

This is the peer reviewed version of the following article:

Spatial diffusion of surnames by long transhumance routes between highland and lowland: A study in Sardinia / Orrù, A.; De lasio, S.; Frederic, P.; Girotti, M.; Boano, R.; Sanna, E.. - In: HOMO. - ISSN 0018-442X. - 69:3(2018), pp. 127-138. [10.1016/j.jchb.2018.06.006]

*Terms of use:*

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

25/04/2024 05:48

(Article begins on next page)

## **Spatial diffusion of surnames by long transhumance routes between highland and lowland: a study in Sardinia**

Orrù A.<sup>1\*</sup>, De Iasio S.<sup>2</sup>, Frederic P.<sup>3</sup>, Girotti M.<sup>1</sup>, Boano R.<sup>1</sup>, Sanna E.<sup>4</sup>

1 Department of Life Sciences and Systems Biology, University of Turin (Italy)

2 Department of Biosciences, University of Parma (Italy)

3 Department of Economics “Marco Biagi”, RECent (Center for Economic Research), University of Modena (Italy)

4 Department of Environmental and Life Sciences, University of Cagliari (Italy)

Abbreviated title: **Spatial diffusion of surnames in Sardinia**

**\*Corresponding author:** Dr. Alessia Orrù, Department of Life Sciences and Systems Biology, Via A. Albertina, 13, University of Turin (Italy).

Tel.: +39 0116704551

e-mail: aorru@unito.it

## **Abstract**

We explored the relationships among Sardinian populations by means of a spatial analysis of surnames in five villages in the historical-geographical zone of Barbagia di Belvì, a mountainous area traditionally devoted to sheep-rearing and the point of departure of transhumance toward lowland areas. We collected the surnames of 19<sup>th</sup> century Sardinian populations through the *Status Animarum* (parish censuses). The structures of past populations were compared with current ones based on surnames reported in telephone directories. The lowland villages studied have been the final destination of transhumance and have a different historical, cultural and economic background. The spread of surnames in Sardinia may have occurred also by means of transhumance, occurring every year along precise routes from the pastoral mountain zones to the agricultural plains.

We used the standardized index of Chen and Cavalli-Sforza to calculate the relationships among the five villages of Barbagia di Belvì (Aritzo, Belvì, Desulo, Gadoni and Tonara). Application of non-metric Multidimensional Scaling to the isonymy matrices showed that the villages of Barbagia di Belvì form a group that has changed very little over time. Transhumance routes were studied by Spatial Autocorrelation (Moran's *I*) applied to surnames. The results suggest that there has been an appreciable admixture between the Sardinian populations of the mountain villages of the central areas and the populations of southern lowland villages.

## **Introduction**

Surnames have a cultural origin and reflect the ethno-historical, socioeconomic and linguistic background of a given population, especially relevant in the context of population mobility (Colantonio et al., 2003; Cheshire, 2014). In “Western” countries, surnames are transmitted from father to children and are similar to a genetic marker, comparable to neutral and highly polymorphic alleles of a Y-chromosome locus (Lasker, 1985; Crow, 1983; King and Jobling, 2009a, b; King et al., 2006; Lucchetti et al., 1987; Colantonio et al., 2003; Sykes and Irven, 2000; Jobling, 2001; Jobling and Tyler-Smith, 2003; Graf et al., 2010; Boattini et al., 2012; Darlu et al., 2012; Martínez-Gonzalez et al., 2012; Winney et al., 2012). It is usual to select autochthonous DNA donors based on the original surnames of the place of residence (Manni et al., 2005; Winney et al., 2012). Using this specific surname-based approach, several genetic studies observed historical population genetic structure or past admixture events that were otherwise invisible (Hill et al., 2000; McEvoy and Bradley, 2006; Bowden et al., 2008; Larmuseau et al., 2012).

The generalized origin of surnames in Italy was decreed by ecclesiastical regulations (*Rituale Romanum*) of the Council of Trent (1545-1563). These rules also prescribed a census of the Catholic population in special registers (*Status Animarum*) indicating the individuals by their first

name and surname.

In many populations, civil sources can also be used: lists of surnames for the entire population (family heads, telephone subscribers, users of other utilities, etc.) are readily available. The relatively easy sampling, similarity with genetically inherited traits and high polymorphism make surnames particularly suitable for the study of microevolutionary processes in populations from the 17<sup>th</sup> century onward (Lucchetti et al., 2008).

