

# Peopling of Three Mediterranean Islands (Corsica, Sardinia, and Sicily) Inferred by Y-Chromosome Biallelic Variability

P. Francalacci,<sup>1\*</sup> L. Morelli,<sup>1</sup> P.A. Underhill,<sup>2</sup> A.S. Lillie,<sup>2</sup> G. Passarino,<sup>3</sup> A. Useli,<sup>1</sup> R. Madeddu,<sup>1</sup> G. Paoli,<sup>4</sup> S. Tofanelli,<sup>4</sup> C.M. Calò,<sup>5</sup> M.E. Ghiani,<sup>5</sup> L. Varesi,<sup>6</sup> M. Memmi,<sup>6</sup> G. Vona,<sup>5</sup> A.A. Lin,<sup>2</sup> P. Oefner,<sup>7</sup> and L.L. Cavalli-Sforza<sup>2</sup>

<sup>1</sup>Dipartimento di Zoologia e Antropologia Biologica, Università di Sassari, 07100 Sassari, Italy

<sup>2</sup>Department of Genetics, Stanford University, Stanford, California 94305-5120

<sup>3</sup>Dipartimento di Biologia Cellulare, Università della Calabria, 87030 Rende, Italy

<sup>4</sup>Dipartimento di Etologia, Ecologia ed Evoluzione, Università di Pisa, 56123 Pisa, Italy

<sup>5</sup>Dipartimento di Biologia Sperimentale, Università di Cagliari, 09100 Cagliari, Italy

<sup>6</sup>Faculté des Sciences et Techniques, Université de Corse, Corte, France

<sup>7</sup>Stanford Sequencing and Technology Center, Palo Alto, California

**KEY WORDS** Y-chromosome; population genetics; Mediterranean populations

**ABSTRACT** An informative set of biallelic polymorphisms was used to study the structure of Y-chromosome variability in a sample from the Mediterranean islands of Corsica and Sicily, and compared with data on Sardinia to gain insights into the ethnogenesis of these island populations. The results were interpreted in a broader Mediterranean context by including in the analysis neighboring populations previously studied with the same methodology. All samples studied were enclosed in the comparable spectrum of European Y-chromosome variability. Pronounced differences were observed between the islands as well as in the percentages of haplotypes previously shown to have distinctive patterns of continental phylogeography. Approximately 60% of the Sicilian haplotypes are also prevalent in Southern Italy and

Greece. Conversely, the Corsican sample had elevated levels of alternative haplotypes common in Northern Italy. Sardinia showed a haplotype ratio similar to that observed in Corsica, but with a remarkable difference in the presence of a lineage defined by marker M26, which approaches 35% in Sardinia but seems absent in Corsica. Although geographically adjacent, the data suggest different colonization histories and a minimal amount of recent gene flow between them. Our results identify possible ancestral continental sources of the various island populations and underscore the influence of founder effect and genetic drift. The Y-chromosome data are consistent with comparable mtDNA data at the RFLP haplogroup level of resolution, as well as linguistic and historic knowledge. *Am J Phys Anthropol* 121:270–279, 2003. © 2003 Wiley-Liss, Inc.

The contemporary genetic structure of *Homo sapiens* in the Mediterranean basin is an amalgam shaped by numerous prehistoric African and European population influences. This study is based on the concept that living populations on the major islands of the western Mediterranean (Sicily, Sardinia, and Corsica) retain a recoverable genetic record reflective of various episodes of human movements and settlement.

The largest island, Sicily, is closest to the mainland and shares considerable history with the Italian peninsula and the populations of Southern Italy. During the historic period, Sicily was populated by different peoples of both Indo-European (Sicels) and non-Indo-European (Sicans, Elymes) origin (Devoto, 1977). Subsequently, significant influences were imparted by the Phoenicians in the West and Greeks in the East. After the Roman period, Sicily was invaded by German tribes such as the Visigoths, and later ruled by the Arabs and Normans. During these centuries, an appreciable immigration from both

Central-North Europe and North Africa is historically documented. The genetic background of Sicily, as inferred by classical markers, is still debated, possibly a result of the different types of genetic profiling used. Some authors described a western-eastern differentiation, reflecting the division of the

Grant sponsor: Italian National Council of Research; Grant number: CNR 97.00712PF36; Grant sponsor: Interreg III; Grant sponsor: MURST 60%; Grant sponsor: 40% Cofin; Grant sponsor: University of Sassari; Grant number: L.27.12.97, art.51; Grant sponsor: NIH; Grant numbers: GM 28428, GM 55273.

\*Correspondence to: Paolo Francalacci, Dipartimento di Zoologia e Antropologia Biologica, Università di Sassari, Via Regina Margherita 15, 07100 Sassari, Italy. E-mail: antropos@uniss.it

Received 25 January 2001; accepted 9 January 2003.

DOI 10.1002/ajpa.10265

island between Indo-European and non-Indo-European populations (Piazza et al., 1988), while others found genetic homogeneity (Rickards et al., 1998).

Corsica and especially Sardinia are more isolated, although the former can be seen from the Italian mainland. The two islands presumably shared a similar history during their early stages of colonization. Among the human remains of this period there are the right temporal and left maxilla bones from Corbeddu Cave in central Sardinia (radiocarbon-dated to  $8,750 \pm 140$  bp) (Spoor and Sondaar, 1986), and the human female skeleton of Araguina-Sennola (Corsica) dated to  $8,300 \pm 130$  bp (Lanfranchi and Weiss, 1997). Some influence from the Iberian Peninsula is apparent, especially for Sardinia (Cao et al., 1989; Paulis, 1995).

Carthaginian and Roman pressure pushed the indigenous non-Indo-European inhabitants of Sardinia (the Nuragians) into the hilly central region of the island. Both external cultural and genetic contacts have been minimal, making this the most conservative region of the island from a linguistic (Contini et al., 1988–1989) and genetic (Piazza et al., 1988; Cappello et al., 1996) perspective. Corsica has always been scarcely populated, barely approaching 100,000 inhabitants until the end of the 18th century (Memmi, 1999), and possibly experienced a contraction in size during the Middle Ages due to the impacts of Saracen incursions and epidemic diseases. The economy of the island flourished following the Pisan rule that left a footprint epitomized by the Neolatin language. Genoan city-state governance replaced Pisan rule for many centuries, which was eventually followed by French control. In the latter period, the population increased to 250,000 inhabitants. Studies on classical genetic markers have given discordant results, with some reports showing the genetic similarity of Corsica and Sardinia (Varesi et al., 1996; Memmi et al., 1998), and others stressing the genetic diversity between the two neighboring islands (Calafell et al., 1996; Moral et al., 1996).

