J Hum Genet (2007) 52:999–1010
DOI 10.1007/s10038-007-0206-6

ORIGINAL ARTICLE

The ins and outs of population relationships in west-Mediterranean islands: data from autosomal *Alu* polymorphisms and *Alu*/STR compound systems

E. Gonzalez-Perez · P. Moral · M. Via · G. Vona · L. Varesi · J. Santamaria · M. Gaya-Vidal · E. Esteban

Received: 25 July 2007 / Accepted: 28 September 2007 / Published online: 24 October 2007

Abstract The islands of the West Mediterranean have played a central role in numerous archaeological, historical and anthropological studies due to their active participation in the history of main Mediterranean civilisations. However, genetic data failed to fit in both their degree of internal differentiation and relationships. A set of 18 *Alu* markers and three short tandem repeats (STRs) closely linked to the CD4, F13B and DM *Alu* have been analysed in seven samples from Majorca, Corsica, Sardinia and Sicily to explore some of these issues. Our samples show a high genetic heterogeneity inside and among islands for the *Alu* data. Global differentiation among islands (F_{ST} 2.2%) is slightly higher than that described for Europeans and North Africans. Both the estimated divergence times among samples and the high population heterogeneity revealed by *Alu* data are compatible with population differences since the first islands' settlement in the Paleolithic period. However, the high within-population diversities and the remarkable homogeneity observed in both STR and *Alu*/STR haplotype variation indicated that, at least since Neolithic times, gene flow has been acting in west Mediterranean. Genetic drift in west-coast Sardinia and gene flow in west Sicily have contributed to their general differentiation, whereas Corsica, Majorca and east Sicily seem to reflect more recent historical relationships from continental south Europe.

Keywords *Alu* insertions · *Alu*/STR haplotypes · Human populations · Mediterranean peopling

Introduction

The Balearic archipelago and the islands of Sardinia, Corsica and Sicily are enclosed in the westernmost part of the Mediterranean basin by the Iberian Peninsula and the Strait of Gibraltar to the west, North Africa to the south and the Strait of Messina and the Italic Peninsula to the east. This region covers an area of about 0.85 million km^2 that embraces a set of populations closely related not only by geography but also by historical relationships, probably since their initial peopling in middle and upper Paleolithic times. Their common historical background includes numerous and almost continuous waves of settlements and conquests by several mainland civilisations. This coming and going of populations in a close geographical area has constituted a challenge for archaeologists, historians, ethnologists and anthropologists alike. The latter have been particularly interested in determining the degree of genetic relationships of a set of populations that, even having a

Electronic supplementary material The online version of this article (doi:10.1007/s10038-007-0206-6) contains supplementary material, which is available to authorised users.

E. Gonzalez-Perez · P. Moral · M. Via · J. Santamaria · M. Gaya-Vidal · E. Esteban
Unitat d'Antropologia, Departament de Biologia Animal, Universitat de Barcelona, Barcelona, Spain

G. Vona
Dipartamento di Biologia Sperimentale, Università di Cagliari, Cagliari, Italy

L. Varesi
Department of Human Genetics, Faculty of Science, University of Corsica, Corte, France

E. Esteban (✉)
Unitat d'Antropologia, Facultat de Biologia, Avda. Diagonal 645, 08028 Barcelona, Spain
e-mail: mesteban@ub.edu

common past of invaders for centuries, still preserve some remarkable differences.

In Majorca, the largest island of the Balearic archipelago, archaeological data suggest peopling since the Paleolithic period. The island was occupied by the Carthaginians before passing to the Romans, who installed a long period of prosperity. From 707, the island was increasingly attacked by Muslim raiders from North Africa. Two centuries later, the Caliphate of Cordoba conquered Majorca, ushering in a new period of prosperity for the island. In the thirteen century, the Catalano–Aragonese launched an invasion with 15,000 men and 1,500 horses, annexing the island to the kingdom. In the archipelago, the mother tongue is the Balearic variation of Catalan, a Romance language spoken in a large part of the former territories of this kingdom. From a genetic point of view, recent data from mtDNA haplotype variability (Picornell et al. 2005) suggest a high similarity among Majorca, other Balearic islands and Spanish populations historically related with the Catalano–Aragonese kingdom. This affinity points to an important gene flow from the mainland without significant bottlenecks involved in the colonisation of the island.

Corsica and Sardinia formed a single land mass in early Paleolithic times. They are now separated by a straight of about 12 km wide. As a result of this close vicinity, these two islands share a common background despite the fact that for many centuries, the contact with different Mediterranean invaders was apparently limited to the coastal flatland territories of the islands. Carthaginians and Romans pushed the indigenous people into the central region of the islands, which explains the fact that, in the case of Sardinia, the centre is the most conservative region linguistically and genetically (Piazza et al. 1988; Cappello et al. 1996). In regards to the vernacular languages, the inhabitants of both islands speak Romance languages, but in the case of Corsica, the language shows a high affinity with the Tuscan dialect although with some internal differentiations, whereas Sardinian is a clearly distinct Romance language, preserving traces of the indigenous pre-Roman languages of the island until this very day.

These two islands also share some demographic features, likely a product of their abrupt geography, that have modeled their genetic structure mainly due to the effect of isolation and genetic drifting. Until the eighteenth century, Sardinia had a population that rarely exceeded 400,000 inhabitants. It was even lower in Corsica: 100,000 inhabitants until the end of the eighteenth century (Day 1987; Gatti 1995; Simi 1997). Genetic studies conducted in Corsica and Sardinia, although numerous, failed to coincide in data. Some studies based on classical markers indicate genetic similarity (Memmi et al. 1998; Vona et al. 2003), whereas others emphasise their genetic heterogeneity (Calafell et al. 1996). More recently, DNA studies have added more data without conclusive results. Francalacci et al. (2003) conducted a survey of Y-chromosome haplotypes in several samples from Corsica, Sicily and central Sardinia. Their main findings underline the differentiation of Sardinians and Sicilians from other Mediterraneans, whereas Corsica remains more similar to continental Italy and French samples, excluding the possibility of significant gene flow from central Sardinia to north-central Corsica. Data from mtDNA (Morelli et al. 2000) have also demonstrated a remarkable discontinuity among central Sardinians and both north Sardinians and Corsicans. On the other hand, a different mtDNA study (Falchi et al. 2006) found genetic similarities among Iberian, Corsican and Sardinian populations. This study confirms the fact that most mtDNA haplogroups in these samples coalesced in Paleolithic dates. Information from autosomal markers also gives controversial results; the maternal genetic similarities among Iberian, Corsican and Sardinian populations seem to be reflected in the high frequencies of ß039 thalassemic mutation (Falchi et al. 2005), whereas a multilocus analysis of autosomal microsatellites (Tofanelli et al. 2001) suggests a remarkable genetic differentiation between Sardinia and Corsica.

