Coll. Antropol. **27** (2003) 2: 523–536
UDC 616.1-056.7:575.17(262)
Original scientific paper

Molecular Variation at Functional Genes and the History of Human Populations – Data on Candidate Genes for Cardiovascular Risk in the Mediterranean

Pedro Moral[1], Neus Valveny[1], Antonio López-Alomar[1], Carla Calo[2], Mostafa Kandil[3], Nordin Harich[3], Emili González-Pérez[1], Marc Via[1], Esther Esteban[1], Jean Michel Dugoujon[4] and Giuseppe Vona[2]

[1] Department of Animal Biology-Anthropology, University of Barcelona, Barcelona, Spain
[2] Department of Experimental Biology, University of Cagliari, Cagliari, Italy
[3] Department of Biology, University Chouaïb-Doukkali, El Jadida, Morocco
[4] Center of Anthropology, UMR 8555, CNRS, University Paul Sabatier, Toulouse, France

ABSTRACT

*A screening of 22 DNA polymorphisms has been performed in western Mediterranean populations (Iberian Peninsula, Morocco, and Central Mediterranean Islands). The analyzed markers correspond to polymorphic sites in several candidate genes for cardiovascular disease including apolipopoteins and their receptors (APOA1, APOB, APOE, APOC1, APOC2, LPA, and LDLR), genes implied in the hemostasis regulation (Factor VII, α and β-fibrinogen, α and β platelet-integrin, tissue plasminogen activator, and plasminogen activator inhibitor-1), and the angiotensin converting enzyme gene. The results are presented of a partial analysis carried out in following population samples: 6 from the Iberian Peninsula, 2 from Morocco, and 3 from Central Islands. The degree of interpopulation diversity was significant and consistent with data from other kind of genetic polymorphisms. The apportionment of the allele frequency variance supported a geographic structure into three main regions: Central Mediterranean Islands, the Iberia Peninsula and North Africa. The genetic distance pattern is compatible with a south-to-north North African influence in the Iberian Peninsula and a remarkable gene flow from sub-Saharan Africa into Morocco. Epidemiologicaly, North Africa is characterized by high frequencies of LPA PNR alleles with high number of repeats (protective for cardiovascular risk) and high frequencies of the APOE*E4 allele (risk factor) as compared with European populations.*

Key words: *apolipoproteins, coagulation, fibrinolysis, DNA markers, Mediterranean populations*

Received for publication December 19, 2002

Introduction

During the last decades, the main focus of attention in the study of population genetic diversity, an important field of biological anthropology, has switched to mitochondrial DNA and Y chromosome variation[1,2] due in part to the theoretical advantages that uniparental inheritance and lack of recombination offer to the interpretation of phylogenetic relationships. However, these two regions represent a very low proportion of the human genome. They function as single loci (single recombination units) and the evolutionary history of a particular gene does not have to be the same that the history of human populations. In other terms, the evolutionary processes leading to present population genetic configuration must have affected the whole genome and, hence, a real understanding of the genetic history of populations will depend on the integrated view of the gene variation taking into account the evolutionary peculiarities of the different genome regions. In this way, recent studies on autosomal DNA sequence variation are providing different and interesting data on the origin and history of human populations[3] that should be taken into account for a global explanation of the origin and history of human groups[4].

A small but important part of our genome is represented by functional genes. Variation in these genes represented by »classical polymorphisms« constituted the first tool used for human population genetic studies. At present, with the advent of molecular biology techniques an extreme supply of genetic markers are available in this kind of genes. From the population point of view it may be argued that the possible implication of selective forces may be an added drawback for historical interpretations. However, this should be a challenge rather than a restriction since the selection is a real component of the human evolutionary history and, excluding mutation changes incompatible with life, its effects on the extant variation in functional regions are likely not much higher than in other genome regions[5,6]. On the other hand, the population distribution of this kind of genes provide additional information that may be relevant to the epidemiological knowledge and as a first approach to understand the significance of gene-environment interactions[7] in the configuration of the biology of human groups.

This paper deals with the population molecular variation at fifteen functional genes listed in Table 1 we have scored through a set of twenty-two polymorphic markers, in the frame of a collaborative study on the genetics of human Mediterranean populations that is carrying out by researchers from Barcelona (Spain), Cagliari (Italy), El Jadida (Morocco), and Paul Sabatier (France) Universities. The analyzed genes are implied in different metabolic pathways, such as lipid metabolism, coagulation and fibrinolysis processes, and are candidate genes for cardiovascular (CVD) diseases. In fact, most of the markers in this study have been almost exclusively tested so far in epidemiological studies searching for CVD susceptibility gene factors, with a few exceptions such as the APOE, the VNTR 3' in APOB, and the three Alu insertions (ApoA1, ACE, and Tpa25) whose geographic distribution has been worldwide examined.

