

ARTICLE

Genetic isolates in Corsica (France): linkage disequilibrium extension analysis on the Xq13 region

Veronica Latini¹, Gabriella Sole¹, Silvia Doratiotto¹, Daniela Poddie¹, Marc Memmi², Laurent Varesi², Giuseppe Vona³, Antonio Cao¹ and Maria Serafina Ristaldi^{*1}

¹*Istituto di Neurogenetica e Neurofarmacologia Consiglio Nazionale delle Ricerche (INN-CNR), Cagliari, Italy;*
²*Università de Corte, Corsica, France;* ³*Dipt. Biologia Sperimentale, Università di Cagliari, Italy*

Genetic isolates with a history of a small founder population, long-lasting isolation and population bottlenecks represent exceptional resources in the identification of genes involved in the pathogenesis of multifactorial diseases. In these populations, the disease allele reveals linkage disequilibrium (LD) with markers over significant genetic intervals, therefore facilitating disease locus identification. This study has been designed to examine the background LD extension in some subpopulations of Corsica. Our interest in the island of Corsica is due to its geographical and genetic proximity to the other Mediterranean island of Sardinia. Sardinian isolates in which the extension of the background LD is particularly high have been recently identified and are now the object of studies aimed at the mapping of genes involved in complex diseases. Recent evidence has highlighted that the genetic proximity between the populations of Corsica and Sardinia is particularly true for the internal conservative populations. Given these considerations, Sardinia and Corsica may represent a unique system to carry out parallel association studies whose results could be validated by comparison. In the present study, we have analyzed the LD extension on the Xq13 genomic region in three subpopulations of Corsica: Corte, Niolo and Bozio, all located in the mountainous north-center of the island. Our results show a strong degree of LD over long distance for the population of Bozio and to a less extent for the population of Niolo. Their LD extent is comparable to or higher than that reported for other isolates.

European Journal of Human Genetics (2004) 12, 613–619. doi:10.1038/sj.ejhg.5201205

Published online 28 April 2004

Keywords: linkage disequilibrium extension; subisolates; Corsica; Sardinia

Introduction

Recent reports have shown that in isolated and conservative populations (isolates) the level of linkage disequilibrium (LD), the nonrandom association of alleles at closely linked loci, is particularly high.^{1–7}

The design and feasibility of genome-wide association studies are critically dependent on the extent of LD. The

higher the extension of LD, the lower are the number of polymorphism markers that have to be examined in order to find an association with a particular disease. Additionally, rare mutations that are diluted in 'open' populations can have high frequencies due to endogamy and founder effect in these isolates.

This study has been designed to examine the extension of the LD on a region of the X chromosome (Xq13), as a measure of the background LD, in some populations of the island of Corsica. Even though the LD extension on the region Xq13 is not representative of the average level of LD of the entire genome,⁸ the analysis of this region has been widely used as a measure of a general LD in a given

*Correspondence: Dr MS Ristaldi, INN-CNR, Via Boccaccio 8, 09047 Selargius (CA), Sardinia, Italy. Tel: +39 070 540342; Fax: +39 070 540722; E-mail: ristald@mcweb.unica.it

Received 6 June 2003; revised 18 February 2004; accepted 20 February 2004

population and to compare the levels of LD between populations.^{1–7,9}

The interest on the isolates of the island of Corsica is based on its closeness to the island of Sardinia, which is the object of several studies, carried out by different groups, aimed at the mapping of disease-related genes. Recent data have shown a genetic closeness between the two islands. The genetic proximity of the populations of Sardinia and Corsica has been inferred on the basis of the gene frequency of serum proteins, isozymes, as well as mitochondrial control region sequence variations. All the data show that Sardinians and Corsicans are genetically much closer to each other than to any other Mediterranean population.^{10–15} Corsica is the third largest island of the Western Mediterranean sea. It is very mountainous with peaks reaching 2710 m in altitude. Until Pleistocene, it was physically linked to Sardinia in a single geological block. About 11 000 years ago, the two islands fell apart and were separated by the strait called 'Bocche di Bonifacio'.

