is the case with most of the populations in its geographical area included in the analysis, even Northern Morocco, Western Morocco, North-Central-South of Tunisia and other Tunisian populations. The samples from Algeria, Sejnane, Takrouna and Zarzis are low outliers, which is interpreted as being indicative of these populations being receivers of a limited gene flow from the area. This result suggests that Zarzis, which is characterized by its low heterozygosity and these populations have either had a smaller effective population size, or they have been more isolated. The samples from Thala and Libya are high outliers, a position that does not necessarily imply that they recently received a greater external contribution, but which is indicative of a series of genetic differences with respect to the populations included in the analysis that may be due to the effect of the genetic contributions inherent to their origin.

Taken together, results on Y chromosome, mtDNA and Alu Insertions in North Africa allow to propose a scenario for this region. The ancient sub-Saharan settlement would have been followed by admixture with Iberian populations. But, as the North African Y chromosome remained dominant in the region, we could argue that this admixture have been realized in one direction: North African men and Eurasian women, explaining the gene flow from Europe and high frequency of European types of mtDNA in North Africa as compared with Y chromosome. This situation would not be the result of drift toward Eurasian mtDNA. Our results on Alu insertions interestingly confirm that this gene flow happened several times probably always on the same direction. These matrimonial exchanges between North Africa and Europe should be considered in a context of patriarchal societies with men attached to territory and women from different regions including Europe. Hence, genetic diversity on one hand and relationship with Europe should have been due to women. This result supports the

important role that migratory movements have played in North African populations, at least since the Neolithic period and suggests their diverse origins.

In conclusion, based on this study, we confirm the different genetic structures of North African populations indicating weak gene flow between them due to high endogamy rates that also should be considered in patriarchal socioeconomic context. In the case of Libya, it would be advisable to study in the future other genetic markers, to be compared with results obtained in this study to elucidate more precisely the role of this region of North Africa in the successive peopling processes of the Mediterranean area.

Acknowledgements

We gratefully acknowledge the blood donors whose availability made this study possible. Thanks also to the anonymous reviewers whose comments improved the presentation of this manuscript.

References:

Acrot S., Fontius J., Deininger P. et al.1995a. Identification and analysis of a young polymorphic Alu element. Biochim Biophys Acta 1263:99–102.

Acrot S, Shaikh T, Kim J, Bennett L. et al. 1995b. Sequence diversity and chromosomal distribution of young Alu repeats. Gene 163:273–278.

Bahri R, Esteban E, Moral P. et al. 2008. New insights into the genetic history of Tunisians: Data from Alu insertion and apolipoprotein E gene polymorphisms. Ann Hum Biol 35:22–33.

Bamshad MJ, Wooding S, Watkins W.S. et al. 2003. Human population genetic structure and inference of group membership. Am J Hum Genet. 72(3):578-89.

Batzer M.A., Arcot S. S., Phinney J. W. et al. 1996. Genetic variation of recent Alu insertions in human populations. *J. Mol. Evol.* 42: 22-29.

Batzer MA, Deininger PL. 2002 *Alu* repeats and human genomic diversity. *Nature Reviews, Genetics*, **3:**370-379.

Batzer MA, Deininger PL. 1991. A human-specific subfamily of Alu sequences. Genomics 9(3):481-7.

Batzer M, Stoneking M, Alegria-Hartman M,et al. 1994. African origin of human-specific polymorphic Alu insertions. Proc Natl Acad Sci USA 91: 12228–12292.

Batzer M. A., Kilroy G.E., Richard P.E., Shaikh T.H., Desselle T.D., Hoppens C.L. &

Bennett, N. R. 1960. Christian and Negro slavery in eighteenth-century North Africa. *J. Afr. Hist.*1:65–82.
Boeke JD. 1997. LINEs and Alus--the polyA connection. Nat Genet. 16(1):6-7.

Camps, G. 1982. Beginnings of pastoralism and cultivation in northwest Africa and the Sahara: Origins of the Berbers. In *The Cambridge History of Africa*, J. D. Clark, ed. Cambridge, U.K.: Cambridge University Press, v. 1, 548–623.