The particular position of Sicily in the centre of the Mediterranean has made the passage through it easier for peoples from virtually all of the Mediterranean and beyond. Before the Roman conquest, Sicily was occupied by remnants of the autochthonous populations of Sicani, Elymi, and Siculi (Indo-European populations that arrived between the second and first millennium BC), as well as by Phoenicians (tenth to eighth century BC) and Greeks (eighth century BC). The Sicilian language has inherited vocabulary and grammatical forms from these earliest settlers of the island as well as from the later colonists and conquerors. In view of their heterogeneous background, the subject of genetic relationships between populations on the island of Sicily is controversial. Some studies based on classical polymorphisms, and later on autosomal DNA markers (Calò et al. 2003; Ghiani et al. 2002; Piazza et al. 1988; Romano et al. 2003), indicated that Sicily is genetically heterogeneous, with a considerable East–West gradient compatible with population settlements occurring at different times. Other authors (Rickards et al. 1998) state that there was no clear geographic clustering within Sicily, rejecting an East–West differentiation.

Although the genetic information here summarised is extensive and covers everything from classical polymorphisms to uniparental and autosomal DNA, as far as we know, none of these studies have tested the four main islands with samples including different geographical areas inside each one jointly. This is the context in which we are presenting our work. We analysed a set of eight autosomal *Alu* polymorphisms and three short tandem

repeats (STRs) closely linked to the CD4, F13B and DM *Alu* markers in seven regions of Majorca, Sardinia, Corsica, and Sicily. We selected these particular markers for two main reasons, the first being the widely contrasted informative nature of *Alu* insertions for the study of human populations (Watkins et al. 2001) due to their stability, low mutation rate and known ancestral state, and the second due to the remarkable degree of information provided by *Alu* markers linked with STRs. The latter are very effective for estimating divergence between populations, although their mutation rate involves a certain degree of homoplasy that can mask the true genetic relationships. Information on haplotype frequencies, together with STR variation on ancestral *Alu* allele background compared with STR variation on the derived *Alu* alleles, has been used to estimate fine genetic relationships between human populations, not only on a large geographical scale (Tishkoff et al. 1996; Ramakrishnan and Mountain 2004), but also at a microgeographical level (Flores et al. 2000; Esteban et al. 2004).

The main objectives of this work are: (1) exploration of the degree of internal variability of Corsica, Sardinia and Sicily for comparison of the results with previous studies that suggested different heterogeneity levels inside these islands, (2) analysis of the genetic relationships among the four islands by maximum usage of different genetic markers such as *Alu* and STRs and (3) use of the qualitative information provided by the *Alu*/STR haplotypes to determine the amount of external gene flow received in the islands as a result of their historical background.

Material and methods

A total of 360 unrelated and healthy autochthonous individuals from seven well-defined rural areas of Majorca, central Sardinia, west-coast Sardinia, central Corsica, west-coast Corsica, east Sicily and west Sicily were analysed. Samples were obtained with the informed consent of the participants, whos four grandparents were born in the same region. The geographical position of the samples is detailed in Fig. 1.

Eight human-specific *Alu* insertion polymorphisms (DM, HS2.43, B65, PV92, D1, F13B, A25 and TPA25) were typed using the primers and polymerase chain reaction (PCR) amplification conditions previously described in Stoneking et al. (1997) and Edward and Gibbs (1992), with minor modifications. As for STRs, CD4 consists of a pentanucleotide $(TTTTC)_n$ repeat amplified according to Tishkoff et al. (1996), with minor modifications. This STR maps approximately 9 kb from the *Alu* marker. The F13B STR is a tetranucleotide repeat $(TTTA)_n$ at 4 kb of the *Alu* marker. Amplification conditions were as described in Nishimura and Murray (1992), with slight modifications. The DM $(CTG)_n$ repeat was amplified according to Brook et al. (1992). In this case the *Alu* polymorphism is located 5 kb telomeric to the repeat. After amplification with fluorescent-labeled primers, PCR products were pooled and electrophoresed on an ABI PRISM 3700 DNA sequencer (Applied Biosystems, Foster City, CA, USA). Genescan and Genemapper 3.0 programs (ABI PRISM, Applied Biosystems) were used to generate fragment sizes and

Fig. 1 Geographical position of the seven insular samples and other west Mediterranean groups used in comparisons

genotypes. Different selected individuals were sequenced for each STR to confirm size lengths and assign the correct repeat number for comparisons with data generated by other authors.

Allele frequencies were computed by direct counting, and Hardy–Weinberg equilibrium was tested by an exact test (Guo and Thomson 1992). Standard gene diversity indices by populations and locus were estimated according to Nei (1987). Locus frequency distributions were compared by an exact test for population differentiation. The program PHASE was used to generate, by means of a Bayesian statistical method, estimates of haplotype frequencies. The v.2.1 implements extensions of the original methods described in Stephens et al. (2001) and Stephens and Donnelly (2003). Linkage disequilibrium estimates between the STR and their respective *Alu* were quantified using the adaptation of Black and Krafsur (1985) algorithms contained in the computer program GENETIX 4.05 (Belkhir et al. 1996–2004). The apportionment of genetic variance was checked by analysis of molecular variance (AMOVA) through the ARLEQUIN computer package (Excoffier et al. 2005). Locus-by-locus fixation indices (F_{ST}, F_{SC} and F_{CT}) were averaged to obtain a global value. The statistical significance of these averages was checked by combining probabilities (Sokal and Rohlf 1997).

Apart from the seven samples included in this study, data from other European and North African samples, mainly from the Mediterranean basin, were collected from the literature. For our seven samples, we reached a database of 18 *Alu* polymorphisms by linking this work to a previous one conducted by our research team (Caló et al. 2005), but the available literature allowed us to create a database of 20 Mediterranean groups only for the following *Alu*: APO, B65, PV92, D1, F13B, A25, TPA25 and ACE. These samples (see Fig. 1 for geographical location) come from the works of Stoneking et al. (1997), Comas et al. (2000) and Garcia-Obregon et al. (2006, 2007). The consulted population data from CD4, F13B and DM STRs was obtained from the ALFRED database (Rajeevan et al. 2005).

Population relationships were approached by means of F_{ST}-related genetic distances analyses (Reynolds et al. 1983) using the PHYLIP 3.6 package (Felsenstein 1989) and depicted through multidimensional scaling from the distance matrix. Genetic distances $(\delta\mu)^2$ for STR data according to Goldstein et al. (1995) were calculated by the computer program Microsat 2 (written by E. Minch and available from: http://www.hpgl.stanford.edu). Population divergence times were estimated according to Goldstein et al. (1995), who proposed an equation to calculate divergence time among two samples by dividing the estimated value of the $(\delta\mu)^2$ distance by twice the product of the mutation rate (β) and the constant size variance (ω) of mutational jumps considering a generation time of 25 years. For divergence time calculations, we assumed that ω is constant with a value of 0.04 (1/25) and β value of 2.8 (10^{-4} (Chakraborty et al. 1997). The time obtained is expressed in years before present (YBP).