According to mitochondrial DNA sequence variations,¹⁵ the Sardinian-Corsica block was peopled in a period between 14 000 and 78 000 years ago (Paleolithic period), through the Tuscany island during the last glaciation (Wurm), when the sea level was lower. After the physical separation (during the Pleistocene period), the population of the two islands fell apart even though a reduced, but constant genetic flow remains between north Sardinia and southern Corsica. For both islands, genetic drift, isolation and low population numbers have played a strong part in their genetic shaping. Sardinia and Corsica were invaded several times, often by the same populations. In the great majority of cases, these invasions were limited to the coast and left slight marks on the gene pool of the native populations. Strong evidence also suggests an internal microgeographic diversity inside Sardinia and Corsica,

with the most conserved population located in the center of the two islands on the mountainous regions.^{4,16–20} The internal conserved populations of Sardinia and Corsica are also genetically closer between the two islands.^{19,20}

Materials and method

The three populations under investigation in the present study belong to an area of northern central Corsica of 14 000 inhabitants (Figure 1). Corte is a small city with about 6000 inhabitants. Bozio and Niolo are geographic areas situated East and West of Corte, respectively. Bozio and Niolo are constituted of several small villages. The total population of Niolo and Bozio is 2400 and 1400 inhabitants, respectively. DNA samples of unrelated male individuals were collected in the region of Niolo ($n=49$), Bozio ($n=51$) and Corte ($n=50$). Samples have been analyzed using seven dinucleotide microsatellite markers on chromosome Xq13.3: DXS983, DXS986, DXS8092, DXS8082, DXS1225, DXS8037 and DXS995. Microsatellites were analyzed by using an ABI Prism 377 DNA analyzer.¹ Genotypes were processed by Genescan v3.1 and Genotype v2.5 software.

The nonrandom allelic association between pairs of microsatellite loci has been tested by an extension of Fisher's exact test on contingency tables.²¹ P -values have been corrected by the step-down Holm–Sidak procedure²² with the formula $P_{corrected} = 1 - (1 - P)^n$, where n is the number of P -values smaller or equal to that being corrected. With this procedure, correction results in a probability with a distribution from 0 to 1, unlike the Bonferroni correction, usually used in multiple testing corrections, in which the corrected P -value can exceed 1. In the present paper, we have re-applied the step-down Holm–Sidak procedure²² to the published P -values of Gavoi,² Talana⁴ and Urzulei.⁵ We were induced to do so

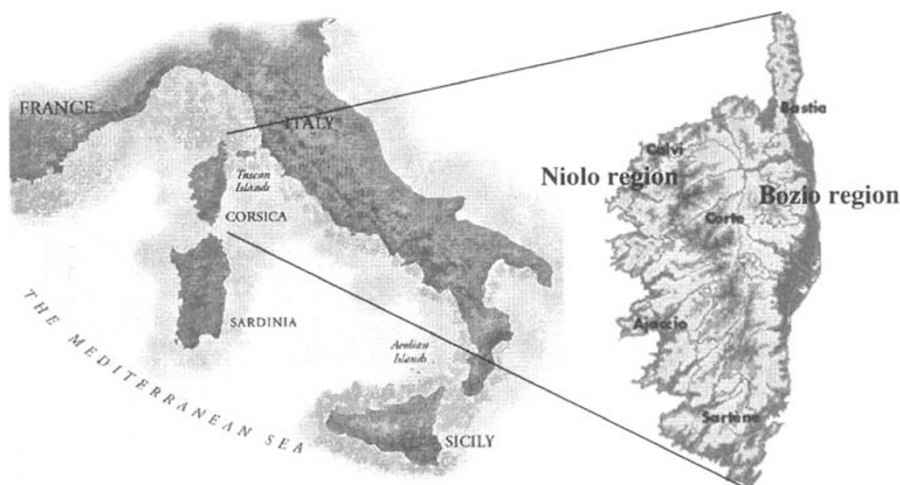


Figure 1 Corsica isolates and West Mediterranean Basin.