Cavalli-Sforza, L. 1991. Genes, languages, peoples. *Sci. Am.* 265:106–110.

Cherni, L., Fernandes, V., Pereira, B. et al. 2008. Post-Last Glacial Maximum expansion from Iberia to North Africa revealed by fine characterisation of mtDNA H haplogroup in Tunisia. *Am J Phys Anthropol* 139(2):253-260

Comas D, Calafell F, Benchemsi N. et al. 2000. Alu insertion polymorphisms in NWAfrica 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, & al. 2004. Alu insertion polymorphisms and the origins of the Aromuns. Ann Hum Gen 68:120–127.
Cost GJ, Boeke JD. 1998. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. Biochemistry. 37(51):18081-93.
Davis, R. C. 2004. *Christian Slaves, Muslim Masters: White Slavery in the Mediterranean, the Barbary Coast, and Italy (1500–1800)*. New York: Macmillan.

Deininger P.L. 1990. Structure and variability of recently inserted *Alu* family members. *Nucl. Acids Res.* 18, 6793–6798.

De Pancorbo MM, López-Martínez M, Martínez-Bouzas C, & al. 2001.The Basques according to polymorphic Alu insertions. Hum Genet. 109(2):224-33

Dewannieux M., Esnault C. & Heidnann T. 2003. Line-mediated retrotransposition of marked Alu sequences. *Nat. Genet* 35: 41-48.

Ennafaa, H., Cabrera, V.M., Abu-Amero, K.K. et al. 2009. Mitochondrial DNA haplogroup H structure in North Africa. *BMC Genet* 10:8.

Ennafaa H, Ben Amor M, Yacoubi-Loueslati B. et al. 2006. Alu polymorphisms in Jerba Island population (Tunisia): Comparative study in Arab and Berber groups. Ann Hum Biol 33:634–640.

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 13:479–491.

Frigi, S., L. Cherni, K. Fadhlaoui et al. 2010. Ancient local evolution of African mtDNA haplogroups in Tunisian Berber populations. *Hum. Biol.* 82(4):367–384.

Frigi, S., H. Ennafaa, Ben Amor M. et al. 2010 a. Assessing human genetic diversity in Tunisian Berber populations by Alu insertion polymorphism. *Ann. Hum. Biol.: 1-10.*

Garcia-Obregon N, Alfonso-Sanchez MA, Perez-Miranda A.M.et al. 2006. Genetic position of Valencia (Spain) in the Mediterranean Basin according to Alu insertions. AmJ Hum Bio 18:187–195.

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

Harpending HC, Ward RH. 1982. Chemical systematics and human populations. In: Nitecki M, editor. Biochemical aspects of evolutionary biology. Chicago: University of Chicago Press. pp. 213–256.

Harpending HC, Jenkins T. 1973. Genetic distance among southern African populations. In: Crawford MH, Workman PL, editors. Methods and theories in anthropological genetics. New Mexico: Albuquerque University. p 177–199.

Houck CM, Rinehart FP & Schmid CW. 1979. A ubiquitous family of repeated DNA sequences in the human genome. J. Mol. Biol. 132(3):289-306

Jobling M. A. & Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nat. Rev. Genet.* 4:598–612.

Jurka J, Smith T. 1988. A fundamental division in the Alu family of repeated sequences.

Proc Natl Acad Sci U S A. 85(13):4775-8.

Kapitonov V & Jurka J. 1996 .The age of Alu subfamilies. J Mol Evol. 42(1):59-65.

Keita S.O.Y (2010). Biocultural Emergence of the Amazigh (Berbers) in Africa: Comment on Frigi et al. (2010). Hum. Biol. 82 (4): 385-393.

Kruskal J.B. 1964. Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29:115-129.

Phillipson, D. 2005. *African Archaeology*. Cambridge, U.K.: Cambridge University Press.

Perna NT, Batzer MA, Deininger PL, & al. 1992. Alu insertion polymorphism: a new type of marker for human population studies. Hum Biol. 64(5):641-8.