The stratification of gene pools of the islands was investigated through analysis of binary polymorphisms (on the nonrecombining portion of the Y chromosome (NRY) in a sample of Corsican and Sicilian human Y chromosomes), and evaluated with previous data obtained with the same methodology in central Sardinia (Semino et al., 2000). Subsequently, the perspectives drawn from patterns of Y-chromosome haplotypes were compared with those derived from studies of mitochondrial DNA (mtDNA) coding and noncoding region sequence variation performed on the same islands (Morelli et al., 2000). The formation of a population or a culture can only be completely understood by an integration of comprehensive genetic, linguistic, historical, archaeological, and anthropological data. For this reason, it is not possible to associate automatically a genetic trait with a population, especially considering the recency and relative homogeneity of the hu-

man species, where most of the alleles are shared among populations. However, DNA segments defined by numerous linked polymorphisms (i.e., haplotypes) allow a detailed population genetic analysis of the history of the molecule, thus providing an index of population subdivision and gene flow in a straightforward manner (Harding et al., 1997). The smaller effective population size of the nonrecombining haploid mtDNA and Y-chromosome loci results in a more rapid divergence relative to the autosomes, providing more precise dissection of historic strata. The nonrandom distribution of distinctive stable haplotypes can display a correspondence with geography, thus providing patterns of affinity and clues concerning past human movements. Such genetic data are independent of, but often compatible with, scenarios inferred from nongenetic knowledge. On the European continent, genetic analyses of protein variants and autosomal DNA polymorphisms indicate a southeast-northwest gradient, which was attributed to the demic diffusion of agriculturalists and pastoralists from the Near East that merged with the preexisting Mesolithic population (Menozzi et al., 1978; Ammerman and Cavalli-Sforza, 1984; Sokal et al., 1991; Barbujani et al., 1994).

The peculiar genetic transmission of these systems, matrilineal for mtDNA and patrilineal for the Y chromosome, provides a useful tool for the inference of human evolution, demographic history, and population migration and admixture (Cann et al., 1987; Jobling and Tyler-Smith, 1995; Hammer and Zegura, 1996). Despite being single nonrecombining DNA molecules susceptible to gender-specific population dynamics, and the fact that their smaller effective population size makes them more sensitive to stochastic processes (Edwards and Beerli, 2000), both haploid molecules have been used for population studies at both global and regional levels. Nonrandom distributions of mtDNA and Y-chromosome haplotypes in Europe have been reported, where frequency gradients were detected for some mtDNA haplogroups, such as H, J, T, and V. These have been interpreted as a reflection of past demographic events, including the possible expansion of agriculturalists from the Near East (Richards et al., 2000), the effect of the recolonization of Europe following the Last Glacial Maximum (LGM) (Torroni et al., 2001), and the persistence of lineages with pre-LGM heritage in the mitochondrial pool (Salas et al., 1998).

Similarly, southeast-northwest clines were recognized in Europe from the analysis of several Y-chromosome microsatellites (Semino et al., 1996; Malaspina et al., 1998, 2001; Casalotti et al., 1999; Rosser et al., 2000). These authors considered such clines consistent with the demic diffusion model of agriculture and the transition to a domestic economy, while a minor east-west cline was proposed as a signal of an expansion from north of the Black Sea (Rosser et al., 2000).

Because of high mutability, equivalently defined microsatellite haplotypes may not necessarily have common ancestry. However, binary polymorphisms on the Y chromosome are generally considered unique events (Jobling and Tyler-Smith, 1995), implying that all chromosomes that shared a derived allele descend from a single male ancestor. Recent progress in deciphering the Y-chromosome binary polymorphism genealogy in contemporary populations (Underhill et al., 2000) provides new impetus for reevaluating views of prehistoric affinity and subsequent diversification. In a manner analogous to what has been done for mtDNA (Torroni et al., 1993, 1996; Macaulay et al., 1999), the different haplotypes have been grouped into arbitrary haplogroups identified by mutations occupying internal positions in the phylogeny (Y Chromosome Consortium, 2002). Semino et al. (2000) used the distribution of Y-chromosome binary haplotypes to infer different demographic episodes associated with the colonization of modern humans in Europe. These pan-European data provide a frame of reference for evaluating populations of the Mediterranean basin.

Here we report on the substructure of Y-chromosome binary haplotypes on three western Mediterranean islands with complex histories of colonization, in an attempt to disentangle the differential ethnogenesis of each population. We also compared paternal Y-chromosome and maternal mtDNA perspectives for any evidence of differential sex-specific influences.

## MATERIALS AND METHODS

DNA was extracted from blood sampled in 85 healthy unrelated male individuals using standard procedures (Maniatis et al., 1990). A total of 51 came from western Sicily, and 34 from Corsica (sampled around the towns of Corte in the center of the island, and Bastia in the north). In addition, one individual from Tuscany and one from Sardinia were also analyzed as control samples, and included in the comparisons with other Mediterranean populations.

Since Y-chromosome haplogroups have a nonrandom geographic distribution, we were able to limit the analysis to geographically relevant markers. In total, 17 fragments from the NRY, which contained 20 variable sites (16 SNPs, 3 small insertion/deletions, and 1 Alu insertion) were amplified. Details concerning the NRY markers are given in Underhill et al. (2001).

A single thermal cycling protocol was used. This involves an initial denaturation at 95°C for 10 min to activate AmpliTaq Gold®, 14 cycles of denaturation at 94°C for 20 sec, primer annealing at 63–56°C using 0.5°C decrements, and extension at 72°C for 1 min, followed by 20 cycles at 94°C for 20 sec, 56°C for 1 min, and 72°C for 1 min, and a final 5-min extension at 72°C. Each 50- $\mu$ l PCR reaction contained 1 U of AmpliTaq Gold® polymerase, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 0.1 mM each of the four deoxyribonucleotide triphos-

phates, 0.2  $\mu$ M each of forward/reverse primers, and 50 ng of genomic DNA.