Table 1 Pairwise LD between markers for Bozio, Niolo and Corte

Locus pair		Distance (cM)	Distance (Mb)	Corte (50)	Niolo (49)	Bozio (51)	Urzulei (49) ^a	Talana (50) ^b	Gavoi (73) ^c
DXS983	DXS995	4.0	12.5	0.060	0.412	0.090			0.046
DXS986	DXS995	2.5	1.0	0.965	0.042	0.026			0.042
DXS8092	DXS995	2.4	7.0	0.323	0.896	0.019			0.308
DXS8082	DXS995	2.3	3.5	0.444	0.001	0.573			0.021
DXS983	DXS1225	2.0	9.5	0.868	0.552	0.006	0.733	0.158	0.017
DXS8037	DXS995	2.0	1.0–3.0	0.063	0.004	0.000			0.014
DXS983	DXS8037	2.0	9.5–11.5	0.605	0.520	0.682	0.257	0.002	0.000
DXS1225	DXS995	2.0	3.0	0.934	0.005	0.102			0.005
DXS983	DXS8082	1.7	9.0	0.944	0.961	0.286	0.130	0.019	0.245
DXS983	DXS8092	1.6	5.5	0.813	0.234	0.589	0.003	0.050	0.000
DXS983	DXS986	1.5	11.5	0.285	0.245	0.130	0.063	0.145	0.003
DXS8037	DXS986	0.5	<2	0.215	0.599	0.017	0.976	0.011	0.003
DXS1225	DXS986	0.5	2.0	0.037	0.156	0.001	0.034	0.006	0.000
DXS8092	DXS1225	0.4	4.0	0.678	0.145	0.902	0.052	0.000	0.000
DXS8092	DXS8037	0.4	4.0–6.0	0.504	0.460	0.131	0.396	0.629	0.000
DXS8082	DXS8037	0.3	0.5–2.5	0.910	0.014	0.253	0.490	0.016	0.004
DXS8082	DXS1225	0.3	>0.5	0.000	0.000	0.000	0.000	0.000	0.000
DXS8082	DXS986	0.2	2.5	0.428	0.120	0.002	0.015	0.642	0.000
DXS8092	DXS986	0.1	6.0	0.640	0.542	0.091	0.159	0.056	0.000
DXS8092	DXS8082	0.1	3.5	0.685	0.381	0.029	0.231	0.005	0.000
DXS1225	DXS8037	0.0	<2	0.067	0.428	0.003	0.320	0.000	0.008

^aAngius et al.⁵^bAngius et al.⁴^cZavattari et al.²

P-values shown in bold are lower than 0.05 after correction for multiple testing by using the Holm–Sidak procedure.¹⁸ The correction of *P*-values in Gavoi, Urzulei and Talana was recalculated to amend the mistakes by Angius et al.^{4,5} and Zavattari et al.²

by the paper of Katoh et al.,⁷ 2002, where the procedure has been re-applied on the data of Zavattari et al.,² showing that the Holm–Sidak method could be inaccurately applied. Indeed the number of *P*-values remaining lower than 0.05 (bold in Table 1) are fewer than those reported.^{2,4,5}

Each *P*-value has been considered significant when <0.05 and suggestive when $0.05 < P < 0.10$. LD has been measured by the normalized disequilibrium coefficient D' ²³ between each marker loci pairs. As small sample size can overestimate the values of D' , we also reported estimated D' corrected by a bootstrap procedure.²⁴ Disequilibrium across each locus was plotted using GOLD software²⁵ (<http://www.sph.umich.edu/csg/abecasis/GOLD/>). In order to test the null hypothesis that the allelic distribution is identical across the populations under investigation, we have carried out a genetic differentiation test using the Genepop 3.3 software (<ftp://ftp.cfe.cnrs-mop.fr/pub/pc/msdos/genepop>).²⁶ For each locus, the test is carried out on a contingency table. The method, based on allelic distribution of alleles in various samples, is described by Raymond and Rousset.²⁷ In the present work, the test is performed automatically for all pairs of populations for all loci.

Results

Table 1 reports the significance of nonrandom allelic association between pairs of microsatellite loci by the pairwise LD based on Fisher's exact test. The data in bold

are those remaining lower than 0.05 after correction using the Holm–Sidak method.²² The data are compared to those of some Sardinian genetic isolates reported in literature.^{2,4,5} The population of Bozio shows 10 out of 21 pairs with significant LD ($P < 0.05$) and two suggestive ($0.05 < P < 0.10$). The largest distance between markers with a significant *P* is 9.5 megabases (Mb) (DXS983–DXS1225). After correction, using the Holm–Sidak method, Bozio shows five pairs with significant LD, with the largest distance between markers with a significant *P* at 2.5 Mb (DXS8082–DXS986). The population of Niolo shows six pairs with significant LD, and two pairs after correction. The largest distance between markers with a significant *P* is 3.5 Mb (DXS8082–DXS995).

The population of Corte shows two pairs with significant LD and three suggestive. The largest distance between markers with a significant *P* is 2 Mb (DXS1225–DXS986). After correction, only one pair remains significant at a distance of 0.5 Mb (DXS8082–DXS1225).