The objectives of the present work are: 1) to explore the variation of this group of genes functionally related to cardiovascular susceptibility in general population samples of Mediterranean origin; and 2) to determine their usefulness as genetic tools for the study of human populations by testing the gene differentiation between the Iberian Peninsula, central Mediterranean Islands, and NW Africa populations, and detecting the possible amount

TABLE 1
LIST OF THE TYPED GENES AND POLYMORPHISMS IN MEDITERRANEAN POPULATIONS

Genes	Symbol	OMIM	Ch. location	Typed markers
Apolipoprotein B	APOB	107730	2p24	Ins/del, Asn/ser, RFLP/XbaI, EcoRI, MspI, VNTR 3'
Apolipoprotein E	APOE	107741	19q13.2	Isoforms (RFLP/HhaI)
Apolipoprotein C1	APOC1	107710	19q13.2	RFLP/HincII
Apolipoprotein C2	APOC2	207750	19q13.2	RFLP/AvaII
Low-density lipoprotein cholesterol receptor	LDLR	606945	19p13.2	RFLP/PvuII
Apolipoprotein (a)	LPA	15220	6q27	(TTTTA)n repeat
Apolipoprotein A1	APOA1	107680	11q23	Alu insertion
Angiotensin-converting enzyme	ACE	106180	17q23	Alu insertion
Factor VII	FVII	227500	13q34	Ins/del; HVR4 repeat
Fibrinogen beta-unit	FGB	134830	4q28	C/T^{-148}, RFLP/BclI
Fibrinogen alpha-unit	FGA	134820	4q28	RFLP/TaqI
Plasminogen activator inhibitor-1	PAI1	173360	7q21.3-q22	RFLP/HindIII
Plaminogen activator, tissue	PLAT	173370	8p12	Alu insertion
Platelet integrin, alpha-2B	GPIIb	273800	17q21.32	Ins/del (9bp)
Platelet integrin, beta-3	GPIIIa	173470	17q21.32	RFLP/TaqI

of gene flow across the two main geographical barriers, the Mediterranean Sea in its westernmost part (the Gibraltar Straits) and the Sahara Desert, which may have conditioned the settlement and differentiation in the Mediterranean region.

Cardiovascular diseases are a major cause of death in the industrialized world[8] showing remarkable regional differences in incidence[9]. In Europe, for instance, the Mediterranean region is characterized by low CVD frequencies as compared with other continental areas, and this low rate has been related with some »protective« factor, both environmental and/or genetic. Some variants of the makers here analyzed have been associated in some, but not all, epidemiological studies with increased risk for cardiovascular accidents usually through its effect on intermediate risk factors such as plasma lipid, Factor VII, fibrinogen, and/or plasminogen levels[10]. At present, the majority of epidemiological studies providing some information on the population distribution of these polymorphisms have been performed in central-north European origin groups. In this context, our study, centered on Mediterranean areas with different CVD prevalences, attempts to fill the important gap in the geographic variation of these genes, as a first approach to understand correlation between population genetic variation and differential frequencies of currently important and common diseases in human populations.

The Mediterranean region embracing South Europe, Near East, and the north -Saharan part of Africa has been the target for interest of human population scholars for a long time. It is commonly

accepted that the origin of present populations living in the northern and southern Mediterranean shores may be traced back to the independent establishment of modern humans from the East with a strong influence of later Neolithic expansions bringing Indo-European and Afro Asiatic languages on both sides[11,12]. Inside this general frame, it is likely that the genetic composition of most present Mediterranean groups may be the result of a complex pattern of isolation periods, mutual interactions, and recurrent colonizations. In this way, the genetic peculiarities of some populations, such as Sardinia, Basques and other islands, with regard to other Europeans, have been repeatedly related to drift and isolation processes[11,13]. A particular issue of the genetic history of Mediterranean populations concerns the north-south relationships. The mentioned linguistic differentiation has been interpreted as acting as an isolation factor enhancing the biological divergences between South Europe and North Africa[14]. On the other hand, several anthropometric and cultural (Iberomarusian and Capsien cultures) data are consistent with remarkable circum -Mediterranean contacts[15] since pre-Neolithic times. Historical records concerning, for instance, to Phoenician settlements, Roman conquests, Islamic expansions, including the occupation of the Iberian Peninsula by North Africans for more than eight centuries (8^{th} to 15^{th} centuries) document the integration of this region into the Mediterranean world. However, the possible demographic impact associated to these population events is controversial. Recent genetic studies focusing on this question seem to support well different interpretations. Opposite hypotheses claim from a common »paleo -North African« origin for Iberians and North African populations[16] to a total demographic isolation across the Gibraltar Straits[17]. In general, apart from some HLA data stressing a particular genetic affinity between Iberians and NW Africa[18,19], most genetic data support a remarkable differentiation between population living in the two Mediterranean sides[20–22], but many of them too (from classical polymorphisms[23], autosomal STR[24], Y chromosome haplogroups[25]) are consistent with a certain degree of gene flow between north and south shores of the Mediterranean Sea. Any case, both the antiquity of north-south population contacts and their demographic impact within the Mediterranean keep being still open questions.

Material and Methods

Population samples

A total of 1,214 autochthonous individuals from 11 populations in the western half of the Mediterranean were analyzed (Figure 1). From each population, blood samples were obtained, preferentially in rural areas, from unrelated healthy people, of both sexes, with their four grandparents originated in the same geographical region, during different campaigns approved by local Ethical Committees. Informed consent was obtained from all participants. The Iberian Peninsula samples included Basques (n=112; Guipuzcoa province, N Spain), Pasiegos (n=96; Pas Valley, N Spain), Central Spanish (n=120; Avila and Segovia provinces), Catalans (n=88; Girona province, NE Spain), and Andalusians (n=100; Almeria and Granada provinces, SE Spain). The South Mediterranean samples comprised Moroccan Berbers (n=138; Khenifra region in Moyen Atlas), Arab-speakers from central-south Morocco (n=101; Doukkala region), and Canary islanders (n=74; Tenerife Island). Finally, the Central Mediterranean Islands were represented by samples from Corsicans (n=100), Sardinians (n=187), and Sicilians (n=98).