After confirmation of successful amplification by agarose gel analysis, each fragment was combined with an equivalent amplicon generated from DNA of a known allelic state. The mixtures were then denatured at 95°C for 1 min, and then reannealed by cooling to 65°C at a rate of 1°C per minute to potentially form heteroduplex and homoduplex molecules. The hybridized molecules were subsequently genotyped by denaturing high-performance liquid chromatography (D-HPLC), as described in Underhill et al. (1997). The temperature used for analysis of each fragment was determined using the melting algorithm available at: <http://insertion.stanford.edu/melt.html>. The homoduplex and heteroduplex chromatographic results were used to score the allelic state of each marker, and hence to construct the haplotype of each sample.

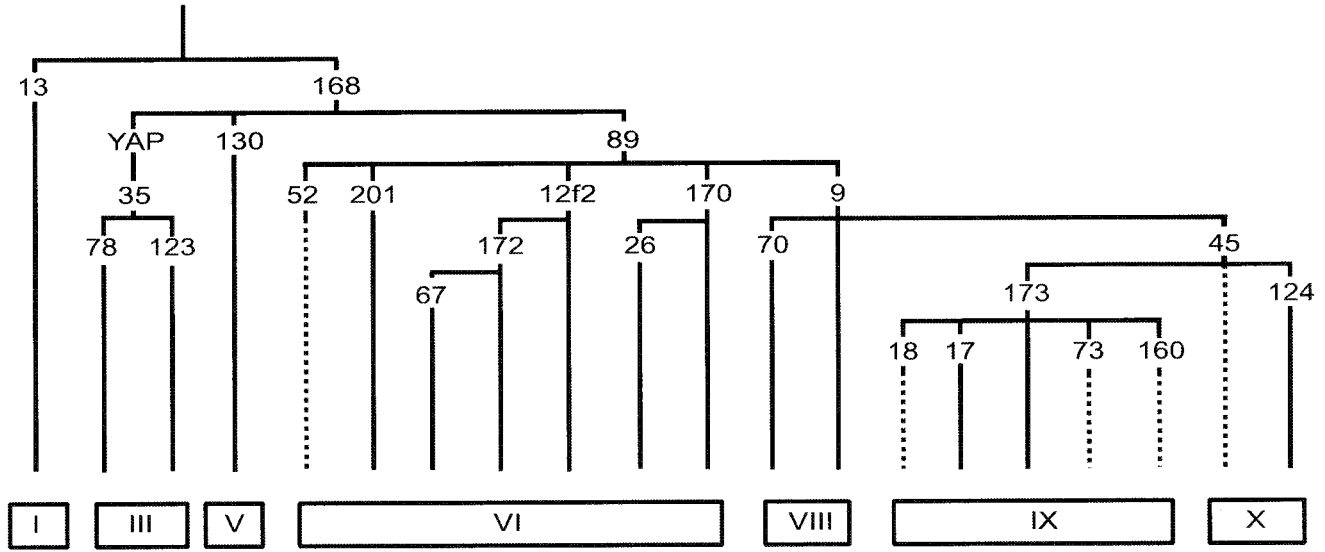
The seven samples showing the M89 marker for which no further downstream markers were detected were subsequently typed at the p12f<sub>2</sub> locus, according to Sun et al. (2000).

Genetic distances were computed from haplotype frequencies according to Nei (1972). A phylogenetic tree was constructed, using the UPGMA method of clustering in the Phylogeny Inference Package, Phylip version 3.2 (Felsenstein, 1989). An unbiased estimate of haplotype diversity,  $h$ , and its variance,  $V(h)$ , were calculated according to the method of Nei (1987, formulas 8.5 and 8.13 therein). The SE of  $h$ ,  $SE(h)$ , was calculated by taking the square root of  $V(h)$ . Genetic relationships between different populations based on haplotype frequencies were further explored by analysis of molecular variance (AMOVA), as implemented in Arlequin (Schneider et al., 2000). Analyses were performed considering haplotype frequencies only. Significance levels of pairwise genetic distances between populations were estimated by 10,000 permutations. Mismatch distributions and  $D$  (Tajima, 1989) and  $F_s$  (Fu, 1997) neutrality statistics were computed for both the Corsica and Sardinia mtDNA RFLP data of Morelli et al. (2000) and the new Sicilian mtDNA RFLP reported here using the Arlequin package. Similar computations were conducted on the companion Y-chromosome data.

## RESULTS

The phylogenetic relationships of the 24 variable Y-chromosome sites defining 20 haplotypes analyzed in this study are shown in Figure 1. Twenty markers were polymorphic. Both the observed mtDNA RFLP and Y-chromosome haplotypes and their frequency by population are specified in Table 1.

The haplotype diversity,  $h$ , and its variance,  $V(h)$ , are shown in Table 2. The three islands show similar values, with a slightly reduced variability in Corsica, possibly due to the smaller sample size. The haplotype diversity values are lowest in the western



**Fig. 1.** Maximum parsimony phylogeny of 24 Y-chromosome markers studied that define 20 haplotypes. Fifteen observed haplotypes are indicated in solid lines. Remaining unobserved 5 haplotypes are shown by dashed lines. Haplogroups are numbered according to Underhill et al. (2000).

**TABLE 1.** Compound Y-chromosome and mtDNA haplotype definitions and frequencies

Haplotype definitions for observed haplogroups: Y-chromosome																							
	13	168	YAP	35	78	123	130	89	201	12f2	172	67	170	26	9	70	45	173	17	124	Sicily	Corsica	Sardinia
13	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
78	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	4
123	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	4
130	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
201	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	6	4	11
12f2	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	6	1	4
172	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	7	1	2
67	0	1	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	9	0	2
170	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	2	3	2
26	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	27
9	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
70	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1
173	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	14	17	18
17	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	1	1	0	2	0	0
124	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	0	1	1
Total																					51	34	78

Haplotype definitions for observed haplogroups: mtDNA												Sicily	Corsica	Sardinia
	12308HinI	9052HaeII	073Aiw44I	7025A/IuI	4595N/a III	8249A/vaII	8994HaeIII	10028A/IuI	14456A/ccI	15606A/IuI	13704B/stNI			
H	0	0	1	1	0	0	0	0	0	0	0	26	24	30
I	0	0	0	0	0	1	0	1	0	0	0	1	0	0
J	0	0	0	0	0	0	0	0	0	0	1	7	6	1
K	1	1	0	0	0	0	0	0	0	0	0	1	7	1
T	0	0	0	0	0	0	0	0	0	1	0	7	5	4
U	1	0	0	0	0	0	0	0	0	0	0	3	6	2
V	0	0	1	0	1	0	0	0	0	0	0	1	0	4
W	0	0	0	0	0	1	1	0	0	0	0	1	0	0
X	0	0	0	0	0	0	0	0	1	0	0	2	5	0
Other	0	0	0	0	0	0	0	0	0	0	0	1	3	4
Total												50	56	46