Results

Variability in west Mediterranean Islands

Alu *polymorphisms*

Allele frequencies for the eight *Alu*s are shown in Table 1. In general, all loci were in Hardy–Weinberg equilibrium after Bonferroni correction (excepting D1 in west Sicily) and showed significant gene diversity differences (Kruskal–Wallis test $p < 0.001$). In regards to F_{ST} values, only the PV92 loci showed moderate genetic differentiation (F_{ST} of 6.7%). Although the samples from Sicily and Sardinia occupied extreme positions in the population variation ranges (see Table 1) for some loci, when average heterozygosities were compared, the Kruskal–Wallis test indicated no remarkable differences ($p = 0.873$) among our samples.

Pairwise population comparisons across the eight loci revealed a remarkable degree of heterogeneity (significant p values for 17 comparisons out of 21) disrupted only for four population comparisons that failed to show significant differences: west-coast Corsica with both central Sardinia and Majorca, and central Corsica with east Sicily and Majorca. The locus that showed the highest number of significant population comparisons (11 out of 21) was consistent with the genetic differentiation revealed by F_{ST} values, PV92. A nonhierarchical AMOVA yielded an average F_{ST} in west-Mediterranean islands of 2.2% ($p < 0.001$). Inside Sardinia, the level of population genetic variance (F_{ST} of 5.5%, $p < 0.001$) was even higher than that observed for the whole of the islands. Both the low number of samples and the extreme allele frequencies shown by the two Sardinian samples in almost all loci probably accounted for this F_{ST} value. However, the same pattern was observed when we recalculated F_{ST} using information from 18 *Alu* polymorphisms: average F_{ST} (3.7%, $p < 0.001$) inside Sardinia was triple that observed among islands (1.3%, $p < 0.01$).

STR gene diversity and Alu*/STR compound systems*

Allele size frequencies of CD4, F13B and DM microsatellites are available as supplementary material from the Web site of the journal. Overall, the three distributions were in Hardy–Weinberg equilibrium after Bonferroni correction, with the only exception of DM STR being in central

Table 1 *Alu* insertion frequencies in west-Mediterranean islands

Alu	Sardinia centre	Sardinia west coast	Corsica centre	Corsica west coast	Sicily east	Sicily west	Majorca	F_{ST} per locus	Variation ranges reviewed populations
DM (2N)	100	74	92	90	80	92	110		Loci not included in the reviewed literature
Alu+	0.607	0.300	0.695	0.567	0.787	0.576	0.527	0.030	
H	0.485	0.442	0.428	0.497	0.339	0.494	0.503	$p = 0.043$	
HS2.43 (2N)	100	98	102	76	102	100	114		0.02 SE Morocco – 0.28 C Sardinia
Alu+	0.280	0.061	0.137	0.158	0.069	0.130	0.096	0.030	
H	0.407	0.116	0.239	0.269	0.129	0.228	0.176	$p = 0.020$	
B65 (2N)	100	46	102	94	102	96	106		0.47 Navarre (Spain) – 0.73 Algerians
Alu+	0.520	0.630	0.598	0.489	0.569	0.583	0.575	0	
H	0.504	0.476	0.485	0.505	0.485	0.491	0.493	NS	
PV92 (2N)	98	90	102	94	98	90	112		0.04 W Sicily – 0.40 S Morocco
Alu+	0.122	0.389	0.245	0.138	0.112	0.044	0.087	0.067	
H	0.217	0.481	0.374	0.241	0.201	0.086	0.307	$p = 0.009$	
D1 (2N)	100	46	102	94	102	96	114		0.11 WC Sardinia – 0.53 W Sicily
Alu+	0.310	0.109	0.323	0.415	0.422	0.531	0.342	0.029	
H	0.432	0.198	0.442	0.491	0.492	0.503	0.454	$p = 0.031$	
F13B (2N)	100	84	102	96	102	94	114		0.29 W Morocco – 0.62 Greek Cypriots
Alu+	0.370	0.333	0.451	0.583	0.471	0.372	0.500	0.009	
H	0.471	0.450	0.500	0.491	0.503	0.472	0.504	NS	
A25 (2N)	100	92	102	94	102	94	114		0.06 C Sardinia – 0.23 SE Morocco
Alu+	0.060	0.109	0.137	0.096	0.108	0.202	0.114	0	
H	0.114	0.196	0.239	0.175	0.194	0.326	0.204	NS	
TPA25 (2N)	100	66	96	94	86	92	114		0.33 WC Sardinia – 0.64 W Sicily
Alu+	0.500	0.333	0.542	0.596	0.465	0.641	0.570	0.012	
H	0.509	0.451	0.502	0.487	0.504	0.465	0.494	$p = 0.052$	
Average H	0.392	0.351	0.401	0.394	0.356	0.383	0.392		Among Islands average F_{ST} 0.022
SD	0.147	0.153	0.109	0.140	0.161	0.155	0.141		$p < 0.001$
CD4									Data from Calo et al. (2005)
Alu+	0.694	0.786	0.716	0.837	0.598	0.700	0.693		

Previously described CD4 *Alu* frequencies on these samples have been included for posterior use in CD4 *Alu*/STR haplotype calculations (Table 3). Variation ranges according to data from the reviewed literature for 20 European and North African populations

2N number of chromosomes analysed, *H* Nei's nonbiased gene diversity *NS* not significant

Sardinia. Allele diversity values and some statistical parameters of allele size distributions, including STR variation on derived chromosomes, are reported in Table 2. Heterozygosity values in the three STRs showed similarly notable levels of within-population variation, but no significant population differences were detected in neither diversity values or allele size frequencies. STR variation in the CD4- and DM-derived chromosomes (those *Alu*-) was extremely lower in all cases, according to previous knowledge about the distribution of these compound systems in modern humans (Tishkoff et al. 1996, 1998). On the contrary, STR variation in the derived F13B chromosomes (those carrying the *Alu* insertion) was high and very similar to that described for the general variation of this STR.

Alu/STR haplotype frequencies are reported in Tables 3, 4 and 5 for CD4, F13B and DM markers, respectively. In the three compound systems, *Alu* and STR alleles were in linkage disequilibrium. Agreeing with that observed for the STR distributions, our samples showed high levels of within-population diversity but weak population differences.