Majumder PP, Roy B, Banerjee S, & al. . 1999 . Human-specific insertion/deletion polymorphisms in Indian populations and their possible evolutionary implications. Eur J Hum Genet 7(4):435-46.

Nasidze I., Risch G., Robichaux M., Sherry S., Batzer M.A. & Stoneking M. 2001. Alu insertion polymorphisms and the genetic structure of human populations from the caucasus. *European Journal of Human Genetics* **9**: 267-272.

Rahmani, N. 2003. *Le Capsien typique et le Capsien supérieur*. Cambridge Monographs in Archaeology 57. BAR International Reports 1157. Cambridge, U.K.: Cambridge University Press.

Rahmani, N. 2004. Technological and cultural change among the last hunter-gatherers of the Maghreb: The Capsian (10,000 BP). *J. World Prehist.* 18:57–105.

Roy-Engel Astrid M, Carroll M, Vogel E. et al. 2001. Alu insertion polymorphisms for the study of human genomic diversity. Genetics 159:279–290.

Romualdi C, Balding D, Nasidze I.S. et al. 2002. Pattern of human diversity, withen and among continents, inferred from biallelic DNA polymorphism. Genome Res 12:602–612.

Salem A.H., Kilroy G.E., Watkins W.S. et al. 2003. Recently Integrated Alu Elements and Human Genomic Diversity. *Mol.Biol.Evol.* 20: 1349-1361.

Santovito A, Selvaggi A, Cervella P,& al. 2007 . Polymorphic Alu insertions in five North-West Italian populations. Am J Hum Biol. 19(4):589-92.

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

Sheppard, P., and D. Lubell. 1990. Early Holocene Maghreb prehistory: An evolutionary approach. *Sahara* 3:63–69.

Stoneking M, Fontius JJ, Clifford S.L. et al. 1997. Alu insertion polymorphisms and human evolution: Evidence for a larger population size in Africa. Genome Res 7:1061–1071.

Tishkoff, S. A. Pakstis AJ, Stoneking M, & al. 2000. Short tandem-repeat polymorphism/Alu haplotype variation at the *PLAT* locus: implications for modern human origins. *Am. J. Hum.Genet.* 67, 901–925.

Varela T, Farin˜a J, Die´guez L.P.et al. 2008. Gene flow and genetic structure in the Galician population (NW Spain) according to Alu insertions. BMC Genetics 9:79.
Venter JC, Adams MD, Myers EW & al. 2001. The sequence of the human genome. Science 291(5507):1304-51
Watkins WS, Ricker CE, Bamshad MJ & al. 2001. Patterns of ancestral human diversity: an analysis of Alu-insertion and restriction-site polymorphisms. Am J Hum Genet 68(3):738-52.

Table 1. Alu insertion frequencies and heterozygosity in studied populations

Population	n	Markers / Rates	B65	ACE	D1	APO-A1	F13B	PV92	TPA-25	Population Het
Libyans	52	Alu ins freq	0.432	0.240	0.317	0.615	0.355	0.269	0.451	
		Het	0.490	0.364	0.433	0.473	0.457	0.393	0.495	0.443
Thala	48	Alu ins freq	0.718	0.604	0.645	0.833	0.500	0.468	0.538	
		Het	0.404	0.478	0.457	0.278	0.500	0.497	0.497	0.444
Smar	64	Alu ins freq	0.335	0.343	0.156	0.656	0.289	0.062	0.601	
		Het	0.445	0.450	0.263	0.451	0.410	0.116	0.479	0.373
Zarzis	86	Alu ins freq	0.110	0.116	0.203	0.651	0.168	0.255	0.302	
		Het	0.195	0.205	0.323	0.454	0.279	0.379	0.421	0.322
Bousalem	47	Alu ins freq	0.276	0.297	0.276	0.638	0.340	0.255	0.531	
		Het	0.399	0.417	0.399	0.461	0.4356	0.448	0.498	0.436

Table 2. Alu insertion frequencies in Mediterranean populations.