Mediterranean populations (Catalonia,  $h = 0.37$ ) and highest in the east (Greece,  $h = 0.81$ ). This pattern may reflect the increase of diversity due to the impact of the arrival of haplotypes in Europe

from the Middle East. Table 3 reports the pairwise  $F_{st}$ , calculated from haplotype frequencies, among the eight Mediterranean populations, their significance level, and the overall  $F_{st}$  statistics. More than

TABLE 2. Percentage of haplotypes, haplotype diversity,  $h$ , and its error,  $se(h)$ , observed in this sample and in reference populations

Population	M35	M172	M89	M201	M170	M26	M173	M17	Other	Sample size	$h$	SE( $h$ )
Sicily	5.9	31.4	11.8	11.8	3.9	0.0	27.5	3.9	3.9	51	0.7983	0.0217
Corsica	14.7	2.9	2.9	11.8	8.8	0.0	50.0	0.0	8.8	34	0.7077	0.0488
NC Italy	2.0	14.0	0.0	10.0	8.0	0.0	62.1	4.0	0.0	50	0.5834	0.0509
Calabria	13.5	21.6	10.8	8.1	0.0	0.0	32.4	0.0	13.5	37	0.8042	0.0223
Sardinia	10.3	5.1	5.1	14.1	2.6	34.6	23.1	0.0	5.2	78	0.7930	0.0187
France	8.7	13.0	4.3	0.0	17.4	0.0	52.2	0.0	4.3	23	0.6841	0.0607
Catalonia	4.2	4.2	0.0	8.3	4.2	0.0	79.2	0.0	0.0	24	0.3687	0.0863
Greece	22.4	21.0	1.3	2.6	7.9	0.0	27.6	11.8	5.2	76	0.8108	0.0131

NC, Northern-Central.

TABLE 3.  $F_{st}$  statistics calculated from haplotype frequencies in 8 mediterranean populations, and matrix of pairwise  $F_{st}$  values (below diagonal) and  $F_{st}$  probability values (1,000 permutations, above diagonal)<sup>1</sup>

Percentage of variation	Among populations 8.50			Within populations 91.50		Significance (10,000 permutations) $P < 0.00001$		
	Sicily	Corsica	N. Italy	Calabria	Sardinia	France	Catalonia	Greece
Sicily		0.00330*	0.00060*	0.49875	0.00000*	0.02420*	0.00000*	0.03570*
Corsica	0.06982		0.13669	0.08239	0.00040*	0.59554	0.04780*	0.00740*
NC Italy	0.09146	0.01665		0.00280*	0.00000*	0.40126	0.24218	0.00000*
Calabria	-0.00388	0.02536	0.07593		0.00020*	0.10809	0.00030*	0.17938
Sardinia	0.09760	0.09807	0.16041	0.07977		0.00020*	0.00000*	0.00000*
France	0.05250	-0.01005	-0.00092	0.02752	0.11879		0.05539	0.03560*
Catalonia	0.19194	0.05496	0.01025	0.15976	0.22571	0.05771		0.00000*
Greece	0.02316	0.04796	0.09640	0.00999	0.09652	0.03973	0.17844	

<sup>1</sup> Asterisk indicates significant  $F_{st}$   $P$  values (significance level = 0.05).

NC, Northern-Central.

TABLE 4. Measures of genetic diversity estimated from Y-chromosome and mtDNA data<sup>1</sup>

Population	N	Hp	D	$F_s$	$P(F_s)$	r	$P(r)$
Y-chromosome							
Sicily	51	11	0.0973	-0.3806	0.4860	0.0407	0.3580
Corsica	34	10	-0.1692	-0.8459	0.3890	0.2023	0.0260
Sardinia	78	13	-0.2303	-0.5474	0.4990	0.0888	0.0330
mtDNA							
Sicily	50	10	-0.5704	-1.9980	0.1980	0.3577	0.0070
Corsica	56	7	0.9942	0.6766	0.7120	0.1986	0.0110
Sardinia	46	7	-0.3353	-0.9716	0.3390	0.2662	0.1840

<sup>1</sup> N, sample size; Hp, number of different haplogroups observed; D, Tajima's D;  $F_s$ , Fu's  $F_s$ ;  $P(F_s)$ ,  $P$ -value for  $F_s$ ; r, Harpending's raggedness index;  $P(r)$ ,  $P$ -value for r.

90% of the total variability of the Mediterranean Y-chromosome gene pool is within populations. The variability among populations, although representing only 8.5 of the total value, is highly significant ( $P < 0.00001$ , 10,000 permutations). This value is about a half of that obtained from the same analysis carried out on the whole of Europe, including the Sicilians and Corsicans studied here, the populations reported in Semino et al. (2000), and a Norwegian sample (G. Passarino, unpublished results), where the variance among populations reached 19.11%.

Sardinia was highly differentiated from all the other populations due in large part to the haplotype defined by M26 that is exceedingly rare elsewhere. Corsica is not significantly different from continental Italian and French populations, while the highest significance value occurs in the comparison of Corsica with Sardinia. Sicily was significantly different from all other populations, except Calabria in

southern Italy, with no significant  $F_{st}$  values in this case.

Using the observed mtDNA and Y chromosome haplotypes defined in Table 1, both mismatch distributions and neutrality statistics D and  $F_s$  were computed (Table 4) to investigate if the mtDNA and Y-chromosome data suggest any population dynamic or natural selection differences. When checked against the respective population sizes, all D values were found to be not significantly different from zero. In addition, based on the  $P$ -values generated by the Arlequin package, the  $F_s$  neutrality statistic of Fu (1997) was insignificant in all cases.

Visual inspection of the mismatch distributions (not shown) did not suggest any major difference in modalities. Harpending's "raggedness" indices from the mismatch distributions, along with their respective  $P$ -values, are also shown for each sex-specific system in Table 4. Only Corsica had significant "raggedness" indices for both loci. For this island, the

indices reveal no detectable difference between the mismatch distributions for both molecules. No conclusions could be drawn using the “raggedness” indices for Sicily and Sardinia.