The number of different CD4 haplotypes in Sardinia and Sicily (seven and nine, respectively) exceeded in number those found in Corsica and Majorca due to the presence of some African characteristic combinations (*Alu*—with both alleles of five and eight repeats) in the former populations. In the particular case of east Sicily, these haplotypes accounted for a frequency of 5.6%. Another haplotype (*Alu*-/10 repeats allele) characteristic of Berber groups (Flores et al. 2000;

Table 2 Variation of CD4, F13 and DM microsatellites in west-Mediterranean islands

	Global STR variation						STR variation on derived *Alu* chromosomes					
	2N	No. alleles	Mean	Variance	Allele range	H	2N	No. alleles	Mean	Variance	Allele range	H
	CD4 pentanucleotide											
Sardinia C	90	4	6.75	4.70	5–10 (5)	0.677	27	3	6.11	0.64	5–10 (6)	0.140
Sardinia WC	94	6	7.61	6.14	5–12 (5)	0.724	26	1	6.00	–	6	0.000
Corsica C	94	5	6.91	5.00	5–11 (5)	0.718	31	2	5.93	0.06	5–6 (6)	0.121
Corsica WC	96	6	7.12	5.77	5–12 (5)	0.683	19	1	6.00	–	6	0.000
Sicily east	94	6	6.84	4.61	5–12 (6)	0.704	36	3	6.14	0.98	5–10 (6)	0.248
Sicily west	96	6	7.57	5.73	5–12 (10)	0.724	28	3	6.53	2.11	5–10 (6)	0.304
Majorca	100	4	6.83	5.06	5–11 (5)	0.684	26	1	6.00	–	6	0.000
	F13B tetranucleotide											
Sardinia C	72	4	8.74	1.63	6–10 (10)	0.720	10	2	8.57	0.70	9–10 (10)	0.520
Sardinia WC	76	4	8.84	1.36	6–10 (9)	0.713	27	3	8.88	0.64	8–10 (8/9)	0.658
Corsica C	94	5	8.80	1.54	6–10 (10)	0.708	41	4	8.32	0.57	7–10 (8)	0.468
Corsica WC	96	4	8.64	1.39	6–10 (8)	0.680	56	3	8.46	0.54	8–10 (8)	0.487
Sicily east	88	5	8.58	1.69	6–10 (8)	0.729	41	4	8.44	0.80	7–10 (8)	0.570
Sicily west	92	5	8.73	1.54	6–10 (10)	0.685	35	4	8.74	0.96	7–10 (8)	0.580
Majorca	94	5	8.74	1.68	6–11 (10)	0.734	44	3	8.54	0.53	8–10 (8)	0.558
	DM trinucleotide											
Sardinia C	94	13	9.75	26.95	5–31 (5)	0.760	32	4	9.68	12.16	5–13 (12)	0.101
Sardinia WC	94	15	10.91	33.15	5–29 (5)	0.799	67	6	12.13	6.33	5–15 (12)	0.015
Corsica C	96	12	9.43	35.13	5–33 (5)	0.704	28	6	12.57	2.18	8–15 (13)	0.014
Corsica WC	96	13	10.02	33.37	5–29 (5)	0.732	41	6	11.51	10.81	5–20 (13)	0.083
Sicily east	94	14	10.82	38.77	5–26 (5)	0.784	19	3	12.05	0.61	11–13 (12)	0.157
Sicily west	96	13	10.02	30.41	5–27 (5)	0.762	35	7	12.43	6.37	5–21 (13)	0.036
Majorca	114	14	11.05	34.27	5–30 (5)	0.796	48	6	10.89	12.61	5–21 (13)	0.063

Allele range in number of repeats. In parentheses: modal allele

2N number of chromosomes, *No. alleles* number of different alleles, variance in repeat number, *H* Nei's nonbiased gene diversity

Esteban et al. 2004) was also found in central Sardinia (1.2%), east Sicily (2.5%) and, with more remarkable frequencies, in west Sicily (4%). The pattern of F13B haplotype frequencies was very similar among islands excluding the comparison between west-coast Sardinia and west-coast Corsica ($p = 0.04$). Some haplotypes were common to all groups, whereas some others were found scattered in certain samples; however, none of these particular haplotypes were detected in Majorca. DM haplotypes also showed a similar pattern of high within-population diversity combined with great population homogeneity.

Leaving aside some occasional differences, the remarkable genetic heterogeneity within and among islands detected for the set of *Alu* markers did not match up with the global homogeneity detected for STR variation and *Alu*/STR haplotype frequencies. Furthermore, in west-Mediterranean islands, the F_{ST} values deduced from STR variation in the three loci as a whole ($F_{ST} = 0.01\%$) or from the three *Alu*/STR compound systems ($F_{ST} = 0.02\%$) were not significantly different from zero.

Genetic relationships in the west-Mediterranean basin

Global relationships in our samples were assessed through F_{ST}-related genetic distance matrices for 18 *Alu* polymorphisms and three *Alu*/STR combinations and through $(\delta\mu)^2$ distances for the three STRs. In all matrices, distance values were significantly different from zero in more than 90% of cases. *Alu* and *Alu*/STR distance matrices were positively correlated (Mantel test, $r = 0.763$, $p = 0.041$) and underlined the genetic differentiation of west-coast Sardinia and west Sicily [see Fig. 2a for the multidimensional scaling (MDS) plot based on *Alu* data]. The first dimension of the MDS plot based on $(\delta\mu)^2$ distances (Fig. 2b) clearly distinguished two population clusters, with west-coast Sardinia and west Sicily as the most differentiated samples within each group. Figure 2b also contains estimates of divergence times among samples; the two main population clusters showed a time separation of around 25,000 YBP, whereas the divergence inside each group was considerably lower.

Table 3 CD4 *Alu*/short tandem repeat (STR) haplotype frequencies and global gene diversity

	Haplotypes with the *Alu* insertion						
	5 (85)	6 (90)	8 (100)	9 (105)	10 (110)	11 (115)	12 (120)
Sardinia C	0.3736 ± 0.0042	0.0593 ± 0.0059			0.2375 ± 0.0038		0.0227 ± 0.0002
Sardinia WC	0.3022 ± 0.0023	0.0286 ± 0.0053		0.0132 ± 0.0001	0.3405 ± 0.0043	0.0524 ± 0.0017	0.0394 ± 0.0001
Corsica C	0.3466 ± 0.0008	0.0123 ± 0.0045			0.2518 ± 0.0052	0.0488 ± 0.0062	
Corsica WC	0.4355 ± 0.0047			0.0104 ± 0.0001	0.3008 ± 0.0036	0.0417 ± 0.0001	0.0104 ± 0.0001
Sicily east	0.2946 ± 0.0059	0.0500 ± 0.0094			0.1875 ± 0.0074	0.0423 ± 0.0017	0.0106 ± 0.0001
Sicily west	0.2761 ± 0.0021	0.0122 ± 0.0041	0.0292 ± 0.0054		0.3209 ± 0.0049	0.0319 ± 0.0006	0.0318 ± 0.0012
Majorca	0.4162 ± 0.0087	0.0164 ± 0.0044			0.2495 ± 0.0080	0.0452 ± 0.0019	

	Haplotypes without the *Alu* insertion				
	5 (85)	6 (90)	8 (100)	10 (110)	Global GD
Sardinia C	0.0128 ± 0.0042	0.2816 ± 0.0059		0.0125 ± 0.0038	0.7751 ± 0.0254
Sardinia WC		0.2214 ± 0.0053			0.8154 ± 0.0227
Corsica C	0.0256 ± 0.0076	0.3068 ± 0.0045			0.7881 ± 0.0238
Corsica WC		0.1971 ± 0.0030			0.7387 ± 0.0338
Sicily east	0.0352 ± 0.0059	0.3330 ± 0.0094	0.0211 ± 0.0012	0.0253 ± 0.0074	0.8193 ± 0.0196
Sicily west	0.0111 ± 0.0021	0.2431 ± 0.0041		0.0408 ± 0.0048	0.8211 ± 0.0194
Majorca		0.2564 ± 0.0044			0.7501 ± 0.0369