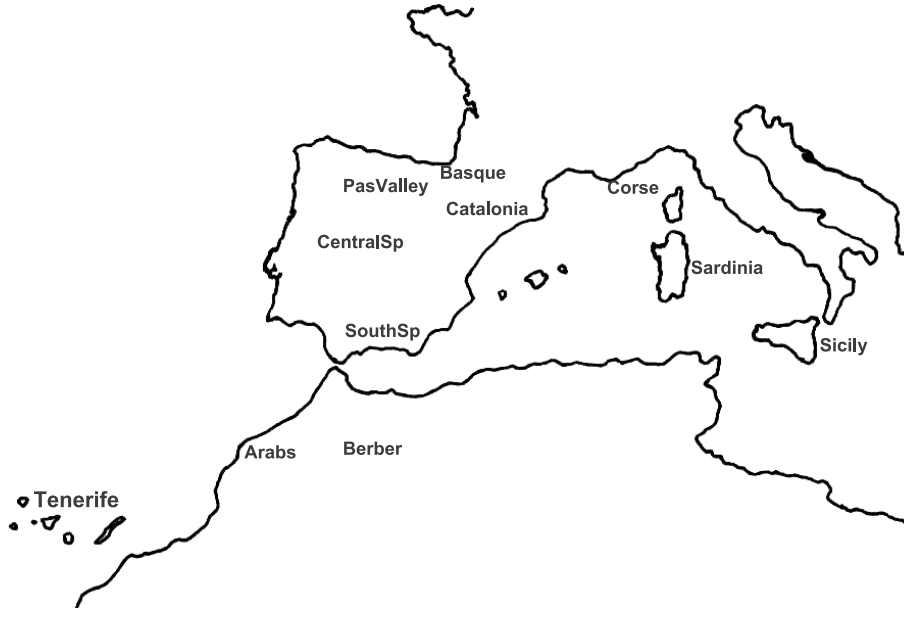

Fig. 1. Geographic location of the Mediterranean samples analyzed in the present study.

Methods

From DNA, extracted from whole blood according the phenol-chloroform standard method and amplified by PCR in a Thermal Cycler (GeneAmp PCR System 9600; Perkin Elmer), a total of 22 polymorphisms (19 biallellic and 3 tandem repeats) were typed using primers and PCR amplification conditions described elsewhere[26,27]. Agarose electrophoresis standard techniques were used for the phenotype identification of the insertion/deletion markers in the APOB, FVII, and GPIIb genes, the Alu insertion polymorphisms, and the RFLPs in the APOB, APOE, APOC1, APOC2, LDLR, LPA, FGB, FGA, GPIIa, and PAI genes (Table 1) from the corresponding digested fragments. The Asn/Ser substitution at the 4311 of the APOB genes was detected by ASO methods. The variants of the VNTR 3' and HVR4 repeat polymorphisms in the APOB and FVII genes, respectively, were typed by agarose (2.5%) electrophoresis and alleles of the (TTTTA)n pentanucleotide in the LPA gene by 10% nondenaturing polyacrylamide gel electrophoresis.

So far the number of polymorphisms typed has not been completed in all population samples (Table 2), in most cases for sample availability reasons, ranging from the APOB Asn/Ser substitution and the LDLR/PvuII analyzed only in the Iberian samples to the APOE, APOC1, APOC2, LPA makers that have been determined in the 11 samples. By populations, five samples (two from Spain, the three Mediterranean Islands, and Moroccan Arabs) were tested for at least 19 polymorphisms. A total of 187 (77.3%) out of the 242 possible (marker x sample) analyses were done that allow a good representation of the population variation of the examined genes.

TABLE 2
SAMPLE SIZES AND POLYMORPHISMS ANALYZED IN THE PRESENT STUDY BY POPULATION

	Populations										
Markers	Basques	Andalusia	Pas Vall.	Center	Catalonia	Canary I.	Berbers	Arabs	Corsica	Sardinia	Sicily
APOB-i/d	112	100	96	120	88	74		101	100	139	95
APOB-XbaI	112	100	95	120	87	74		101	98	131	98
APOB-MspI	111	98	94	120	84	71		95			
APOB-EcoRI	111	100	95	120	88	74		101	99	130	92
APOB-asn/ser	112	100	95	120	88	74		101			
APOB-vntr 3'	112	99	96	119	88	73		101	95	138	97
APOE-HhaI	111	99	96	120	88	72	121	101	99	136	94
APOC1-HincII	111	100	90	120	88	74	120	101	100	146	96
APOC2-AvaII	112	100	94	114	88	74	117	98	100	139	91
LPA-Penta	112	100	96	120	88	74	138	101	96	149	93
LDLR-PvuII	108	100	86	101	78	71					
Alu-ACE	100	100	93	92			135	92	61	106	51
Alu-APOA1	108	97	87	86			116	83	43	108	51
Apu-TPA25	102	98	95	95			136	98	46	186	50
FVII-i/d	88	97					121	98	46	182	50
FVII-HVR4	91	95					81	83	46	159	50
FGB-HindIII	90	100					122	98	46	182	50
FGB-BclI	85	98					121	80	46	106	50
FGA-TaqI	86	100					121	93	56	183	50
GPIIIa-TaqI	85	98					121	66	46	187	50
GPIIb-i/d	91	100					119	96	46	186	50
PAI-HindIII	78	99					120	83	46	185	50

Statistical analysis

Allele frequencies were calculated by direct counting and deviations from Hardy-Weinberg expectations were assessed by an exact test. Haplotype frequencies in the APOB, Factor VII, Fibrinogen, and integrin loci, and APOE-C1-C2 cluster were estimated by the expectation-maximization algorithm, and linkage disequilibrium was tested from genotype frequency distributions[28]. Gene diversity for population and locus was estimated according to the Nei's methods[29]. For each marker, population genetic differentiation was assessed by G and exact tests and the extent of genetic variation of the markers analyzed by the Nei's coefficients of gene diversity. Calculations were preformed using the Arlequin and Genepop v.3.3 statistical packages[30,31].