Data for Bozio and Niolo are similar to those reported by others for some Sardinian genetic isolates such as Talana and Urzulei (Table 1).^{4,5} The other Sardinian isolate examined to date, Gavoi, shows instead a greater number of significant pairwise comparisons (Table 1).²

Figure 2 shows the extent of the LD graphically, between pairwise microsatellite markers inside the genomic region under investigation in the population of Niolo, Corte and Bozio. Red indicates $D' = 1$; dark blue $D' = 0$. As shown in Figure 2, the population of Niolo shows the highest *D'*

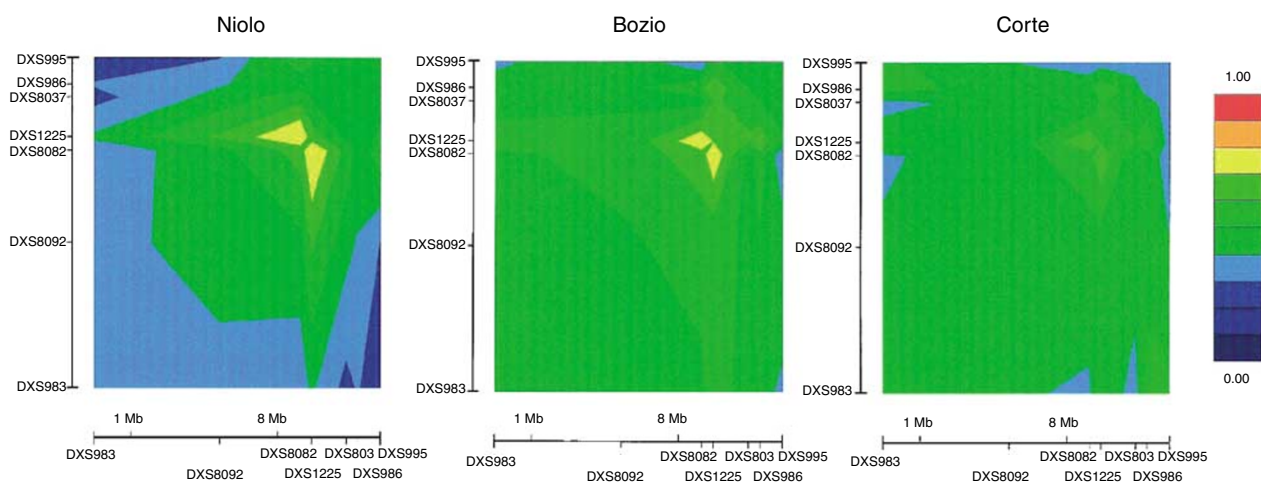


Figure 2 Distribution of LD for all the studied markers. Distances (Mb) are those reported in Table 1. The figure was drawn by the GOLD program (v. 1.1): colors reflect D' values from red ($D' = 1$) to deep blue ($D' = 0$).

Table 2 Original D' values and D' derived by use of a bootstrap correction for all the isolates studied

Locus pair	NIOLO		CORTE		BOZIO	
	D'	Corrected D'	D'	Corrected D'	D'	Corrected D'
DXS983 DXS8092	0.378	0.260	0.411	0.263	0.435	0.336
DXS983 DXS8082	0.327	0.209	0.389	0.249	0.481	0.404
DXS983 DXS1225	0.421	0.303	0.478	0.349	0.605	0.541
DXS983 DXS8037	0.267	0.136	0.364	0.236	0.404	0.333
DXS983 DXS986	0.311	0.213	0.526	0.419	0.445	0.357
DXS983 DXS995	0.214	0.105	0.525	0.464	0.381	0.331
DXS8092 DXS8082	0.475	0.362	0.447	0.326	0.533	0.433
DXS8092 DXS1225	0.618	0.551	0.433	0.309	0.513	0.382
DXS8092 DXS8037	0.442	0.342	0.451	0.359	0.494	0.418
DXS8092 DXS986	0.393	0.275	0.469	0.358	0.473	0.360
DXS8092 DXS995	0.299	0.173	0.411	0.327	0.486	0.416
DXS8082 DXS1225	0.809	0.760	0.686	0.622	0.781	0.727
DXS8082 DXS8037	0.560	0.505	0.408	0.318	0.464	0.393
DXS8082 DXS986	0.461	0.368	0.470	0.370	0.586	0.526
DXS8082 DXS995	0.566	0.527	0.350	0.235	0.316	0.200
DXS1225 DXS8037	0.483	0.379	0.505	0.422	0.615	0.541
DXS1225 DXS986	0.413	0.284	0.583	0.501	0.664	0.602
DXS1225 DXS995	0.531	0.468	0.370	0.263	0.515	0.435
DXS8037 DXS986	0.432	0.329	0.442	0.340	0.498	0.438
DXS8037 DXS995	0.455	0.392	0.360	0.285	0.448	0.405
DXS986 DXS995	0.425	0.362	0.308	0.172	0.459	0.402

D' values greater than 0.5 are in bold.

value (0.81; Table 2); however, the D' values decrease with the distance. The population of Bozio shows a maximum D' value (0.78; Table 2), which is slightly lower to that of Niolo. Nevertheless, the D' values decline at a much lower rate with the distance and remain at intermediate values, higher than 0.5 for a large part of the Xq13 region. On the contrary, the maximum D' value for the population of Corte (0.65; Table 2) is lower than those of Bozio and Niolo, decreasing to under 0.5 for the most part of the interval.