	n	TPA-25	ACE	APO	F13B	PV92	D1	B65	References
Europe									
Germany	70	0.514	0.464	0.867	0.049	0.100	0.310	0.348	Romualdi et al.(2002)
France	53	0.56	0.48	0.99	0.42	0.23	0.46	0.57	Stoneking et al. (1997)
Swiss	43	0.45	0.37	0.94	0.48	0.20	0.34	0.58	Stoneking et al. (1997)
Basque	96	0.568	0.443	0.953	0.484	0.188	0.380	0.604	Comas et al. (2000)
Catalonia	60	0.608	0.300	0.983	0.500	0.175	0.350	0.525	Comas et al. (2000)
Andalusia	67	0.590	0.470	0.985	0.448	0.194	0.306	0.552	Comas et al. (2000)
Valencia	101	0.556	0.387	0.940	0.476	0.232	0.322	0.529	Garcia-Obregon et al. (2006)
Greek Cypriots	48	0.53	0.39	0.95	0.62	0.25	0.27	0.65	Stoneking et al. (1997)
Turk Cypriots	33	0.58	0.33	0.98	0.39	0.33	0.35	0.64	Stoneking et al. (1997)
Albania	60	0.557	0.467	1.000	0.600	0.203	0.267	0.670	Comas et al. (2004)
Albania Aromuns	49	0.500	0.367	1.000	0.684	0.306	0.316	0.638	Comas et al. (2004)
Genova	30	0.450	0.067	1.000	0.167	0.233	0.433	0.267	Santovito et al. (2007)
North Africa									
N.Morocco	111	0.617	0.333	0.910	0.338	0.333	0.288	0.608	Comas et al. (2000)
W.Morocco	140	0.575	0.314	0.929	0.293	0.343	0.304	0.614	Comas et al. (2000)
SE.Morocco	49	0.510	0.265	0.847	0.306	0.398	0.194	0.510	Comas et al. (2000)
Sahara	58	0.397	0.284	0.836	0.371	0.310	0.259	0.534	Comas et al. (2000)
Algeria	47	0.532	0.266	0.915	0.315	0.287	0.149	0.734	Comas et al. (2000)
Tunisia	48	0.604	0.240	0.875	0.344	0.313	0.245	0.594	Comas et al. (2000)
Takrouna	33	0.469	0.106	0.863	0.212	0.121	0.348	0.560	Frigi et al. (2010a)
Sejnane	47	0.521	0.276	0.882	0.180	0.170	0.436	0.531	Frigi et al. (2010a)
NSC Tunisia	96	0.572	0.278	0.932	0.345	0.392	0.306	0.531	Bahri et al. (2008)
Middle East									
Syria	70	0.507	0.400	0.926	0.275	0.176	0.290	0.314	Romualdi et al.(2002)

Table 3. Results of the analyses of molecular variance for seven polymorphic Alu insertions.

Groups	Composition of groups	% of variance among groups	% of variance among populations within groups	% of variance within populations
All North African populations	LYB,ALG,NMO,WMO,SMO,SAH, SEJ,TAK,BSM,SMR, THA,ZRZ,TUN,NSC	-	2.01*	97.99
All Tunisian populations	SEJ,TAK,BSM,SMR, THA,ZRZ,TUN,NSC	-	2.05*	97.95
Berber populations	SEJ,TAK,SMR,THA	-	5.34*	94.66
Arab populations	BSM,TUN,ZRZ	-	2.71*	97.29
Berbers vs Arabs	SEJ,TAK,SMR,THA vs BSM,TUN,ZRZ	-0.74 ns	4.78*	95.96
North populations vs South populations	SEJ,TAK,BSM,THA vs SMR,ZRZ	-0.12 ns	4.07*	95.05
Tunisians vs North Africans	All Tunisian populations vs LYB,ALG,NMO,WMO,SMO,SAH	0.2 ns	3.92*	95.88
Western-Central Europe vs Eastern Mediterranean Europe vs North Africa vs the admixture fourth group	BAS,CAT,CAN,VAL,AND,SWI,FRA vs ARA,GCY,TCY,ALB vs LYB,ALG,NMO,WMO,SMO,SAH, SEJ,TAK,BSM,SMR,ZRZ,TUN,NSC vs SEJ,TAK,SYR,GEN,GER	3.60*	1.1*	95.30

ns: non significant; * ($p<0.05$)