Pairwise genetic distances between populations were estimated according Reynolds coefficient[32] for different marker combinations and represented as neighbor-joining trees (NJ) whose consistency was checked by bootstrap resampling analysis (1000 iterations)[33]. Population relationships were also depicted by principal component analysis (PCA) from the correlation matrix of the original allele frequencies.

Population genetic structure was approached by an Fst hierarchical analysis of the genetic variance attributable to different geographic population subdivisions by the analysis of the molecular variance (AMOVA) using the Arlequin software. Estimates of gene flow across the Gibraltar Straits and Sahara Desert were obtained applying the Bertorelle and Excoffier method[34].

Results

All loci were polymorphic in all tested populations and a total of 70 alleles were detected, 39 corresponding to biallelic markers plus APOE and 31 to repeat polymorphisms (VNTR 3' APOB: 20 alleles; LPA repeat: 7, and FVII/HVR4: 4 alleles). Seventeen out of 187 tests for Hardy-Weinberg equilibrium showed significant departures, but after application of Bonferroni correction, only five (2.7%) tests were significant (TPA25, FVII/HVR4, and GPIIIa/TaqI in Sardinia; PAI/HindIII in Corsica, and GPIIIa/TaqI in Moroccan Berbers) that could likely be explained by random statistical fluctuations. In general, allele frequencies ranged around European values when data were available[35–38]. As most conspicuous features inside the Mediterranean scope it may be outlined the relatively high frequencies of APOE*E4 allele and LPA alleles with high number of repeats found in North African populations as well as the different frequency pattern in the APOC2 locus as compared with Europeans (around 0.5 for each allele)[39]. From an epidemiological point of view, high APOE*E4 frequencies could be an indicator of future increased risk for cardiovascular diseases in these populations taking into account the rapid social and lifestyle changes currently occurring in Morocco, however, the scope of this risk is difficult to assess since it might be overcome to some degree by the »protective« effect reported for LPA alleles with high number of repeats.

The haplotype variation observed in the Mediterranean for the five gene regions considered is indicated in Table 3.

The number of different haplotypes is logically depending on the number and variability of the loci in each region, being in our case higher in apolipoprotein than in hemostasis regions, although the number of haplotypes with frequencies higher than 5% is similar in the five regions analyzed. Linkage disequilibrium was tested in a total of 23 combinations of pairs of loci and significant values were found between markers from the APOB, APOE-C1, FVII, and Fibrinogen genes with a con-

siderable variation between populations. It is noteworthy that the amount of linkage disequilibrium in the three island populations from Corsica, Sardinia, and Sicily was lower than in the remaining Mediterranean groups (Figure 2).

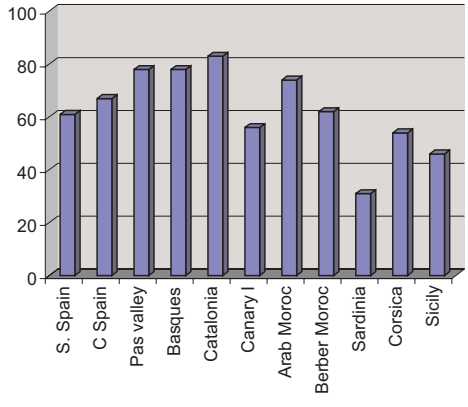

Fig. 2. Proportion (%) of linkage disequilibrium in the Mediterranean populations.

Gene diversity

A gene diversity analysis (Table 4) indicates that, as expected, the more important part of the total genetic variation found for all markers (Ht=0.385) corresponds to within-population diversity (Hs=0.378) and only a small, but significant, part 2.0% to interpopulation differentiation.

As for within-population diversities, the average heterozygosity for all 22 markers is 0.378 with a variation range between 0.178 (APOB/MspI) and 0.778 (APOB/VNTR) clearly depending on the type of marker (Table 3). The gene diversity differences between markers were significant (Kruskal-Wallis test, $p<0.0001$). Mean population heterozygosities across markers ranged between 0.354 (Sardinia) and 0.419 (Sicily), but no significant differences between samples were found (Kruskal-Wallis test, $p=0.988$).

Between-population heterogeneity as measured by Gst (a measure equivalent to Wright's Fst) and tested by chi-square is indicated Table 4. The average Gst value for the 22 markers is 0.020. This value is of the same order of magnitude that other reported for other markers[40] and even higher than those from hypervariable DNA marker previously described[24], indicating that the analyzed markers are a useful tool for detecting population genetic differentiation even at a microgeographical level. Sixteen markers were significantly heterogeneous at the 1 per cent level; an additional four markers were significant at the 5 per cent level, and only three markers (APOE/ HhaI, APOC1/HincII, and LDLR/PvuII) were not significantly heterogeneous across populations.