In this paper we analyze the spread of the typical surnames from the central area of Sardinia.

Studies on the distribution of surnames in Sardinia have demonstrated that: they act like neutral alleles (Zei et al., 1983a, b, 1986); the migration matrices determined from the temporal differences in surname distributions agree with what is known about the history of the population and its mobility (Wijsman et al., 1984); the temporal variation of the frequency of surnames is 10 times lower than the spatial variation (Zei et al., 1986); the similarities among village populations based on the relative distribution of erythrocyte markers agree with the similarities among villages based on surname distributions (Lucchetti et al., 1987); the isonymic relationships among various Sardinian villages follow the isolation by distance model and the isonymy among nearby villages increases in time (Sanna et al., 2001, 2006). Moreover the genetic homogeneity of Sardinians, shown by the distribution of some Y-chromosome genetic markers, is not confirmed when the samples are formed on the basis of the zone of origin of the surnames of sampled individuals rather than their zone of residence (Zei et al., 2003).

It is also noteworthy that the isonymic difference/similarity among various municipalities of Sardinia is conditioned by their geographical location and historical background (Sanna et al., 2006; Orrù et al., 2008), while the Sardinian municipalities form a particular and well defined cluster when compared with other Mediterranean communities (Lucchetti et al., 2008).

The island has presented two different but complementary socioeconomic and cultural realities through the centuries. One, typical of mountainous areas, is characterized mainly by pastoral activities, while the other, typical of lowland areas, is characterized mainly by agricultural activities (Sanna and Danubio, 2008). This productive system remained the same until the second half of the 20<sup>th</sup> century (Sanna, 2006).

There were substantially two types of transhumance in Sardinia: short-range, involving movements within the same municipal territory or toward nearby villages; long-distance, from the western slopes of the Gennargentu Massif in central Sardinia to the south-western plains (Le Lannou, 1941; Ortu, 1988; Sanna and Danubio, 2008).

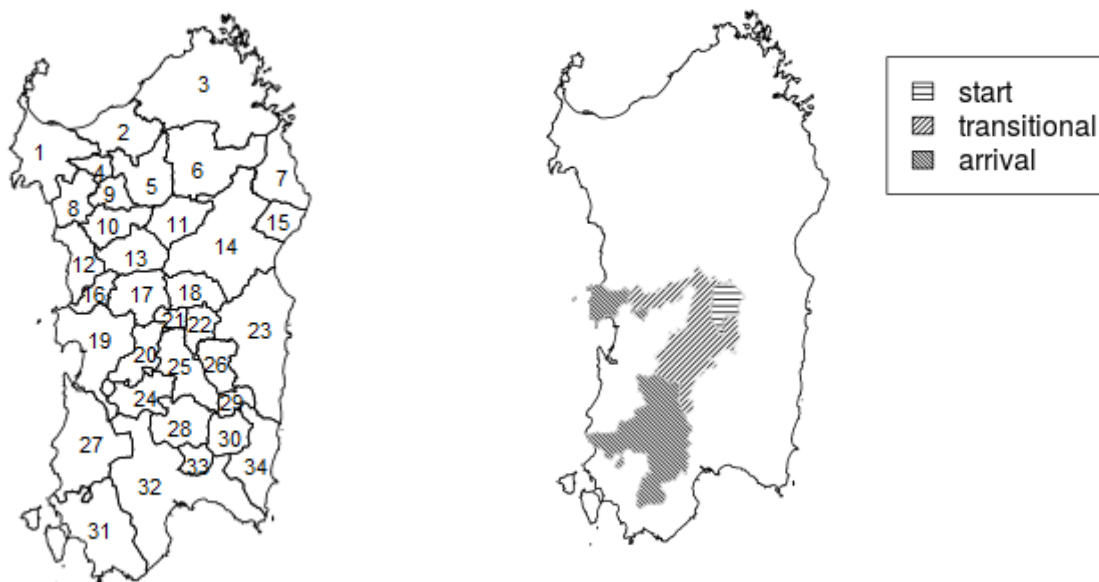
The Sardinian shepherds who practiced long-distance transhumance remained for about six months in the plains, where the flocks wintered: transhumance started in December and ended at the end of

April (Meloni, 1982; Sanna and Danubio, 2008). The transhumance covering the longest distances started in the villages forming Barbagia di Belvì and Barbagia di Ollolai, with migrations to the wintering areas of up to 150 km. To reach the wintering areas, the flocks of Barbagia had to cover many kilometers of highlands and plains densely populated by farmers; sometimes, en route, this created conflicts with the owners of the lands bordering the winter pastures (Le Lannou, 1941).