LYB Libya; ALG Algeria; NMO North-Morocco; SAH Sahara; SMO Southeast-Morocco; WMO West-Morocco; BSM Bou Salem; THA Thala ; SMR Smar; ZRZ Zarzis ; SEJ Sejnane ; TAK Takrouna ;
NSC North-South-Center Tunisians ; TUN Tunisians ; AND Andalusia ; ALB Albania, ARA Aromuns (Albania) ; BAS Basques (Spain) ; CAT Catalonia ; FRA France ; GER Germany ; GCY Greek Cypriots ; SWI Swiss, TCY Turk Cypriots ; SYR Syria ; GEN Genova.

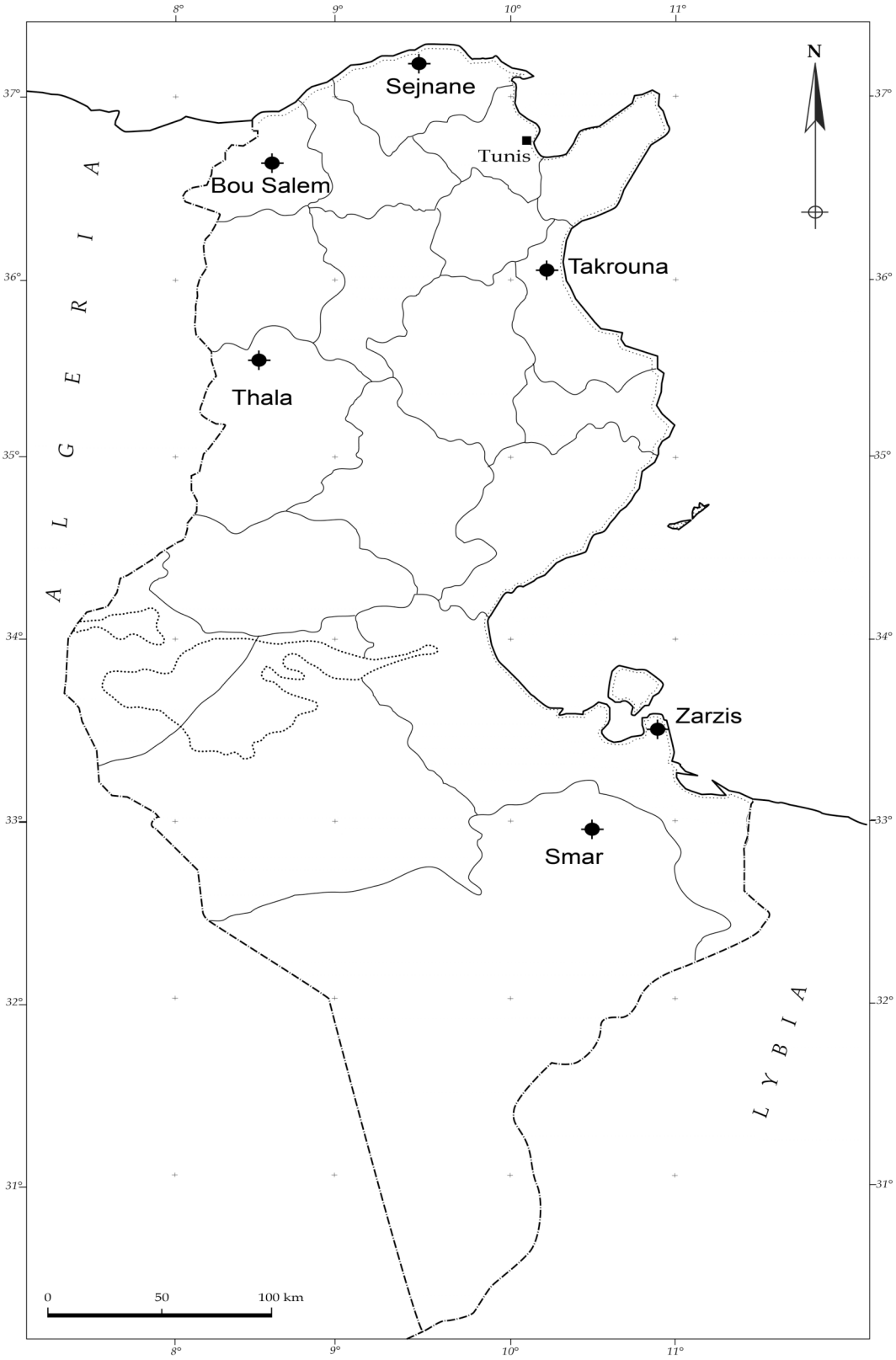

Figure 1. Location of Tunisian populations

BSM Bou Salem*; THA Thala* ; SMR Smar*; ZRZ Zarzis* ; SEJ Sejnane ; TAK Takrouna ; NSC North-South-Center Tunisians; TUN Tunisians (Locations of NSC and TUN are not indicated in articles)
*: studied populations

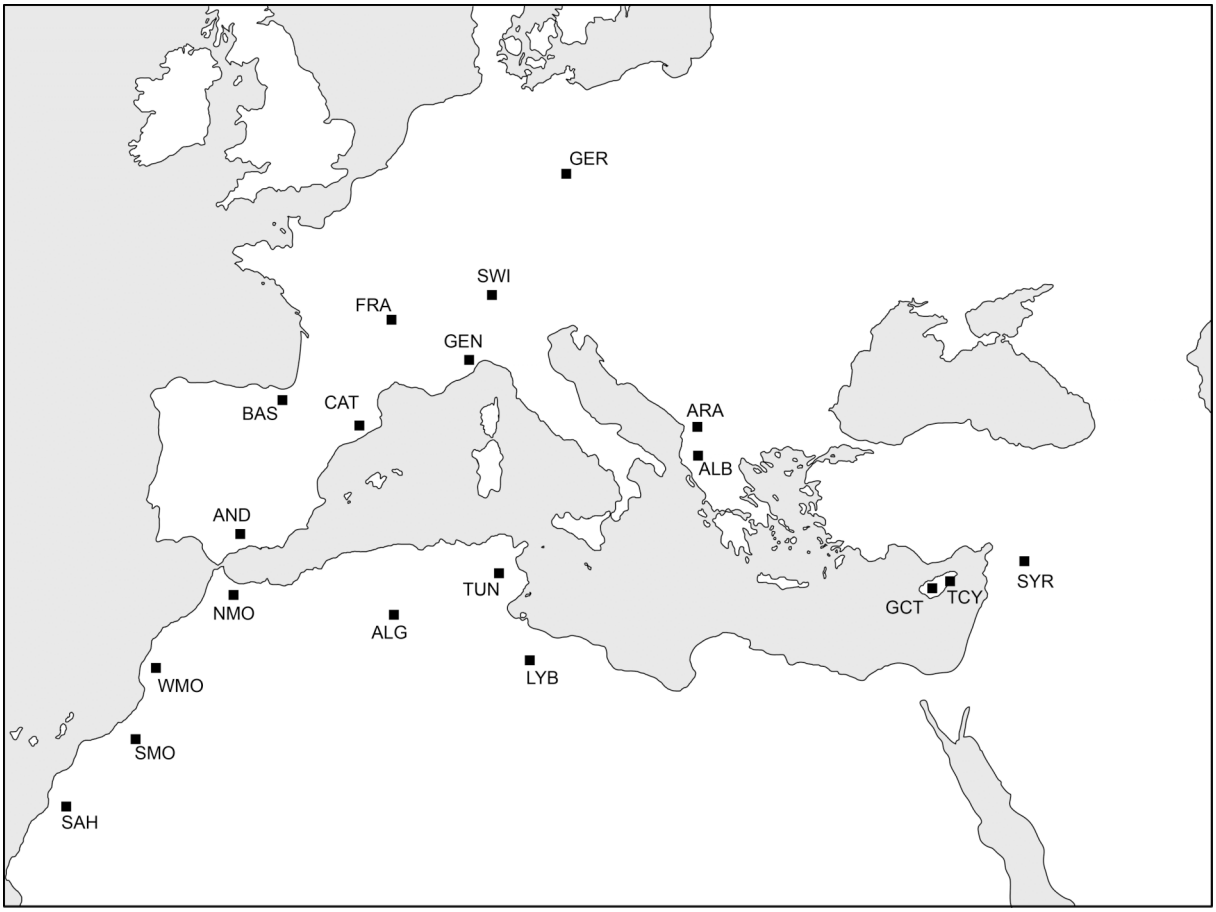

Figure 2. Geographical location of populations used in these analyses.