### DISCUSSION

Although the genetic landscape of Europe reflects gene diversification and microgeographic demography, including both size fluctuations, range expansions, and isolation by distance since the LGM (Hewitt, 2000), the associated phylogeography of molecular antecedents to post-LGM diversification provides an opportunity to recover insights of population origins from earlier times (Underhill et al., 2001). The sequential accumulation of bifurcating mutations in the phylogeny is unequivocal, as supported by a consistency index (Kluge and Farris, 1969) of 0.98 for the 218-marker phylogeny of Underhill et al. (2001). However, determining an accurate temporal chronology of specific haplotypes is difficult, given the considerable uncertainties regarding both population dynamics and Y-chromosome microsatellite evolutionarily relevant mutation rates (see note 23, Semino et al., 2000). The level of microsatellite variation associated with specific binary haplotypes provides an index of their relative ages, regardless of the algorithms and assumptions used. Finally, leveraging paleoclimatic and archaeological knowledge provides an independent context for phylogeography (Foley, 1998). A previous survey on the European Y chromosome (Semino et al., 2000) showed haplotype frequency distributions indicating a pronounced substructure affiliated with geography. The habitation of the western Mediterranean islands during the past 10,000 years has a complex history. Nonetheless, the availability of corresponding haplotype data from Europe and adjacent regions (Semino et al., 2000) provides the opportunity to either exclude or localize possible continental sources of post-LGM hunter-gathers and agriculturalists to the islands.

Semino et al. (2000) showed that the major component radiating from Iberia following the LGM involved RFLP 49a,f ht 15, a derivative of the M173 haplotype. Support for this scenario was obtained using independent samples and the P25 = DYS194<sub>469</sub> transversion (Wilson et al., 2001). Another relevant haplotype involves the M170 transversion that arose in situ in Central Europe from an undifferentiated M89 ancestor who may have arrived in Europe via the Levant or Western Asia prior to the LGM. The contemporary distribution of M170-associated lineages is consistent with post-LGM dispersal from the Balkans (Semino et al., 2000). The subsequent derived M170-related M26 mutation defines a haplotype which, while absent in nearby Corsica, occurs at a uniquely high frequency (34.6%) in Sardinia, probably due to founder effect. A similar frequency for this haplotype was reported for another set of Sardinian samples (Passarino et al., 2001). The M26 haplotype also occurs in Basques

at <10% (Semino et al., 2000), and it is reported with lower incidence in Andalusia, northwestern Africa (Bosch et al., 2001), and England (D. Goldstein, personal communication). Assessment of related microsatellite diversity will be necessary to better understand the origin of M26 lineages and their migratory relationship between continental Europe and Sardinia.

The various haplotypes individually defined by either M35, p12f2, M172, or M201 show a cline, with maximum frequency in the Middle East and Asia Minor decreasing towards southwestern Europe. The p12f2 derivative has been considered a signature of demographic events associated with the transition to agriculture (Santachiara-Benerecetti et al., 1993; Semino et al., 1996; Quintana-Murci et al., 2001).

Phylogeography and microsatellite diversity suggest that the M17-defined haplotype and its compound haplotype HG3 analogue (Rosser et al., 2000; Quintana-Murci et al., 2001) differentiated from an M173 predecessor north of the Black Sea during or after the LGM and subsequently radiated into Eastern Europe and Western Asia.

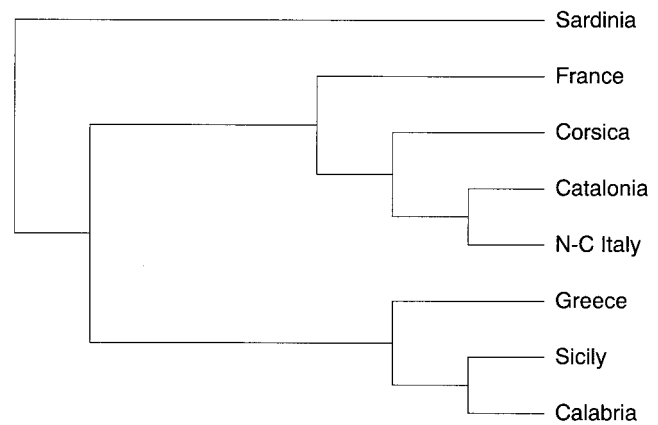
To simplify our presentation, the 20 observed haplotypes are discussed on the basis of their most derived mutation (Fig. 1). The comparison of continental and island Mediterranean Y-chromosome haplotype composition provides insights into the colonization of the major islands of the Mediterranean. Regarding Corsica, the M173 haplotype approaches 60%, while the various 12f2, M35, M172, and M201 haplotypes represent about 30% of the variation. These results are similar to those observed in central-northern Italy. The ratio of 12f2, M35, M172, and M201 haplotypes is also similar to that observed in Sardinia, but the allelic partition of the M173 and M170 components is quite different. In fact, Sardinia shows a very high frequency of the M170-M26 haplotypes (and consequently a lower incidence of M173). In addition, the majority of M170-associated Sardinian chromosomes also show the M26 allele, while this mutation is absent in the 3 Corsican individuals with M170 ancestry. This observation excludes significant gene flow from central Sardinia to central-northern Corsica, and the  $F_{st}$  probability values show a highly significant difference in the comparison between the two islands (Table 3). A larger number of individuals analyzed for the microsatellite YCAIIb by Scozzari et al. (2001) showed an analogous pattern, supporting our hypothesis. A remarkable genetic differentiation between Sardinia (southern and central populations) and Corsica (samples from the north and south of the island) was also pointed out by multilocus analysis of autosomal markers (less influenced than a single-locus system by stochastic variation) such as STRs (Tofanelli et al., 2001) and Alu polymorphisms (Moral et al., 1999).

The frequencies observed in the Corsican population seem to converge toward those detected in con-

tinental Italy, as shown by the  $F_{st}$  probability value that does not reach the significance level (Table 3). This may be due to other distinctive gene flows that overlaid a previous genetic heritage shared with Sardinia. This result is consistent with linguistic data showing that the Corsican language is more closely related to Tuscan than to the Sardinian language (Grimes, 1996). The pattern of affinity is also supported by mtDNA RFLP data (Morelli et al., 2000), but not by pairwise difference analysis (Varesi et al., 2000). This discrepancy could be due in part to the different origin of the samples and partly to the difference in the analytical approach, as the first work is based on phylogenetically relevant polymorphisms of the whole mitochondrial genome, and the second paper focused on the first hypervariable segment of the control region.