A hierarchical analysis of the variance of the allele and haplotype frequencies on the basis of different geographical population subdivisions show a clear geographic structure of the gene variation. The

TABLE 3
HAPLOTYPE VARIATION IN THE MEDITERRANEAN REGION

Gene region	No. samples	No. haplotypes		Haplotypes >5%
		Mean	Range	
APOB(i/d-XbaI-EcoRI-vntr3')	10	26.2	19–34	5.7
APOE-C1-C2	11	9.2	7–11	3.1
FVII(i/d-HVR4)	7	4.7	4–6	3.3
FG(HindIII-BclI-TaqI)	7	6.4	4–8	3.9
Integrin(TaqI-i/d)	7	4.0	4	4.0

TABLE 4
GENETIC HETEROGENEITY FOR 22 MARKERS IN THE MEDITERRANEAN REGION

Locus	No. samples	Hs	Ht	Gst	Comp. p
APOB-i/d	10	0.405	0.410	0.013	0.001±0.000
APOB-XbaI	10	0.492	0.495	0.010	0.015±0.004
APOB-MspI	7	0.178	0.181	0.018	0.003±0.001
APOB-EcoRI	10	0.290	0.293	0.011	0.011±0.003
APOB-asn/ser	7	0.350	0.355	0.016	0.001±0.000
APOB-vntr3'	10	0.778	0.784	0.012	0.000±0.000
APOE-HhaI	11	0.265	0.266	0.006	0.078±0.010
APOC1-HincII	11	0.237	0.239	0.001	0.952±0.007
APOC2-AvaII	11	0.486	0.497	0.021	0.000±0.000
LPA-Penta	11	0.457	0.464	0.014	0.000±0.000
LDLR-PvuII	6	0.370	0.371	0.007	0.152±0.014
Alu-ACE	8	0.447	0.470	0.048	0.000±0.000
Alu-APOA1	8	0.130	0.132	0.017	0.004±0.001
Alu-TPA25	9	0.493	0.500	0.014	0.000±0.000
FVII-i/d	7	0.290	0.302	0.040	0.000±0.000
FVII-HVR4	7	0.394	0.423	0.070	0.000±0.000
FGB-HindIII	7	0.337	0.345	0.022	0.001±0.001
FGB-BclI	7	0.364	0.268	0.016	0.002±0.001
FGA-TaqI	7	0.345	0.351	0.016	0.014±0.004
GPIIIa-TaqI	7	0.375	0.382	0.019	0.001±0.000
GPIIb-i/d	7	0.451	0.460	0.020	0.000±0.000
PAI-HindIII	7	0.481	0.491	0.022	0.003±0.001
Average		0.378	0.385	0.020	0.000±0.000

Hs=mean intrapopulation heterozygosisty; Ht=gene diversity in the entire population; Gst=gene diversity among populations relative to the total genetic variation in the entire population; Comp. p = population differentiation tested by an exact test.

61% of the markers exhibit a variance pattern (within-region variation < between -region variance) consistent with the sample classification into three regions: Iberia, North Africa, and Central Mediterranean Islands. The other two tested population groupings (North/ South and East/West) are less supported (48% and 26% of all markers, respectively).

Population relationships

The relationships among the seven population samples for which more data were available were examined from genetic distances shown in Table 5. These populations include Basques and South Spain from the Iberian Peninsula, Arab- and Berber-speakers from Morocco, and Corsica, Sardinia, and Sicily, representing the three sub-regions from the Western Mediterranean. In order to test the population information yielded by the two main kind of genes (apolipoproteins and hemostasis), calculations were done from apolipoprotein markers (8 single markers and two haplotype systems with a total of

71 gene frequencies) and hemostasis markers (9 single markers and 3 haplotype systems, with a total of 40 gene frequencies) separately, and then using all markers together. In the three cases, distance estimates support a same intra-region pattern corresponding the lowest values (from 0.005 to 0.010) to between-Moroccan comparisons while the whole formed by the three Central Mediterranean Islands appears as the more heterogeneous region (average distances from 0.039 to 0.074). Average between-region distances show the greatest differentiation between Iberia and Central Islands (0.0398) from hemostasis genes, and between North Africa and Central Islands (0.0312 and 0.0374) from both apolipoprotein and all markers. On the other hand, the lowest average inter-region distances based on hemostasis and all loci frequencies correspond to Iberia-North Africa comparisons. The distances of the two Iberian populations to Moroccans present remarkable differences, South Spain showing consistently (in the three analyses) lower distance than Basques especially to Arab -speaker Moroccans (from 2 to 5 times lower). In fact, in most cases distances of South Spain to the two Moroccan samples are equivalent, or even lower, than that between the two Iberian samples (Table 5).

Consistently, the NJ tree in Figure 3 divides the populations in three main groups: Iberia, Central Islands, and North Africa. The node clustering North Africans is well supported according to the close genetic affinity uncovered by genetic distances whereas the other nodes show less bootstrap support. A correspondence analysis (data not shown) gave a similar population pattern to that displayed in the NJ tree.

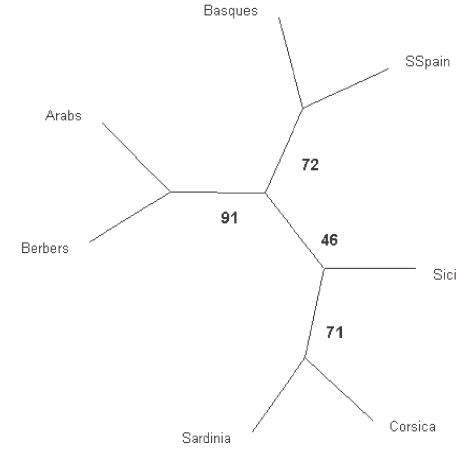

Fig. 3. Neighbor-joining tree from the genetic distances based on allele frequencies of 19 polymorphisms.

TABLE 5
GENETIC DISTANCES BETWEEN MEDITERRANEAN POPULATIONS (x 10000)

	S. Spain	Basques	Arabs	Berbers	Corsica	Sardinia	Sicily
S. Spain		104/216	133/52	207/129	138/242	110/311	175/253
Basques	171		243/311	307/309	233/481	280/734	212/367
Arabs	75	280		56/106	267/255	218/348	311/260
Berbers	162	314	94		391/389	377/407	310/428
Corsica	210	380	282	429		126/419	131/229
Sardinia	247	548	320	403	332		164/742
Sicily	257	305	322	487	215	597	

Above diagonal: distances based on apolipoprotein allele frequencies/distances from hemostasis makers. Below diagonal: distances based on both kinds of markers.