TNI Tunisia, LYB Libya, ALG Algeria, AND Andalusia, ALB Albania, ARA Aromuns (Albania), BAS Basques (Spain), CAT Catalonia, FRA France, GER Germany, GCY Greek Cypriots, NMO North-Morocco, SAH Sahara, SMO Southeast-Morocco, SWI Swiss, TCY Turk Cypriots, WMO West-Morocco, SYR Syria, GEN Genova.

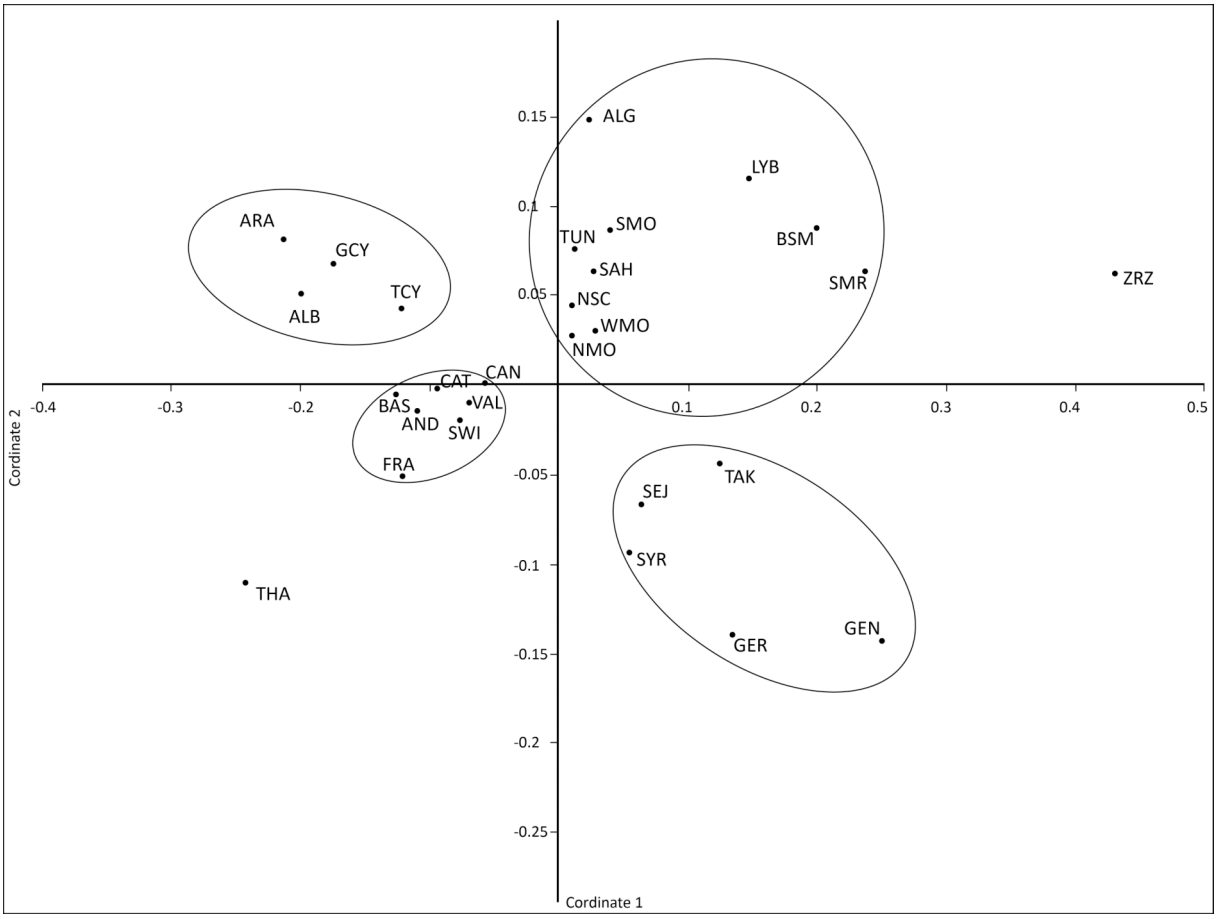

Figure 3. Non-metric multidimensional scaling (MDS) applied to