The Sardinian sample came from the hilly central region of the island, called Barbagia. Previous studies on classical markers (Cappello et al., 1996) and mtDNA (Morelli et al., 2000) showed a significant differentiation of this sample from surrounding populations (especially the population from the northern area of the island, called Gallura, that is genetically and linguistically similar to Corsica), indicating the persistence of more conservative features in Barbagia. In fact, the incidence of Western haplotypes was higher there, with a signature of an ancient gene flow from the Iberian peninsula characterized by of the mitochondrial haplogroup V (Morelli et al., 2000). In fact, mtDNA haplogroup V was considered by Torroni et al. (2001) a Mesolithic Iberian marker, although this haplogroup was absent in ancient samples (Izagirre and De La Rúa, 1999). The data are also confirmed by the Y-chromosome polymorphisms, with a 2/3 ratio of M173 vs. 12f2, M35, M172, and M201 haplotypes, and the presence of the M26 haplotype.

A more significant impact of the haplotypes of Middle East affiliation is reflected in the Y-chromosome variation in Sicily. The ratio is reversed in central-northern Italy and in the islands of Corsica and Sardinia, being 2:1 in favor of the predominant Middle East haplotypes. This ratio is similar to that observed in Calabria and Greece, in agreement with the geographic and historic proximity of these three regions. The  $F_{st}$  probability value is not significant in comparison with Calabria (Table 3). Sicily also shows a high incidence of M67 (which is absent in Corsica), a subsequently derived lineage from the M172 haplotype. This marker was already observed in the Middle East (Underhill et al., 2000), indicating the impact of people of Mediterranean and Middle Eastern ancestry on the island. In addition, 43% of the Sicilian haplotypes show the 8-Kb allele at the p12f2 locus, proposed as representative of populations associated with the transition to agriculture by various authors (Santachiara-Benecetti et al., 1993; Semino et al., 1996; Quintana-Murci et al., 2001). The Sicilian sample showed two chromosomes defined by the M17 allele. The same haplo-



**Fig. 2.** UPGMA tree of Mediterranean populations. N-C, Northern-Central.

type was also observed in Northern Italy, though it has possibly followed different paths: from northeast in this case, and through the historically well-known Greek influence (where the markers reach the frequency of 11.8%) in the case of Sicily, as shown by the fourth principal component of autosomal genes in Europe (Cavalli-Sforza et al., 1994).

The relationships between the various European populations are illustrated in the tree shown in Figure 2. This tree reveals a split between three major population groups, with relatively good agreement between the geographic and genetic distances. The northwestern Mediterranean populations cluster together, separating from the cluster including the southern Italian and Greek populations. The central Sardinian sample represents the outlier in this tree, reflecting the remarkable influence of genetic drift and of founder effect in its population history.

## CONCLUSIONS

Drawing inferences concerning population history from variation associated with a single sex-specific, nonrecombining DNA sequence must be done with caution. However, the rapid between-population divergence of NRY biallelic variation, and its ability to preserve elements of common ancestry of prehistoric time spans, make it a unique tool for investigating the ethnogenesis of populations, including the relative degree of admixture with preexisting populations, especially in situations of complex demography common in the Mediterranean basin. Integration of binary haplotype-specific associated microsatellite diversity provides evidence concerning the polarity of dispersals. In association with mtDNA analysis, companion high-resolution studies of Y-chromosome diversity provide information concerning possible male/female differences in demographic history. However, even though the two genetic systems share an analogous mode of transmission, they are not directly comparable in terms of absolute frequencies. While asymmetric patterns of male and female migratory behaviour may occur (Seielstad et al., 1998; Oota et al., 2001),

the effect of possible differences in the male vs. female effective population size should also be considered (since males and females may have a different reproductive potential). Also, different mutation rates within and between loci have significant impact on Y-chromosome and mtDNA variability. The quantitative comparison of both loci suggests little difference regarding male and female demography on all three islands. This differs from the results reported for other data (Perez-Lezaun et al., 1999; Pereira et al., 2001). Such differences may reflect different microevolutionary histories or the different types of data used, including sample size as well as Y-chromosome microsatellites vs. binary and mtDNA noncoding sequences vs. RFLP data. In the three populations studied here, the higher percentage (30–60%) of Y-chromosome haplotypes of Middle East affiliation with respect to corresponding mtDNA haplotypes (less than 20%, as reported by Richards et al., 2000) does not necessarily imply a difference in the sex ratio of the newcomer farmers in these islands (since a similar proportion is also present in the Near East), but perhaps nonrandom mating phenomena. In fact, the phylogeography suggests that Y-chromosome haplotypes seem to turn over faster, diagnostic of more bottlenecks with respect to mtDNA, and thus accentuating the frequency clines of Y-chromosome haplotypes relative to mtDNA haplogroups. The three main Mediterranean islands studied show evidence of different patterns of human peopling, with Corsica and Sicily closely associated with neighboring continental populations, while Sardinia shows a marked feature of isolation, with some possible ancient contact with the Iberian Peninsula. These data are in substantial agreement with the trend observed with mtDNA data (Morelli et al., 2000), suggesting that there was no gender differentiation in the population pathways. The linguistic data and historic events of the islands also support this interpretation.

The study of binary polymorphisms on the Y chromosome has proved effective in the reconstruction of the history of the peopling of the central Mediterranean islands. Further, more detailed analyses focused on intraregional variation could shed light on microevolutionary processes and improve our knowledge of the genetic relationships of the Corsican, Sardinian, and Sicilian populations, among themselves and in the Mediterranean framework.

#### ACKNOWLEDGMENTS

This work was supported by grants from the Italian National Council of Research (CNR 97.00712PF36 to L. Terrenato), Interreg III and MURST 60% to P.F., and 40% Cofin (to P.F., G.P., and G.V.). L.M. was awarded a research contract (L.27.12.97, art.51) and support from the Young Researcher Fund from the University of Sassari. NIH grants GM 28428 and GM 55273 went to L.L.C.-S. and P.A.U. We thank O. Semino for sharing unpublished data for the Sardinian samples.