In order to visualize and compare the strength of LD over distance among Sardinian and Corsican isolates, Figure 3 shows the LD trend along 10Mb of the X chromosome region under investigation for the three populations examined and in the Sardinian isolates of Talana (Figure 3a) and Gavoi (Figure 3b). D' value of 1 indicates complete LD, 0 indicates no LD. The degree of LD needed for effective mapping depends on several factors, nevertheless D' values higher than 0.5 are considered useful.²⁸

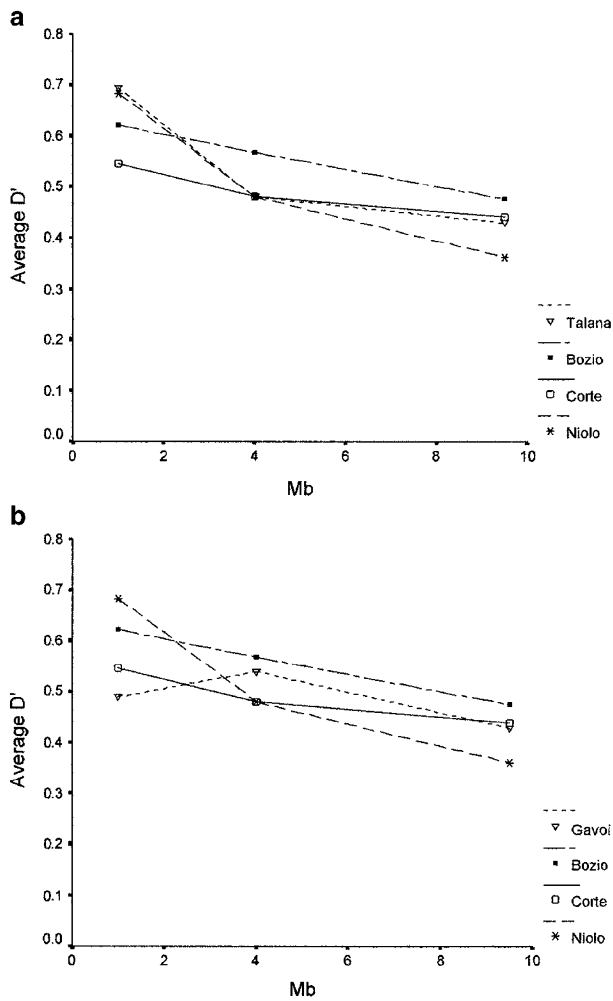


Figure 3 Comparison of the LD extension evaluated as average multiallelic D' values versus stratified physical distances for the chromosome Xq13 region in Corsica and Sardinia isolates. Each point represents the average of pairwise comparisons of six loci. (a) Average D' values have been calculated without marker DXS995 to allow comparison with the Sardinian isolate of Talana as reported by Angius *et al.*⁴ (b) Average D' values were calculated including marker DXS995 to allow comparison with the Sardinian isolate of Gavoi as reported by Zavattari *et al.*²

In Figure 3a, the average D' values over the distance have been calculated excluding the pairwise D' values of the marker DXS995, which have not been analyzed by Angius *et al.*⁵

In Figure 3b, the average D' values over the distance have been calculated including marker DXS995 to allow comparison with the genetic isolate of Gavoi as reported by Zavattari *et al.*² The average D' values for the population of Niolo, Bozio and Corte are slightly lower than those of Figure 2a obtained excluding marker DXS995.

Table 3 Genetic differentiation for each population pair

Population pair	χ^2	P-value
Niolo–Corte	39.54	0.00030
Niolo–Bozio	18.11	0.20199
Corte–Bozio	46.53	0.00002

P-values for each population pair across all loci (Fisher's method).

We could not compare the Sardinian isolate of Urzulei to our samples since D' data are not published.⁵

In summary, Figure 3 a and b shows that average D' values along 10 Mb for the central isolated population of Corsica and those of central Sardinia, reported by others, are similar with the exception of Corte (see below). In particular, the populations of Bozio and Gavoi show average D' values close to or higher than 0.5 over very long distances (Figure 3b), while those of Niolo and Talana show high values of average D' for relatively short distances (1 Mb) with a decrease over distance (Figure 3a).