STR alleles are expressed in number of repeats; parentheses show the size in base pairs. Haplotypes with frequencies lower than 1% are excluded

Table 4 F13B *Alu*/short tandem repeat (STR) haplotype frequencies and global gene diversities

	Haplotypes with the *Alu* insertion			
	7 (176)	8 (180)	9 (184)	10 (188)
Sardinia C		0.2503 ± 0.0143	0.0529 ± 0.0105	0.0846 ± 0.0129
Sardinia WC		0.1329 ± 0.0083	0.1296 ± 0.0114	0.0985 ± 0.0128
Corsica C	0.0205 ± 0.0033	0.3056 ± 0.0063	0.0585 ± 0.0095	0.0592 ± 0.0086
Corsica WC		0.3925 ± 0.0125	0.1094 ± 0.0092	0.0813 ± 0.0087
Sicily east	0.0341 ± 0.0001	0.2800 ± 0.0148	0.0623 ± 0.0120	0.0895 ± 0.0114
Sicily west	0.0109 ± 0.0001	0.2028 ± 0.0185	0.0353 ± 0.0101	0.1310 ± 0.0170
Majorca		0.2729 ± 0.0061	0.1266 ± 0.0151	0.0603 ± 0.0144

	Haplotypes without the *Alu* insertion				
	6 (172)	8 (180)	9 (184)	10 (188)	Global GD
Sardinia C	0.1101 ± 0.0036	0.0552 ± 0.0143	0.1554 ± 0.0105	0.2904 ± 0.0129	0.8279 ± 0.0237
Sardinia WC	0.0832 ± 0.0012	0.0477 ± 0.0083	0.2592 ± 0.0114	0.2487 ± 0.0128	0.8294 ± 0.0204
Corsica C	0.0821 ± 0.0050		0.1224 ± 0.0095	0.3450 ± 0.0086	0.7763 ± 0.0273
Corsica WC	0.0832 ± 0.0011	0.0450 ± 0.0125	0.0468 ± 0.0092	0.2416 ± 0.0087	0.7726 ± 0.0288
Sicily east	0.1136 ± 0.0002	0.0722 ± 0.0148	0.0967 ± 0.0117	0.2514 ± 0.0114	0.8239 ± 0.0217
Sicily west	0.0865 ± 0.0021	0.1776 ± 0.0185	0.0951 ± 0.0101	0.2603 ± 0.0170	0.8239 ± 0.0193
Majorca	0.1157 ± 0.0037		0.1180 ± 0.0151	0.2907 ± 0.0144	0.8236 ±0.0206

STR alleles are expressed in number of repeats; parentheses show the size in base pairs

Heterogeneity within west-Mediterranean islands has been examined in a wider context (Fig. 3a) to determine its true significance. F_{ST}-related genetic distances among our samples and a set of related populations for 8 *Alu* polymorphisms ranged from the lowest value of 0.0018 between two Spanish samples (northeast Spain and Navarre) to the highest 0.1379 (between west-coast Sardinia and west Sicily). Average genetic distances inside west-

Table 5 DM *Alu*/short tandem repeat (STR) haplotype frequencies and global gene diversities

	Haplotypes with the *Alu* insertion							
	5 (77)	10 (92)	11 (95)	12 (98)	13 (101)	14 (104)	15 (107)	20 (122)
Sardinia C	0.3440 ± 0.0228	0.0357 ± 0.0001				0.1071 ± 0.0001	0.0276 ± 0.0149	0.0357 ± 0.0001
Sardinia WC	0.2352 ± 0.0252							
Corsica C	0.5292 ± 0.0064	0.0217 ± 0.0001		0.0110 ± 0.0012		0.0679 ± 0.0068		
Corsica WC	0.3843 ± 0.0081	0.0222 ± 0.0001	0.0111 ± 0.0001	0.0135 ± 0.0056	0.0249 ± 0.0053	0.0333 ± 0.0001		
Sicily east	0.4280 ± 0.0108		0.0272 ± 0.0049	0.0466 ± 0.0103	0.0584 ± 0.0091	0.0504 ± 0.0032	0.0128 ± 0.0001	0.0183 ± 0.0110
Sicily west	0.4123 ± 0.0113	0.0236 ± 0.0041	0.0236 ± 0.0043			0.0262 ± 0.0114	0.0152 ± 0.0053	0.0109 ± 0.0001
Majorca	0.2815 ± 0.0179		0.0252 ± 0.0121	0.0492 ± 0.0039	0.0483 ± 0.0152	0.0290 ± 0.0070		0.0161 ± 0.0043

	21 (125)	22 (128)	24 (134)	25 (137)	26 (140)	28 (146)	Global GD
Sardinia C	0.0357 ± 0.0001						0.8466 ± 0.0534
Sardinia WC		0.0500 ± 0.0001					0.8789 ± 0.0432
Corsica C	0.0326 ± 0.0002					0.0109 ± 0.0001	0.7504 ± 0.0418
Corsica WC	0.0111 ± 0.0001	0.0111 ± 0.0001				0.0111 ± 0.0001	0.8005 ± 0.0333
Sicily east		0.0893 ± 0.0023		0.0253 ± 0.0021	0.0128 ± 0.0001		0.7862 ± 0.0428
Sicily west	0.0217 ± 0.0001			0.0109 ± 0.0001			0.8051 ± 0.0350
Majorca	0.0220 ± 0.0148	0.0103 ± 0.0111	0.0161 ± 0.0058				0.8719 ± 0.0194

	Haplotypes without the *Alu* insertion							
	5 (77)	11 (95)	12 (98)	13 (101)	14 (104)	15 (107)	20 (122)	21 (125)
Sardinia C	0.1203 ± 0.0228	0.0357 ± 0.0001	0.1329 ± 0.0190	0.0682 ± 0.0103				
Sardinia WC	0.0648 ± 0.0252	0.0500 ± 0.0001	0.2409 ± 0.0193	0.2000 ± 0.0001	0.0500 ± 0.0001	0.0943 ± 0.0159		
Corsica C		0.0652 ± 0.0008	0.0433 ± 0.0012	0.0972 ± 0.0024	0.0734 ± 0.0068	0.0109 ± 0.0001		
Corsica WC	0.0713 ± 0.0081	0.0444 ± 0.0001	0.0754 ± 0.0056	0.1974 ± 0.0053	0.0333 ± 0.0001		0.0111 ± 0.0001	
Sicily east		0.0497 ± 0.0049	0.0816 ± 0.0103	0.0698 ± 0.0091				
Sicily west	0.0116 ± 0.0113	0.1068 ± 0.0043	0.0958 ± 0.0047	0.1181 ± 0.0040	0.0173 ± 0.0114		0.0109 ± 0.0001	0.0109 ± 0.0001
Majorca	0.0958 ± 0.0179	0.0502 ± 0.0121	0.1017 ± 0.0039	0.1498 ± 0.0152	0.0181 ± 0.0070			0.0157 ± 0.0148