analyze genetic relationships among twenty-seven populations. Stress value is 0.134.

The acronyms are the same as those used in Figures 1 and 2.

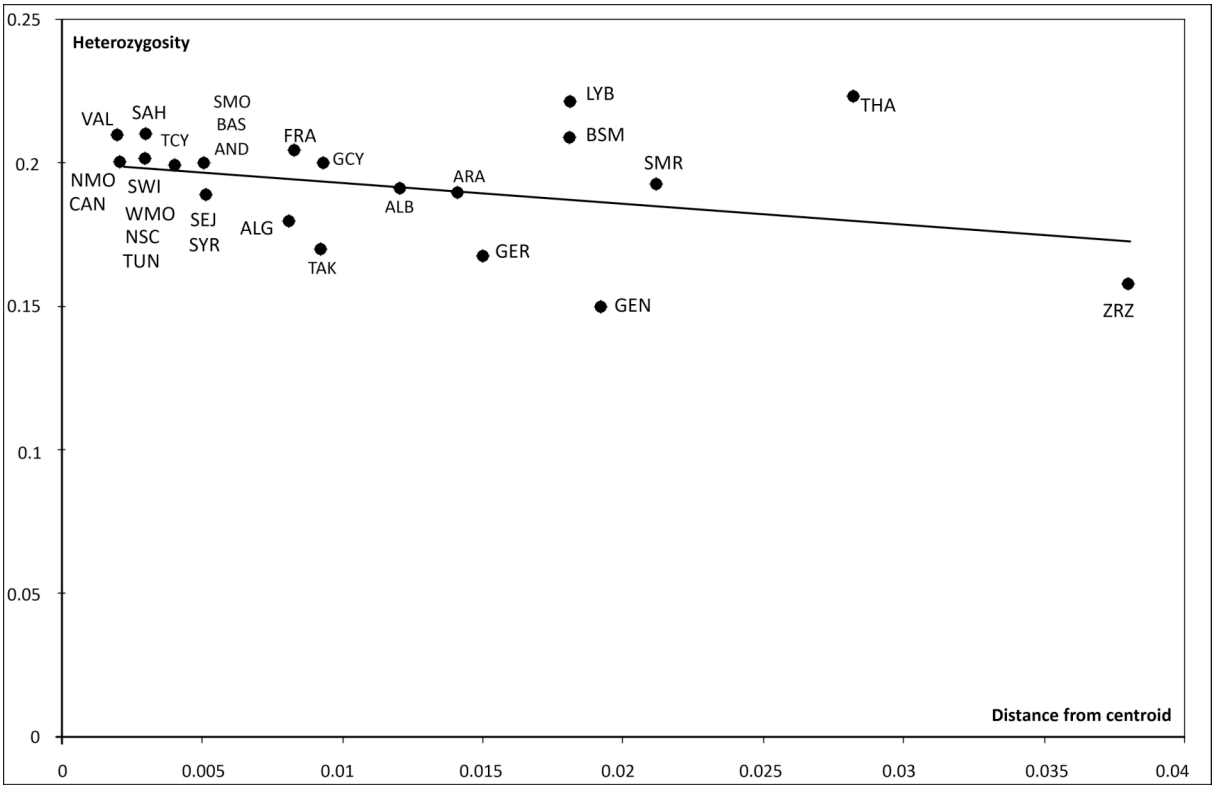

Figure 4. Heterozygosity and distance from centroid in twenty-seven populations from the European Mediterranean area. The line represents the expected regression according to the model of Harpending and Ward (1982).