It is noteworthy that the population of Gavoi shows greater number of significant P-values and lower average D' values than other populations such as Bozio. This difference is most likely due to the larger sample size used in the work of Zavattari *et al.*² Indeed the power of Fisher's test and D' values are affected by sample size.²⁹

In order to evaluate the bias of D' estimates due to small sample size, we corrected the estimated D' value by use of the bootstrap procedure. As shown in Table 2, there is a general decrease in D' values after correction. Nevertheless, most of the D' values greater than 0.5 are still higher than 0.5 after correction. Only the D' values that were very close to 0.5 before correction go below this value afterwards. These data confirm the high degree of LD in the populations of Niolo and Bozio.

Finally, we have also carried out a genetic differentiation test, by Genepop 3.3, based on the allelic distribution of alleles in the three samples. Results (Table 3) show no differentiation between the populations of Niolo and Bozio ($P=0.2$), whereas both populations show significant P-values against Corte ($P=3 \times 10^{-4}$ and 2×10^{-5} for Niolo and Bozio, respectively).

Discussion

Our results show a high degree of LD for the population of Bozio and Niolo, and a lower degree for Corte. This high degree of LD has most likely been created by genetic drift and has been maintained by isolation and slow growth of the populations. Corte shows a lower degree of LD. This result is most likely explained by the fact that Corte, being the historical capital of Corsica, has been relatively more open to genetic flow compared to Niolo and Bozio regions. Very little is known about the biodemographic history of

the populations under investigation, since data based on the ecclesiastical records are not available.

The Niolo region is located in a mountainous area in the central northwest of Corsica, is composed of several villages and its population size has been relatively constant, with a very slow growth rate in the last decades. Its economy is mainly based on the presence of snow resorts. The Bozio region is located in a mountainous region in the central northeast of Corsica and is composed of several villages, very close to each other, which are being depopulated owing to migration toward the coast and outside the island owing to the poor economy of the area. The central area of the island, on the whole, has been the stronghold of the Corsican indigenous population against the different invaders.

Genetic differentiation test shows no differentiation among the populations of Niolo and Bozio, which are both different from Corte (Table 3). We believe that there has not been a significant genetic flow between the populations of Bozio and Niolo and Corte since populations are separated by strong natural barriers. The genetic proximity between Bozio and Niolo, pointed out by the genetic differentiation test, is most likely explained by the fact that the three populations derive from a common ancestral genetic pool, which has been lost in the population of Corte, more open, and has probably been preserved in the populations of Bozio and Niolo.

Isolates have been suggested to be useful in the identification of genetic regions involved in common diseases.^{30–32} In this context, it is worth noting that the general population of the island of Sardinia is considered suitable to carry out association studies aimed at the identification of genes involved in the pathogenesis of complex diseases, given the well-documented genetic isolation¹⁸ and the high frequency of some genetic diseases.^{33–36} Nevertheless, Eaves *et al*⁹ have shown that background LD extension on the general population of Sardinia is not very high and comparable to that of the general population of the United Kingdom. The reason for such apparent contradiction is to be found in the earlier-mentioned peculiar genetic structure of the Sardinian, as well as Corsica, population, which is characterized by an extraordinary degree of microdifferentiation.^{12,17,20} To the observed internal variability has contributed the following: (1) the internal geographical barriers; (2) the strict isolation and the accompanying great level of endogamy and inbreeding; and (3) the endemic presence of malaria and other diseases as well as famines, which exerted a strong selective pressure. Nevertheless, it is possible to delimitate inside the two islands homogeneous areas with high LD level as reported here and by several authors.^{12,17}

In approaching these studies, we need to bear in mind that the identification of a genomic region associated to a disease must be confirmed in more than one subisolate for at least two reasons: (1) different loci may be responsible

for the same disease phenotype in different populations and (2) a casual genetic identity must be distinguished from a true identity by descent.^{37–39} To this end, the internal founding populations of Corsica and Sardinia may represent an interesting system to validate association study results.²⁹ The two populations derive from a common founding genetic pool, and similar evolutionary forces, such as isolation, consanguinity and bottleneck due to famines and epidemics, have shaped them over time. In addition, the two populations share a similar dietary regime and climate (Mediterranean). All these considerations suggest that they may have 'selected' the same kind of allele associate to particular common diseases.