STR alleles are expressed in number of repeats; parentheses show the size in base pairs. Haplotypes with frequencies lower than 1% are excluded

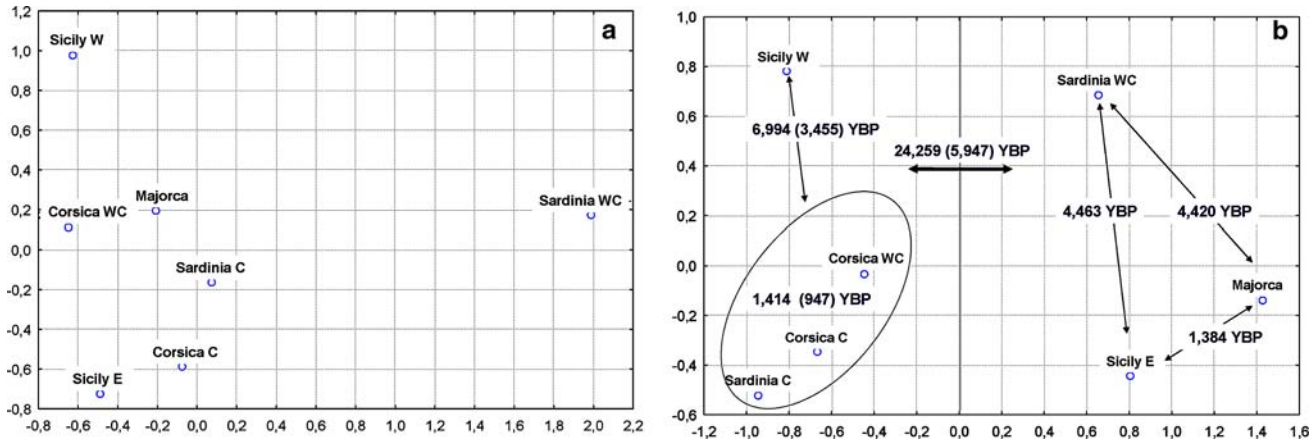

Fig. 2 **a** Plot of multidimensional scaling (MDS) (stress = 0.008) applied to the F_{ST} genetic distance matrix based on 18 *Alu* markers. **b** Plot of MDS (stress < 0.001) applied to the $(\delta\mu)^2$ genetic distance matrix based on three short tandem repeats (STRs). Years before present (YBP) estimated through the distance values are indicated for the main groups; for averaged YBP, standard deviations (SDs) are indicated in *parentheses*

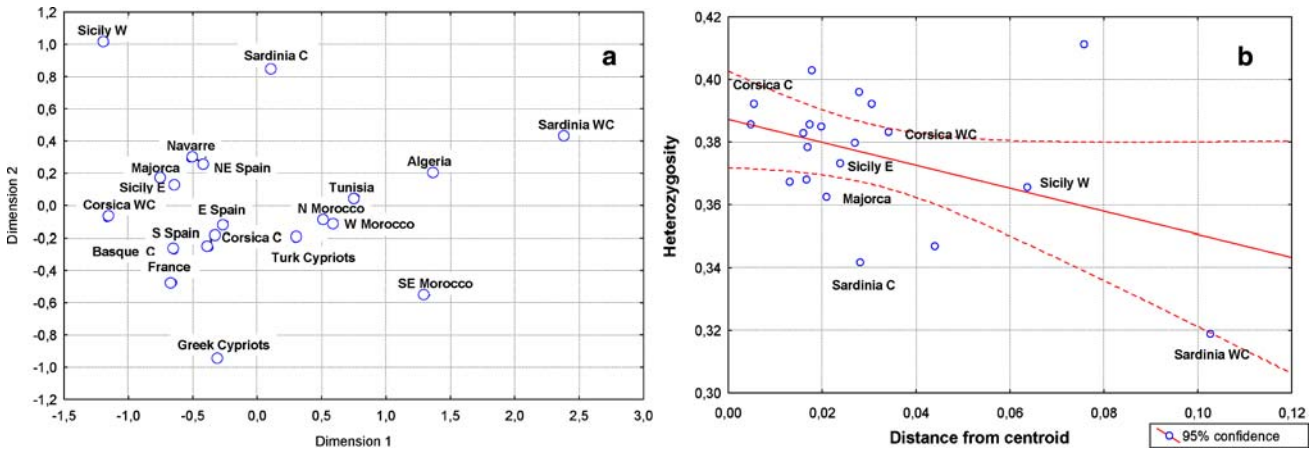

Fig. 3 **a** Plot of multidimensional scaling (MDS) (stress = 0.091) applied to the F_{ST} genetic distance matrix based on eight *Alu* markers. **b** Position of west-Mediterranean islands in the heterozygosity vs. distance from the centroid plot based on *Alu* polymorphisms

Mediterranean islands (average distance dm = 0.044) were considerably higher than the average distances among southwestern Europeans (dm = 0.010) and North Africans (dm = 0.013). On average, west-Mediterranean islands showed the highest between-group distance with North Africans (0.045). The fraction of genetic variance resulting from differences among these two groups measured through the across-loci average F_{CT} value was 1.79% ($p < 0.001$).

When population relationships were depicted through an MDS plot (Fig. 3a), west-coast Sardinia, central Sardinia and west Sicily occupied a peripheral position in the upper part of the graphic, whereas the other samples were closely related to the Spanish and French samples. The Sardinian differentiation may be explained by the fact that, in comparison with the whole population correlation between distance from centroid and heterozygosity (Fig. 3b), they showed less heterozygosity than that expected under the Harpending and Ward (1982) model, suggesting either a greater influence of genetic isolation or smaller effective population size.

Discussion

For west-Mediterranean islands, the autosomal *Alu* and STR data reported here are the first to be described and jointly discussed in order to shed light on some of the most controversial issues of west-Mediterranean population relationships, namely, the internal degree of heterogeneity within islands, the particular affinities and/or differences among islands, the amount of external gene flow received and finally, the divergence times among these regions.

Concerning *Alu* markers, the seven west-Mediterranean samples show noticeable levels of genetic diversity, with

the only exceptions being east Sicily and west-coast Sardinia, wich have the lowest average heterozygosities. Genetic differentiation inside and among islands is extremely high as can be deduced from both the results of pairwise population comparisons (17 out of 21 cross-loci population comparisons are statistically significant) and global F_{ST} values. A general trend to low gene diversity in Sardinia (Fig. 3b) joined with discrepant patterns of *Alu* allele frequencies among samples could be consistent with such differentiation. The global degree of differentiation among islands (2.2%, $p < 0.001$) is even slightly higher than that reported in Europeans (1.9%) or North Africans (1.5%, Comas et al. 2000; 2.3%, Gonzalez-Perez et al. 2003) for a comparable set of *Alu* markers and samples.