Discussion

In this paper we have explored the molecular variation in functional genes related to cardiovascular diseases in several general population samples from three different areas of the Western half of the Mediterranean region. The fact that all markers have not yet typed in all population samples might affect as much only side aspects such as the genetic diversity of some particular region as the Iberian Peninsula or the geographical variation of the LDLR gene, but it does not determine the results relative to the objectives of this study. The variation found in the functional genes examined reveal a degree of between-population genetic heterogeneity and a pattern of population relationships consistent with that from other genetic markers that make them good tools for the study of genetic differentiation between human groups. This analysis provides new molecular data to the knowledge of the genetic structure of present Western Mediterranean populations corroborating the geographic structure of the gene variation in this region. In all the analyses performed, a clear differentiation of the Central Islands from the rest was detected and also, even in a lesser degree, between Iberian and Moroccan populations.

The Central Island samples analyzed show the greatest genetic distances both to Iberians and North Africans, appearing Sardinia as the most differentiated population. The genetic peculiarity of Sardinia within the European picture is well known and has been repeatedly evidenced from autosomal[11] and uniparental markers[41], and interpreted as the result of periods of genetic drift and geographic isolation, with a limited genetic impact of the series of the conquests of the island during its history[42]. By regions, our genetic distance analysis reveals that the group formed by the three island populations is the most heterogeneous likely related to different settlement histories and geographic factors as it has been pointed out in previous studies[13,43]. In contrast to reported mitochondrial data[44], our results do not indicate any particularly closer genetic distance between Corsicans and Basques in agreement with other autosomal surveys.

As far the Iberia to North Africa relationships, the Spanish samples analyzed present closer genetic distances to Moroccan (average distance 0.0208) than to Central Mediterranean Island populations (average distance 0.0324) supporting a certain genetic affinity between populations living on both sides of the Gibraltar Straits. Concerning this question, previous genetic studies and interpretations are controversial ranging from a common origin for Iberians and North Africans[16] to a substantial genetic discontinuity of Paleolithic roots[14,17]. Genetic similarities between both populations have been suggested from several HLA data[18,19,45] and other anthropological traits[46]. Furthermore, data on Y-chromosome[25], classical markers[23], and autosomal STR[24] are consistent with a gene flow between these regions.

An important historical event was the conquest of the Iberian Peninsula by North Africans in the 8th century that lasted in the southern part of the Peninsula for eight centuries and entailed an important cultural and technological legate although the demographical impact of this conquest is a controversial issue. According to historical information, this North African cultural influence was stronger in the southern part of the Peninsula than in the northern areas, here represented by the South Spain and Basque samples, respectively. The fact that the genetic differentiation of Moroccans from South Spain is lesser than that from Basques (genetic distances more

than two times shorter) suggests a certain correlation between the different North African influence in the Peninsula and the genetic variation observed for the frequencies of the analyzed markers. A raw estimate[34] of the Moroccan gene flow into South Spain based on the allele frequencies of 20 markers and assuming Basques and Arab-speaker Moroccans as parental populations, gives a bootstrap (10,000 iterations) average contribution of mR=0.458 (bootstrap S.D.= 0.080). Obviously, a certain bias may be not discarded due to the selection of parental populations; nevertheless our data indicate a substantial North African contribution to the south part of the Iberia Peninsula. Values of the same order of magnitude have been reported from a large battery of Alu polymorphisms typed by us in the same populations, but that are different from other genetic studies using the same kind of polymorphisms[17], raising old questions such as that the histories of genes and populations have not to be coincident or the potential sample dependence of population results focusing on genetically similar groups and/or wide geographical areas with a relative heterogeneity as can be North Africa.

The two Moroccan samples analyzed, in spite of belonging to different linguistic groups, show the shortest genetic distances indicating that these cultural differences are not correlated with the present data on the variation in the autosomal markers examined and agree with other studies supporting that the Arabization in NW Africa from the Arab invasions in the 7th and 11th centuries was a rather cultural than demographic process[20].

Concerning the sharp differences reported in other studies between North African and South European populations, we assessed the possible sub-Saharan influence into North Africa as a potential factor of differentiation. From frequency data from a Sub-Saharan population (mainly from an Ivory Coast sample studied by us for several polymorphisms considered in the present study) the amount of Sub-Saharan gene flow into North Africa was 0.277. This value agrees with previous genetic surveys[47,48], suggesting an alternative factor that may help to understand the present genetic differences between the two coasts of the Western Mediterranean.

In summary, our study of 22 polymorphisms in functional genes shows a degree of population diversity that allow a clear genetic differentiation between human Mediterranean groups comparable to other genetic markers and that may be applied to trace back genetic histories. The general consistency of the regional genetic structure found in Western Mediterranean confirms the usefulness of this kind of markers in population studies. Although the potential implication of selection processes may theoretically represent an inconvenient for historical interpretations, it is difficult to think of an effective selective role on the gene variation here analyzed since the most plausible selective mechanism should be through cardiovascular diseases affecting mainly after the reproductive lifetime. Consequently, data from this study complete our knowledge of the genetic composition of Mediterranean human groups providing a new view on particular and controversial issues such as the genetic influence across the Gibraltar Straits.

Acknowledgements

This work was supported by PB98 -1235-C03-01 and BMC2002-01224 from the Spanish Ministry of Ciencia y Tecnologia and 2001SGR00089 from Generalitat de Catalunya grants.