#### LITERATURE CITED

- Ammerman AJ, Cavalli-Sforza LL. 1984. The Neolithic transition and the genetics of populations in Europe. Princeton: Princeton University Press.
- Barbujani G, Pilastro A, De Domenico S, Renfrew C. 1994. Genetic variation in North Africa and Eurasia: Neolithic demic diffusion vs. Paleolithic colonization. *Am J Phys Anthropol* 95:137–154.
- Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, Bertranpetit J. 2001. High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. *Am J Hum Genet* 68:1019–1029.
- Calafell F, Bertranpetit J, Rendine S, Cappello N, Mercier P, Amoroso JP, Piazza A. 1996. Population history of Corsica: a linguistic and genetic analysis. *Ann Hum Biol* 23:237–251.
- Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. *Nature* 325:31–36.
- Cao A, Gossens M, Pirastu M. 1989. Beta thalassaemia mutations in Mediterranean populations. *Br J Haematol* 1:309–312.
- 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. *Am J Hum Genet* 60:125–141.
- Casalotti R, Simoni L, Belledi M, Barbujani G. 1999. Y-chromosome polymorphisms and the origin of the European gene pool. *Proc R Soc Lond [Biol]* 266:1959–1965.
- Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton: Princeton University Press.
- Contini M, Cappello N, Griffo R, Rendine S, Piazza A. 1988–1989. Géolinguistique et géogénétique: une démarche interdisciplinaire. *Geolinguistique* 4:129–197.
- Devoto G. 1977. *Il linguaggio d'Italia*. Milano: Rizzoli Editore.
- Edwards SV, Beerli P. 2000. Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54:1839–1854.
- Felsenstein J. 1989. PHYLIP—phylogeny inference package (version 3.2). *Cladistics* 5:164–166.
- Foley R. 1998. The context of human evolution. *Genome Res* 8:339–347.
- Fu XY. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925.
- Grimes BF. 1996. *Ethnologue: languages of the world*. Dallas: SIL International.
- Hammer MF, Zegura SL. 1996. The role of Y chromosome in human evolutionary studies. *Evol Anthropol* 5:116–134.
- Harding RM, Fullerton SM, Griffiths RC, Bond J, Cox MJ, Schneider JA, Moulin DS, Cleeg JB. 1997. Archaic African and Asian lineages in the genetic ancestry of modern humans. *Am J Hum Genet* 60:772–789.
- Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913.
- Izagirre N, De La Rua C. 1999. An mtDNA analysis in ancient Basque populations: implications for haplogroup V as a marker for a major Paleolithic expansion from southwestern Europe. *Am J Hum Genet* 65:199–207.
- Jobling MA, Tyler-Smith C. 1995. Father and sons: the Y chromosome and human evolution. *Trends Genet* 11:449–456.
- Kluge AG, Farris JS. 1969. Quantitative phyletics and the evolution of anurans. *Syst Zool* 18:1–32.
- Lanfranchi F, Weiss MC. 1997. L'aventure humaine préhistorique en Corse. Ajaccio: Albiana.
- Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonnè-Tamir B, Sykes B, Torrioni A. 1999. The emerging tree of west Eurasian mtDNAs; a synthesis of control-region sequences and RFLPs. *Am J Hum Genet* 64:232–249.
- Malaspina P, Cruciani F, Ciminelli BM, Terrenato L, Santolamazza P, Alonso A, Banyko J, Brdicka R, Garcia O, Gaudiano C, Guanti G, Kidd KK, Lavinha J, Avila M, Mandich P, Moral P, Qamar R, Mehdi SQ, Ragusa A, Stefanescu G, Caraghin M, Tyler-Smith C, Scozzari R, Novelletto A. 1998. Network anal-