Populations with high LD extension are well suited for a rough LD mapping of extended genomic regions, whereas they are probably unsuited for fine mapping since markers found far away from the disease-associated locus may also show a significant association with the disease genes. Therefore, we suggest a multistep procedure as proposed by Kaessmann *et al*⁶ in attempting to use isolates of the Sardinia and Corsica Islands in association studies:

- Identification of a large genomic region containing the disease-associated locus in a small subisolated population of the central region of Corsica and replication in a small subisolated population of central Sardinia.
- The disease-associated region could be more finely mapped in a recently expanded population.
- Fine mapping could be carried out in open populations in which the extent of LD is low, such as the general population of Sardinia^{4,5} and possibly Corsica (the LD extension on Xq13.3 in the general population of Corsica is under investigation).
- Open population in which the level of background LD is very low, such as the African population, could be used for the final mapping, in agreement with the common diseases/common variant hypothesis and the 'out of Africa' theory of human evolution.⁴⁰

In this study, we have used microsatellite markers to assess background LD. The use of microsatellite markers in LD mapping is being substituted by single nucleotide polymorphisms. Nevertheless, microsatellite markers remain a valuable tool for the first screening of background LD in a given population, as shown by this study and by others,^{1–7} and can be useful for studies of complex traits in isolated populations where the extent of LD is particularly high.²⁹

In conclusion, our data indicate that the population of the central region of Corsica could be well suited for the mapping of genes involved in the pathogenesis of complex diseases. We also believe that Sardinia and Corsica and their isolates could represent an interesting system to carry out association studies given the common genetic derivation and similarity of selective pressures.

Acknowledgements

We thank Traci Primm for reviewing the English version of this manuscript. This work has been supported by Ministero della Sanità, Italy. Progetti speciali art. 12 bis comma 6, decreto legislativo D. lgs 229/99 and Diagnosi, prevenzione e terapia di malattie monogeniche e multifattoriali nella popolazione Sarda, Regione sardecna