REFERENCES

1. RICHARDS, M., V. MACAULAY, Am. J. Hum. Genet., 68 (2001) 1315. — 2. HAMMER, M. F., T. M. KARAFET, A. J. REDD, H. JARJANAZI, S. SANTACHIARA-BENERECETTI, H. SOODYALL, S. L. ZEGURA, Mol. Biol. Evol., 18 (2001) 1189. — 3. YU, N., Z. ZHAO, Y. X. FU, N. SAMBUUGHIN, M. RAMSAY, T. JENKINS, E. LESKINEN, L. PATTHY, J. B. JORDE, T. KUROMORI, W. F. LI, Mol. Biol. Evol., 18 (2001) 214. — 4. TEMPLETON, A., Nature, 7 (2002) 46. — 5. KHOGALI, S. S., B. M. MAYOSI, J. M. BEATTIE, W. J. MCKENNA, H. WATKINS, POULTON, Lancet, 357 (2001) 1265. — 6. CHARCHAR, F. J., M. TOMASZEWSKI, S. PADMANABHAN, B. LACKA, M. N. UPTON, G. C. INGLIS, N. H. ANDERSON, A. MCCONNACHIE, E. ZUKOWSKA-SZCZECHOWSKA, W. GRZESZCZAK, J. M. CONNELL, G. C. WATT, A. F. DOMINICZAK, Hypertension, 39 (2002) 353. — 7. HAGBERG, J. M., K. R. WILUND, R. E. FERRELL, Physiol. Genomics, 4 (2000) 101. — 8. LLOYD-JONES, D. M., M. G. LARSON, A. BEISER, D. LEVY, Lancet, 353 (1999) 89. — 9. MURRAY, C. J. L., A. D. LOPEZ, Lancet, 349 (1997) 1269. — 10. BENTZEN, J., T. JORGENSEN, M. FENGER, Clin. Genet., 61 (2002) 126. — 11. CAVALLI-SFORZA, L. L., P. MENOZZI, A. PIAZZA: The history and geography of human genes. (Princeton University Press, Princeton, 1994). — 12. RENFREW, C., Cambridge Archaeol. J., 1 (1991) 3. — 13. VONA, G., P. MORAL, M. MEMMI, M. E. GHIANI, L. VARESI, Am. J. Phys. Anthropol., (2002) (in press). — 14. SIMONI, L., P. GUERESI, D. PETTENER, D. BARBUJANI, Hum. Biol., 71 (1999) 399. — 15. JACKES, M., D. LUBELL, C. MEIKLEJOHN, Curr. Anthropol., 38 (1997) 839. — 16. ARNAIZ-VILLENA, A., J. MARTINEZ-LASO, J. ALONSO-GARCIA, Hum. Biol., 71 (1999) 725. — 17. COMAS, D., F. CALAFELL, N. BENCHEMSI, A. KELAL, G. LEFRANC, M. STONEKING, M. A. BATZER, J. BERTRANPETIT, A. SAJANTILLA, Hum. Genet., 107 (2000) 312. — 18. SANCHEZ-MAZAS, A.: Prehistoric Iberia. (Plenum Press, London, 2000). — 19. ARNAIZ-VILLENA, A., J. MARTINEZ-LASO, E. GOMEZ-CASADO, N. DIAZ-CAMPOS, P. SANTOS, A. MARTINHO, H. BREDA, Immunogenetics, 47 (1997) 37. — 20. BOSCH, E., F. CALAFELL, F. R. SANTOS, A. PEREZ-LEZAUN, D. COMAS, N. BENCHEMSI, C. TYLER-SMITH, J. BERTRANPETIT, Am. J. Hum. Genet., 65 (1999) 1623. — 21. RANDO, J. C., F. PINTO, A. M. GONZALEZ, M. HERNANDEZ, J. M. LARRUGA, V. M. CABREARA, H. J. BANDELT, Ann. Hum. Genet., 62 (1998) 531. — 22. HARICH, N., E. ESTEBAN, A. LOPEZ-ALOMAR, G. VONA, P. MORAL, Ann. Hum. Biol., 29 (2002) 473. — 23. KANDIL, M., P. MORAL, E. ESTEBAN, L. AUTORI, D. ZAOUI, C. CALO, L. VACCA, G. VONA, Hum. Biol., 71 (1999) 791. — 24. BOSCH, E., F. CALAFELL, A. PEREZ-LEZAUN, J. CLARIMON, D. COMAS, E. MATEU, R. MARTINEZ-ARIAS, B. MORERA, Z. BRAKEZ, O. AKHAYAT, A. SEFINAI, A. CHAMBÓN-THOMSEN, J. BERTRANPETIT, Eur. J. Hum. Genet., 8 (2000) 360. — 25. SCOZZARI, R., F. CRUCINAI, A. PANGRAZZIO, P. SANTOLAMAZZA, G. VONA, P. MORAL, V. LATINI, L. VARESI, M. MEMMI, V. ROMANO, M. GENNARELLI, J. JARUSELSKA, R. WILLEMS, J. PARIK, C. TYLER-SMITH, M. JOBLING, A. TORRONI, Hum. Immunol., 62 (2001) 871. — 26. VALVENY, N.: Genetic risk factor for heart disease: Polymorphisms of the lipid metabolism genes. (University of Barcelona, Spain, 2000). — 27. LOPEZ-ALOMAR, T.: Polymorphisms of the hemostasis genes: Population variation in Western Mediterranean and association with risk for heart disease. (University of Barcelona, Spain, 2002). — 28. WEIR, B. S., Genetic data analysis II. (Sinauer Publ., Sunderland, 1996). — 29. NEI, M.: Molecular evolutionary genetics. (Columbia University Press, New York, 1987). — 30. SCHNEIDER, S., D. ROESSLI, L. EXCOFFIER, Arlequin ver. 2.000: A software for population genetics data analysis. (University of Geneva, Switzerland, 2000). — 31. RAYMON, M., F. ROUSSET, J. Hered., 86 (1995) 248. — 32. REYNOLDS, J., B. S. WEIR, C. C. COCKERMAN, Genetics, 195 (1983) 767. — 33. FELSESTEIN, J., Cladistics, 5 (1989) 164. — 34. BERTORELLE, G., L. EXCOFFIER, Mol. Biol. Evol., 15 (1998) 1298. — 35. CORBO, R. M., R. SCACCHI, Ann. Hum. Genet., 63 (1999) 301. — 36. CORBO, R. M., R. SCACCHI, L. MUREDDU, G. MULAS, S. CASTRECHINI, P. RIVASI, Hum. Biol., 6 (1999) 933. — 37. SEIXAS, S., M. J. TROVOADA, J. ROCHA, Hum. Biol., 6 (1999) 1001. — 38. DESTRO-BISOL, G., C. CAPELLI, M. BELLEDI, Hum. Biol., 5 (2000) 733. — 39. HARICH, N., E. ESTEBAN, A. LOPEZ-ALOMAR, A. CHAFIK, P. MORAL Clin. Genet., 62 (2002) 240. — 40. GIRALDO, M. P., E. ESTEBAN, M. P. ALUJA, R. M. NOGUES, C. H. BACKES-DURO, J. M. DUGOUJON, P. MORAL, Ann. Hum. Genet., 65 (2001) 537. — 41. MORELLI, L., M. G. GROSSO, G. VONA, L. VARESI, A. TORRONI, P. FRANCALACCI. Hum. Biol., 72 (2000) 585. — 42. MORAL, P., G MAROGNA, M. SALIS, V. SUCCA, G. VONA, Am. J. Phys. Anthropol., 93 (1994) 441. — 43. RICKARDS, O., C. MARTINEZ-LABARGA, G. SCANO, G. F. DE STEFANO, G. BIONDI, G. PACACI, H. WALTER, Hum. Biol., 7 (1998) 669. — 44. VARESI, L., M. MEMMI, M. C. CRISTOFARI, G. E. MAMELI, C. M. CALO, G. VONA, Am. J. Hum. Biol., 12 (2000) 339. — 45. GOMEZ-CASADO, E., P. MORAL, J. MARTINEZ-LASO, A. GARCIA-GOMEZ, L. ALLENDE, C. SILVERA-REDONDO, J. LONGAS, M. GONZALEZ, M. KANDIL, J. ZAMORA, A. ARNAIZ -VILLENA, Tissue Antigens, 55 (2000) 239. — 46. MORAL, P., M. KANDIL, D. ZAOUI, F. LUNA, A. LOPEZ, A. FERNÁNDEZ, E. ESTEBAN, N. VALVENY: Tendencias actuales de investigación en la antropología física española. (University of Leon Press, Spain, 2000). — 47. DIOS, S., J. R. LUIS, J. C. CARRIL, J. L. CAEIRO, Hum. Biol., 73 (2001) 675. — 48. FERNÁNDEZ-SANTADER, A., M. KANDIL, F. LUNA, P. MORAL, Hum. Biol., 74 (2002) 695.