- yses of Y-chromosomal types in Europe, northern Africa, and western Asia reveal specific patterns of geographic distribution. *Am J Hum Genet* 63:847–860.
- Malaspina P, Tsopanomalou M, Duman T, Stefan M, Silvestri A, Rinaldi B, Garcia O, Giparaki M, Plata E, Kozlov AI, Barbujani G, Vernesi C, Papola F, Ciavarella G, Kovatchev D, Kerimova MG, Anagnou N, Gavrila L, Veneziano L, Akar N, Loutradis A, Michalodimitrakis EN, Terrenato L, Novelletto A. 2001. A multistep process for the dispersal of a Y chromosomal lineage in the Mediterranean area. *Ann Hum Genet* 65:339–349.
- Maniatis T, Fritsch EF, Sambrook J. 1990. *Molecular cloning: a laboratory manual* (second edition). Cold Spring Harbor: Cold Spring Harbor Laboratory.
- Memmi M. 1999. *Etude génétique de la population corse et sa relation avec les populations "euro-méditerranéennes"*. Doctoral thesis. Corte: Université de Corse.
- Memmi M, Moral P, Calò 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.
- Menozzi P, Piazza A, Cavalli-Sforza LL. 1978. Synthetic maps of human gene frequencies in Europeans. *Science* 201:786–792.
- Moral P, Memmi M, Varesi L, Mameli GE, Succa V, Gutierrez B, Lutken N, Vona G. 1996. Study on the variability of seven genetic serum protein markers in Corsica (France). *Anthropol Anz* 54:97–107.
- Moral P, Bao M, Gonzalez E, Lopez-Alomar A, Varesi L, Memmi M, Vona G. 1999. Alu polymorphisms in Corsica and Sardinia: new evidence for the genetic position of Corsican population within western Mediterranean. *Antropol Contemp* 22:77–86.
- Morelli L, Grosso MG, Vona G, Varesi L, Torroni A, Francalacci P. 2000. Frequency distribution of mitochondrial haplogroups in Corsica and Sardinia. *Hum Biol* 72:585–595.
- Nei M. 1972. Genetic distance between populations. *Am Nat* 106:283–292.
- Nei M. 1987. *Molecular evolutionary genetics*. New York: Columbia University Press.
- Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilineal versus patrilineal residence. *Nat Genet* 29:20–21.
- Passarino G, Underhill PA, Cavalli-Sforza LL, Semino O, Pes G, Carru C, Ferrucci L, Franceschi C, Deiana L, Baggio G, De Benedictis G. 2001. Use of Y chromosome binary markers to study the high prevalence of males in Sardinian centenarians and the genetic structure of the Sardinian population. *Hum Hered* 52:136–139.
- Paulis G. 1995. *Lingue e popoli della Sardegna preistorica e protostorica*. *Antropol Contemp* 18:37–40.
- Pereira L, Dupanloup I, Rosser ZE, Jobling MA, Barbujani G. 2001. Y-chromosome mismatch distributions in Europe. *Mol Biol Evol* 18:1259–1271.
- Perez-Lezaun A, Calafell F, Comas D, Mateu E, Bosch E, Martinez-Arias R, Clarimon J, Fiori G, Luiselli D, Facchini F, Pettener D, Bertranpetit J. 1999. Sex-specific migration patterns in Central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA. *Am J Hum Genet* 65:208–219.
- Piazza A, Cappello N, Olivetti E, Rendine S. 1988. A genetic history of Italy. *Ann Hum Genet* 52:203–213.
- Quintana-Murci L, Krausz G, Zerjal T, Sayar SH, Hammer MF, Mehdi SQ, Ayub Q, Qamar R, Mohyuddin A, Radhakrishna U, Jobling MA, Tyler-Smith C, McElreavey K. 2001. Y-chromosome lineages trace diffusion of people and language in southwestern Asia. *Am J Hum Genet* 68:537–542.
- Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, Villems R, Thomas M, Rychkov S, Rychkov O, Rychkov Y, Gölge M, Dimitrov D, Hill E, Bradley D, Romano V, Cali F, Vona G, Demaine A, Papiha S, Triantaphyllidis C, Stefanescu G, Hatina J, Belledi M, Di Rienzo A, Novelletto A, Oppenheim A, Nørby S, Al-Zaheri N, Santachiara-Benerecetti S, Scozzari R, Torroni A, Bandelt HJ. 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* 67:1251–1276.
- 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.
- Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Corte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Nørby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Roots S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 67:1526–1543.
- Salas A, Comas D, Lareu MV, Bertranpetit J, Carracedo A. 1998. mtDNA analysis of the Galician population: a genetic edge of European variation. *Eur J Hum Genet* 6:365–375.
- Santachiara-Benerecetti AS, Semino O, Passarino G, Torroni A, Brdicka R, Fellous M, Modiano G. 1993. The common Near-Eastern origin of Ashkenazi and Sephardi Jews supported by Y-chromosome similarity. *Ann Hum Genet* 57:55–64.
- Schneider S, Roessli D, Excoffier L. 2000. ARLEQUIN version 2.0: a software for population genetics data analysis. Geneva: Genetics and Biometry Laboratory, University of Geneva. <http://anthropologie.unige.ch/arlequin>
- Scozzari R, Cruciani F, Pangrazio A, Santolamazza P, Vona G, Moral P, Latini V, Varesi L, Memmi MM, Romano V, De Leo G, Gennarelli M, Jaruzelska J, Villems R, Parik J, Macaulay V, Torroni A. 2001. Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region. *Hum Immunol* 62:871–884.
- Seielstad MT, Minch E, Cavalli-Sforza LL. 1998. Genetic evidence for a higher female migration rate in humans. *Nat Genet* 20:278–280.
- Semino O, Passarino G, Brega A, Fellous M, Santachiara-Benerecetti AS. 1996. A view of the Neolithic demic diffusion in Europe through two Y chromosome-specific markers. *Am J Hum Genet* 59:964–968.
- Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Limborska S, Kouvatsi A, Marcikiae M, Primorac D, Santachiara-Benerecetti S, Cavalli-Sforza LL, Underhill PA. 2000. The genetic legacy of Palaeolithic *Homo sapiens sapiens* in extant Europeans: a Y-chromosome perspective. *Science* 290:1155–1159.
- Sokal RR, Oden NL, Wilson C. 1991. Genetic evidence for the spread of agriculture in Europe. *Nature* 351:143–145.
- Spoor CF, Sondaar PY. 1986. Human fossils from the endemic island fauna of Sardinia. *J Hum Evol* 15:399–408.
- Sun C, Skaletsky H, Rozen S, Gromoll J, Nieschlag E, Oates R, Page DC. 2000. Deletion of azospermia factor a (AZFa) region of human Y chromosome caused by recombination between HERV15 proviruses. *Hum Mol Genet* 9:2291–2296.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Tofanelli S, Taglioli L, Varesi L, Paoli G. 2001. The STR-based genetic profile of the population from Corsica island (France). *Forensic Sci Int* 121:33–38.
- Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC. 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet* 53:563–590.
- Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari A, Obinu D, Savontaus ML, Wallace DC. 1996. Classification of European mitochondrial DNA from an analysis of 3 European populations. *Genetics* 144:1835–1850.
- Torroni A, Bandelt HJ, Macaulay V, Richards M, Cruciani F, Rengo C, Martinez-Cabrera V, Villems R, Kivisild T, Metspalu E, Parik J, Tolk HV, Tambets K, Forster P, Karger B, Francalacci P, Rudan P, Janicijevic B, Rickards O, Savontaus ML, Huoponen K, Laitinen V, Koivumäki S, Sykes B, Hickey E,

- Novelletto A, Moral P, Sellitto D, Coppa A, Al-Zaheri N, Santachiara-Benerecetti AS, Semino O, Scozzari R. 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. *Am J Hum Genet* 69:844–852.
- Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ. 1997. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high performance liquid chromatography (DHPLC). *Genome Res* 7:996–1005.
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Hurlbut A, Bonn -Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Soergel D, Sherry S, Feldman MW, Davis RW, Cavalli-Sforza LL, Oefner PJ. 2000. Y Chromosome variation and the history of human populations. *Nat Genet* 26:358–361.
- Underhill PA, Passarino G, Lin AA, Shen P, Miraz n-Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65:43–62.
- Varesi L, Memmi M, Moral P, Mameli GE, Succa V, Vona G. 1996. La distribution de quatorze marqueurs g n tiques dans la population de l'ile de Corse (France). *Bull Mem Soc Anthropol Paris* 8:5–14.
- Varesi L, Memmi M, Cristofari MC, Mameli GE, Cal  CM, Vona G. 2000. Mitochondrial control region sequence variation in the Corsican population (France). *Am J Hum Biol* 12:339–351.
- Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB. 2001. Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc Natl Acad Sci USA* 98:5078–5083.
- Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y-chromosome binary haplotypes. *Genome Res* 12:339–348.