References

- 1 Laan M, Paabo S: Demographic history and LD in human population. *Nat Genet* 1997; **17**: 435–438.
- 2 Zavattari P, Deidda E, Whalen M et al: Major factors influencing LD by analysis of different chromosome regions in distinct populations: demography, chromosome recombination frequency and selection. *Hum Mol Genet* 2000; **9**: 2947–2957.
- 3 Varilo T, Laan M, Hovatta I, Wiebe V, Terwilliger JD, Peltonen L: Linkage disequilibrium in isolated populations: Finland and young sub-population of Kuusano. *Eur J Hum Genet* 2000; **8**: 604–612.
- 4 Angius A, Melis MP, Morelli L et al: Archival, demographic and genetic studies define a Sardinian sub-isolate as a suitable model for mapping complex traits. *Hum Genet* 2001; **109**: 198–209.
- 5 Angius A, Bebbere D, Petretto E et al: Not all isolates are equal: linkage disequilibrium analysis on Xq13.3 reveals different patterns in Sardinian sub-populations. *Hum Genet* 2002; **111**: 9–15.
- 6 Kaessmann H, Zollner S, Gustaffson A et al: Extensive LD in small human populations in Eurasia. *Am J Hum Genet* 2002; **70**: 673–685.
- 7 Katoh T, Mono S, Ikuta T et al: Genetic isolates in East Asia: a study of linkage disequilibrium in the X chromosome. *Am J Hum Genet* 2002; **71**: 395–400.
- 8 Taillon-Miller P, Bauer-Sardina I, Saccone NL et al: Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. *Nat Genet* 2000; **25**: 324–328.
- 9 Eaves IA, Merriman T, Barber RA et al: The genetically isolated populations of Finland and Sardinia may not be a panacea for linkage disequilibrium mapping of common disease genes. *Nat Genet* 2000; **25**: 320–323.
- 10 Vona G, Memmi M, Varesi L, Mameli GE, Succa V: A study of several genetic markers in the Corsican population (France). *Anthropol Anz* 1995; **53**: 125–132.
- 11 Varesi L, Vona G, Memmi M, Marongiu F, Ristaldi MS: β -Thalassemia mutations in Corsica. *Hemoglobin* 2000; **24**: 239–244.
- 12 Vona G, Memmi M, Calò CM et al: Genetic structure of the Corsican population (France): a review; in: Research Signpost (eds): Recent Research Developments in Human Genetics: 2002, Vol 1, pp 147–164.
- 13 Grimaldi MC, Crouau-Roy B, Contu L, Amoros JP: Molecular variation of HLA class I genes in the corsican population: approach to its origin. *Eur J Immunogenet* 2002; **29**: 101–107.
- 14 Vona G, Moral P, Memmi M, Ghiani ME, Varesi L: Genetic structure and affinities of the Corsican population (France): classical genetic markers analysis. *Am J Hum Biol* 2003; **15**: 151–163.
- 15 Varesi L, Memmi M, Cristofari MC, Mameli GE, Calò MC, Vona G: Mitochondrial control-region sequence in the corsican population (France). *Am J Hum Biol* 2000; **12**: 339–351.
- 16 Ranque J, Vicoli R, Silicati V et al: Le peuplement humain en Corse. Variété des populations insulaires. Etude sèro-antropologique. *Bull Soc Corse Biol Hum* 1961; **1**: 115–168.
- 17 Cappello N, Rendine S, Griffo R et al: Genetic analysis of Sardinia I data on 12 polymorphisms in 21 linguistic domains. *Ann Hum Genet* 1996; **60**: 125–141.
- 18 Vona G: The peopling of Sardinia (Italy): history and effects. *Int J Anthropol* 1997; **12**: 71–87.
- 19 Calò CM, Vacca L, Vona G: Analysis of selected microsatellites in 4 regions of Corsica (Francia). *Antropol Contemp* 1999; **87–98**, (Monografia).
- 20 Latini V, Vona G, Ristaldi MS et al: β -Globin cluster haplotypes in Corsica and Sardinia populations. *Hum Biol* 2003; **75**: 855–871.
- 21 Slatkin M: Linkage disequilibrium in growing and stable populations. *Genetics* 1994; **137**: 331–336.
- 22 Lautenberger JA, Stephens JC, O'Brien SJ, Smith MW: Significant admixture linkage disequilibrium across the FY locus in African Americans. *Am J Hum Genet* 2000; **66**: 969–978.
- 23 Lewontin RC: The interaction of selection and linkage. General considerations; heterotic models. *Genetics* 1964; **49**: 49–67.
- 24 Efron B, Tibshirani RJ: *An Introduction to the Bootstrap*. New York: Chapman & Hall; 1993.
- 25 Abecasis GR, Cookson WO: Gold – graphical overview of linkage disequilibrium. *Bioinformatics* 2000; **16**: 182–183.
- 26 Raymond M, Rousset F: GENEPOP (version 1.2): population genetics software for exact tests and ecumenisms. *J Hered* 1995; **86**: 248–249.
- 27 Raymond M, Rousset F: An exact test for population differentiation. *Evolution* 1995; **49**: 1280–1283.
- 28 Reich DE, Cargill M, Bolk S et al: Linkage disequilibrium in the human genome. *Nature* 2001; **411**: 199–204.
- 29 Mohlke KL, Lange EM, Valle TT et al: Linkage disequilibrium between microsatellite markers extends beyond 1cM on chromosome 20 in Finns. Choice in mapping genes for complex diseases. *Nat Genet* 1999; **23**: 394–404.
- 30 Ardlie KG, Kruglyak L, Seielstad M: Patterns of LD in the human genome. *Nat Genet* 2002; **3**: 299–309.
- 31 Marrosu MG, Murru R, Murru MR et al: Dissection of the HLA association with multiple sclerosis in the founder isolate population of Sardinia. *Hum Mol Genet* 2001; **11**: 1221–1226.
- 32 Sheffield VC, Stone EM, Carmi R: Use of isolated inbred human populations for identification of disease genes. *TIG Rev* 1998; **14**: 391–396.
- 33 Wright AF, Carothers AD, Pirastu M: Population multiple sclerosis in the founder isolated population of Sardinia. *Hum Mol Genet* 2001; **10**: 2907–2916.
- 34 Contu D, Morelli L, Zavattari P et al: Sex-related bias and exclusion mapping of the non recombinant portion of chromosome Y in human type 1 diabetes in the isolated founder population of Sardinia. *Diabetes* 2002; **51**: 3573–3576.
- 35 Angius A, Petretto E, Maestrale GB et al: A new essential hypertension susceptibility locus on chromosome 2p24–p25, detected by genomewide search. *Am J Hum Genet* 2002; **71**: 893–905.
- 36 Gianfrancesco F, Esposito T, Ombra MN et al: Identification of a novel gene and a common variant associated with uric acid nephrolithiasis in a Sardinian genetic isolate. *Am J Hum Genet* 2003; **72**: 1479–1491.
- 37 Risch NJ: Searching for genetic determinants in the new millennium. *Nature* 2000; **405**: 847–856.
- 38 Cardon LR, Bell JI: Association study designs for complex diseases. *Nat Rev* 2001; **2**: 91–99.
- 39 Heath S, Robledo R, Beggs W: A novel approach to search for identity by descent in small samples of patients and controls from the same mendelian breeding unit: a pilot study on myopia. *Hum Hered* 2001; **52**: 183–190.
- 40 Tishkoff SA, Williams SM: Genetic analysis of African populations: human evolution and complex disease. *Nat Rev* 2002; **3**: 611–621.