P. Moral

Division of Anthropology, Department of Animal Biology, University of Barcelona, Av. Diagonal 645, 08028 Barcelona, Spain

MOLEKULARNE VARIJACIJE FUNKCIONALNIH GENA I POVIJEST LJUDSKIH POPULACIJA – PODACI O KANDIDATSKIM GENIMA ZA ČIMBENIKE RIZIKA ZA RAZVOJ KARDIOVASKULARNIH BOLESTI NA MEDITERANU

S A Ž E T A K

Na uzorcima 6 populacija zapadnog Mediterana (Iberijski poluotok, Maroko, otoci središnjeg Mediterana) analizirana su 22 polimorfizma DNK molekule. Analizirani biljezi smješteni su na polimorfnim lokusima nekoliko kandidatskih gena za čimbenike rizika za razvoj kardiovaskularnih bolesti, uključujući apolipoproteine i njihove receptore (APOA1, APOB, APOE, APOC1, APOC2, LPA i LDLR), gene uključene u regulaciju hemostaze (faktor VII, α i β-fibrinogen, α i β platelet-integrin, tkivni aktivator plazminogena i inhibitor-1 aktivatora plazminogena) te gen enzima angiotenzin konvertaze. Prikazani su rezultati djelomične analize koja je provedena na nekoliko populacijskih uzoraka iz svake regije (6 s Iberijskog poluotoka, 2 iz Maroka te 3 sa središnjih otoka). Stupanj međupopulacijske različitosti bio je značajan i sukladan podacima drugih genetičkih polimorfizama. Udio varijance učestalosti alela podupire zemljopisnu strukturu u tri osnovne regije: središnji mediteranski otoci, Iberijski poluotok i Sjeverna Afrika. Genetičke udaljenosti podupiru sjevernoafričke utjecaje od juga prema sjeveru Iberijskog poluotoka, te značajan tok gena iz subsaharske Afrike u Maroko. Što se tiče epidemioloških osobitosti, Sjevernu Afriku karakterizira visoka učestalost LPA PNR alela s visokim brojem ponavljajućih sljedova (zaštitnih za kardiovaskularni rizik) i visokom učestalošću APOE*E4 alela (čimbenika rizika) u usporedbi s europskim populacijama.