West-coast Sardinia and west Sicily are clearly differentiated from all samples (Figs. 2a, 2b, 3a). The action of genetic drift in relatively small population groups could have contributed to their differentiation. Although we cannot ignore that historical, linguistic and some genetic evidence (Piazza 1988) in Sardinia point to differences in population settlements among central and coastal areas due to the confinement of the original population, the Nuragici, into the internal regions as a result of the Carthaginian and Roman invasions. In the case of Sicily, historical records also indicate an important retreat of the original Sicanian population due to the arrival of the continental Italian Sicels. Our results concur with evidence based on Y-chromosome haplotypes (Francalacci et al. 2003) and mtDNA (Morelli et al. 2000) that point out Sardinia and Sicily as the most differentiated populations in the west-Mediterranean basin.

The relative heterogeneity among the remaining insular samples revealed by the plot based on 18 *Alu* markers (Fig. 2a) is less evident when other Mediterranean groups are added to the MDS analysis. The proximity of Corsica, Majorca and east Sicily to continental samples indicated by the set of eight *Alu* markers (Fig. 3a) has also been suggested by data from mtDNA (Falchi et al. 2006; Picornell et al. 2005).

Genetic differentiation in west-Mediterranean islands is not evident by STR variation. Although all samples show notable levels of within-population diversity, neither significant population differences nor remarkable levels of genetic variance have been detected in any of the three analysed STRs. The discrepancies observed between results indicated by *Alu* markers and STR variation may derive from the different nature of these two polymorphisms. The former are unique events far from the effect of random fluctuations caused by mutation and probably reflect the ancestral origin of populations. A population split from this ancestral group with enough time to accumulate STR variation due to the high microsatellite mutation rates, together with the homogenising effect of gene flow, could explain the observed discrepancies in genetic heterogeneity and F_{ST} values among these two genetic markers.

Gene flow among west-Mediterranean islands and beyond seems to have been outstanding. STR variation on the three loci coincide in showing high heterozygosity values in all samples; in most cases, STR variation parameters are higher than those reported for mainland Europeans (Tishkoff et al. 1996, 1998; Esteban et al. 2004). Insularity has not acted as a strong barrier to gene flow, at least among west-Mediterranean islands and mainland southern Europe, according to the merged historical background of these samples. However, historical records also point out North African influences. We have not detected any remarkable affinity among west-Mediterranean islands and North Africans. But this fact does not exclude some particular examples of African gene flow. Traces of African contributions to the gene pool of some islands can be deduced from the frequency of several CD4 *Alu*/STR haplotypes. The relatively high contribution of African-characteristic haplotypes in Sicily (8.16% in the east sample and 5.16% in the west sample) in comparison with the other islands (less than 2.5%) agrees with the strategic geographic position and the historical background of this island. Majorca, however, which was under Islamic rule for more than three centuries, does not exhibit any trace of African haplotypes. This fact agrees with other genetic data (Picornell et al. 2005) reinforcing the historical evidence that documented an important repopulation of the island by Spaniards after the Catalano–Aragonese conquest.

We conclude with some data of divergence times among samples, even though these estimations represent maximum values, because they are based on the assumption that the measured STR variation has developed locally, and we know that gene-flow processes in the west Mediterranean could have added some bias to time calculations. Time estimates (Fig. 2b) separate our samples by a time range of around 24,259 ± 6,211 YBP in two groups: central Sardinia, central Corsica, west-coast Corsica and west Sicily vs. west-coast Sardinia, east Sicily and Majorca. An average date of 5,973 ± 2,815 YBP separates west Sicily and west-coast Corsica from the remaining populations inside their respective groups. These dates are compatible with the population heterogeneity revealed by *Alu* data, suggesting that some differences among our samples could be traced back to the first settlement of the islands, likely reflecting genetic drift and/or genetic isolation processes. On the other hand, the high within-population diversities and the remarkable STR and *Alu*/STR homogeneity among islands suggest that, at least since Neolithic times, gene flow has been active in the west-Mediterranean basin. Genetic drift in west-coast Sardinia and gene flow in west Sicily have probably stressed their general genetic differentiation.

Acknowledgments We thank all of the anonymous islanders for their participation in the study, and all those who contributed to the sampling, for their valuable collaboration. This work was supported by grant CGL2005–03391 from the Spanish Ministry and grant 2005SGR00252 from the Generalitat de Catalunya. The work of EGP was financed by grant 2001FI00177 from the Generalitat de Catalunya. We give special thanks to A.D. Hadley Loera for careful revision of the manuscript.

References

Belkhir K, Borsa P, Chikhi L, Rafauste N, Bonhomme F (1996–2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France)

Black WC, Krafsur E (1985) A FORTRAN program for the calculation and analysis of two-locus linkage disequilibrium coefficients. Theor Appl Genet 70:491–496

Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K et al (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein family member. Cell 68:799–808

Calafell F, Bertranpetit J, Rendine S, Cappello N, Mercier P, Amoros JP, Piazza A (1996) Population history of Corsica: a linguistic and genetic analysis. Ann Hum Biol 23:237–251

Calo CM, Garofano L, Mameli A, Pizzamiglio M, Vona G (2003) Genetic analysis of a Sicilian population using 15 short tandem repeats. Hum Biol 75:163–178

Calo CM, Piras IS, Moral P, Falchi A, Ghiani ME, Varesi L, Vona G (2005) Analisi Molecolare delle popolazioni del Mediterraneo traverso 11 inserzioni Alu. Antropo 9:1–12. http://www.didac.ehu/antropo

Cappello N, Rendine S, Griffo R, Mameli GE, Succa V, Vona G, Piazza A (1996) Genetic analysis of Sardinia: I. Data on 12 polymorphisms in 21 linguistic domains. Ann Hum Genet 60:125–141

Chakraborty R, Kimmel M, Stivers DN, Davidson LJ, Deka R (1997) Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proc Natl Acad Sci USA 95:1041–1046

Comas D, Calafell F, Berchemsi N, Helal A, Lefranc G, Stoneking M, Batzer MA, Bertranpetit J, Sajantila A (2000) Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary through the Gibraltar Strait. Hum Genet 107:312–319. doi:10.1007/s004390000370

Day J (1987). Gli uomini e il territorio: i grandi orientimenti del popolamento sardo dall'XI al XVIII secolo. Storia dei sardi e della Sardegna. Vol. II Milano Jaka Book, pp 13–47

Edwards MC, Gibbs RA (1992) A human dimorphism resulting from loss of an Alu. Genomics 14:590–597

Esteban E, González-Pérez E, Harich N, Lopez-Alomar A, Via M, Luna F, Moral P (2004) Genetic relationships among Berbers and South Spaniards based on CD4 microsatellite/Alu haplotypes. Ann Hum Biol 31:202–212. doi:10.1080/03014460310001652275

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

Falchi A, Giovannoni L, Calo CM, Piral IS, Moral P, Paoli G, Vona G, Varesi L (2006) Genetic history of some western Mediterranean human isolates through mtDNA HVR1 polymorphisms. J Hum Genet 51:9–14. doi:10.1007/s10038-005-0324-y

Falchi A, Giovannoni L, Vacca L, Latini V, Vona G, Varesi L (2005) Beta-globin gene cluster haplotypes associate with beta-thalassemia on Corsica island. Am J Hematol 78:27–32. doi:10.1002/ajh.20199

Felsenstein J (1989) PHYLIP-phylogeny inference package. Cladistics 5:164–166

Flores C, Maca-Meyer N, Gonzalez AM, Cabrera VM (2000) Northwest African distribution of the *CD4/Alu* microsatellite haplotypes. Ann Hum Genet 64:321–327

Francalacci P, Morelli L, Underhill PA, Lillie AS, Passarino G, Useli A, Madeddu R, Paoli G, Tofanelli S, Calo CM, Ghiani ME, Varesi L, Memmi M, Vona G, Lin AA, Oefner P, Cavalli-Sforza LL (2003) Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. Am J Phys Anthrop 121:270–279. doi:10.1002/ajpa.10265

Garcia-Obregon S, Alfonso-Sanchez MA, Perez-Miranda AM, Vidales C, Arroyo D, Peña JA (2006) Genetic position of Valencia (Spain) in the Mediterranean basin according to *Alu* insertions. Am J Hum Biol 18:187–195. doi:10.1002/ajhb.20487

Garcia-Obregon S, Alfonso-Sanchez MA, Perez-Miranda AM, De Pancorbo MM, Peña JA (2007) Polymorphic *Alu* insertions and the genetic structure of Iberian Basques. J Hum Genet 52:317–327. doi:10.1007/s10038-007-0114-9

Gatti AM (1995) Tra demografia e genetica: populazione e matrimonio in Sardegna. Antropologia Contemporanea 18:67–80

Ghiani ME, Calo CM, Autuori L, Mameli GE, Succa V, Vacca L, Cerutti N, Rabino Massa E, Vona G (2002) New data on the genetic structure of the population of Sicily: analysis of the Alia population (Palermo, Italy). Am J Hum Biol 14:289–299. doi:10.1002/ajhb.10036

Goldstein DB, Ruiz-Linares A, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. Genetics 139:463–471

González-Pérez E, Via M, Esteban E, Lopez-Alomar A, Mazieres S, Harich N, Kandil M, Dugoujon JM, Moral P (2003) Alu insertions in the Iberian Peninsula and NW Africa: genetic boundaries or melting pot? Coll Antropol 27:491–500

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

Harpending H, Ward RH (1982) Chemical systematics and human population. In: Nitecki M (ed) Biochemical aspects of evolutionary biology. University of Chicago Press, Chicago, pp 213–252

Memmi M, Moral P, Calo CM, Autuori L, Mameli GE, Succa V, Varesi L, Vona G (1998) Genetic structure of southwestern Corsica (France). Am J Hum Biol 10:567–577

Morelli L, Grosso MG, Vona G, Varesi L, Torroni A, Francalacci P (2000) Frequency distribution of mitochodrial haplogroups in Corsica and Sardinia. Hum Biol 72:585–595

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nishimura DY, Murray JC (1992) A tetranucleotide repeat for the F13B locus. Nucleic Acids Res 20:1167

Piazza A, Cappello N, Olivetti E, Rendine S (1988) A genetic history of Italy. Ann Hum Genet 52:203–213

Picornell A, Gomez-Barbeito L, Tomas C, Castro JA, Ramon MM (2005) Mitochondrial DNA HVRI variation in Balearic populations. Am J Phys Anthrop 128:119–130. doi:10.1002/ajpa.10423

Rajeevan H, Cheung KH, Gadagkar R, Stein S, Soundararajan U, Kidd JR, Pakstis AJ, Miller PL, Kidd KK (2005) ALFRED: an allele frequency database for microevolutionary studies. Evolut Bioinf 1:1–10

Ramakrishnan U, Mountain J (2004) Precision and accuracy of divergency time estimates from STR and SNPSTR variation. Mol Biol Evol 21:1960–1971. doi:10.1093/molbev/msh212

Reynolds J, Weir BS, Cockerman CC (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105:767–779

Rickards O, Martinez-Labarga C, Scano G, De Stefano GF, Biondi G, Pacaci M, Walter H (1998) Genetic history of the population of Sicily. Hum Biol 70:699–714

Romano V, Cali F, Ragalmuto A, D'Anna RP, Flugy A, De Leo G, Giambalvo O, Lisa A, Fiorani O, Di Gaetano C, Salerno A, Tamouza R, Charron D, Zei G (2003) Autosomal microsatellite and mtDNA genetic analysis in sicily (Italy) Ann Hum Genet 67:42–53. doi:10.1046/j.1469-1809.2003.00007.x

Simi P (1997) L'apport de la documentation de la Franciscorsa à la connaissance de la démographie de la Corse. Bull Soc Hist Nat Corse 679:25–43

Sokal RR, Rolf FJ (1997) Biometry. The principles and practice of statistics in biological research, 3rd edn. WH Freeman, New York

Stephens M, Donnelly P (2003) A comparison of bayesian methods for haplotype reconstruction. Am J Hum Genet 73:1162–1169. doi:0002-9297/2003/7305-0018$15.00

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989. doi:0002-9297/2001/6804-0020$02.00

Stoneking M, Fontius JJ, Clifford S, Soodyall H, Arcot SS, Saha N, Jenkins T, Tahir MA, Deininger PL, Batzer MA (1997) Alu insertion polymorphisms and human evolution: evidence for a larger population size in Africa. Genome Res 7:1061–1071.doi: 10.1101/gr.7.11.1061

Tishkoff SA, Dietzsch E, Speed W, Pakstis AJ, Kidd JR, Cheung K, Bonne-Tamir B, Santachiara-Benerecetti AS, Moral P, Krings M, Pääbo S, Watson E, Risch N, Jenkins T, Kidd KK (1996) Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. Science 271:1380–1387. doi: 10.1126/science.271.5254.1380

Tishkoff SA, Goldmans A, Calafell F, Speed WC, Deinard AS, Bonne-Tamir B, Kidd JR, Pakstis AJ, Jenkins T, Kidd KK (1998) A global haplotype analysis of the myotonic dystrophy locus: implications for the evolution of modern humans and for the origin of myotonic dystrophy mutations. Am J Hum Genet 62:1389–1402. doi:0002-9297/98/6206-0016$02.00

Tofanelli S, Taglioli L, Varesi L, Paoli G (2001) STR-based genetic profile of the population of Corsica island (France). Forensic Sci Int 121:3–38. doi:10.1016/S0379-0738(01)00516-3

Vona G, Moral P, Memmi M, Ghiani ME, Varesi L (2003) Genetic structure and affinities of the Corsican population (France): classical genetic markers analysis. Am J Hum Biol 15:151–163. doi:10.1002/ajhb.10133

Watkins WS, Ricker CE, Bamshad MJ, Carroll ML, Nguyen SV, Batzer MA, Harpending HC, Rogers AR, Jorde LB (2001) Patterns of ancestral human diversity: an analysis of Alu-insertion and restriction-site polymorphisms. Am J Hum Genet 68:738–752. doi:0002-9297/2001/6803-0018$02.00