The spread of surnames in Sardinia may also have occurred by means of long-distance transhumance, occurring every year along precise routes from the pastoral mountain zones to the agricultural plains.

Historically, the territory of the island is divided into 34 zones (Fig. 1). This categorization (Ghiani Moi, 1964) provides important information about the territorial distribution of people in Sardinia: people living in the same zone have likely shared the same habits, living conditions, food, meteorological features, etc. We considered a subset of these zones, i.e. only the ones involved in transhumance. The point of departure of transhumance toward lowland areas was the zone of Barbagia di Belvì, a mountainous area traditionally devoted to sheep-rearing. From the genetic point of view, the population of this zone is considered the most conservative of the island (Cappello et al., 1996; Calò et al., 1998; Vona and Calò, 2006).

Figure 1



- |                      |               |              |
|----------------------|---------------|--------------|
| 2 Anglona            | 30 Gerrei     | 14 Nuorese   |
| 20 Arborea           | 11 Goceano    | 1 Nurra      |
| 22 Barbagia di Belvì | 27 Iglesiente | 23 Ogliastra |
| 4                    |               |              |

18	Barbagia di Ollolai	5	Logudoro	33	Parteolla
26	Barbagia di Seulo	21	Mandrolisai	12	Planargia-Montiferru
15	Baronia di Orosei	13	Marghine	29	Quirra
7	Baronia di Siniscola	24	Marmilla	25	Sarcidano
32	Campidano di Cagliari	9	Meilogu	34	Sarrabus
19	Campidano di Oristano	6	Monte Acuto	31	Sulcis
10	Campu Giavesu	16	Monte Urtigu	17	Tirso
4	Coracense	8	Monteleone	28	Trexenta
3	Gallura				

We obtained the frequencies of surnames of these mountain populations in the 19<sup>th</sup> century and in recent times and compared them with the current distribution of surnames in the lowland villages. These lowland populations live in a cultural and economic environment very different from that of the populations of Barbagia di Belvì.

The aims of this paper are: 1) to highlight the relationships among the Sardinian populations in relation to transhumance, and 2) to determine if transhumance resulted in appreciable admixture between the pastoral populations of the mountains and the agricultural populations of the plains.

## Methods

The populations considered were those of the villages of Aritzo, Belvì, Desulo, Gadoni and Tonara, which form the historical-geographical zone of Barbagia di Belvì (Ghiani Moi, 1964), and the populations of the lowland areas, which were the traditional destination of transhumance.

In order to describe the structure of surnames in the ancient population of Barbagia di Belvì, we used the books of the *Status Animarum*, the ecclesiastic books of censuses in which the priest recorded the name and surname of the head of the household and the other family members.

The registers of all the villages of Barbagia are available only since 1840; before that, there are no years for which the parish books of all five villages are present contemporaneously. Nevertheless, we extended the analysis by considering the *Status Animarum* books of 1815 for Aritzo, Belvì and Gadoni, of 1816 for Desulo, and of 1811 for Tonara; in addition, we consulted the *Status Animarum* of 1844 for Desulo. We compared these data with the recent distribution of surnames in the transit and destination villages of transhumance obtained from telephone subscribers in 1993. The telephone directory lists the people who reflect the family groups living in a given area, so much so that this device has been frequently used in various demo-ecological and genetic analyses of Italian populations (Boattini et al., 2012; Risso et al., 2015).

Figure 2 shows the geographical distribution of the villages of Barbagia di Belvì and their long-distance transhumance routes.

Figure 2



Table 1 reports the number of records of the *Status Animarum* registered in the years considered; from each act, we extracted the surname of the householders to identify the original surnames in Barbagia di Belvì. This table also reports the number of surnames present in the telephone directories of 1993.

Table 1 - Villages of Barbagia di Belvì. Years examined, number of householders, number of individuals present in the telephone directories. In brackets, the percentage of different surnames with respect to the total number of individuals considered

	<i>Status Animarum</i>					Telephone directories
	YEARS					YEAR
	1811	1815	1816	1840	1844	1993
<i>Villages</i>						
<b>Aritzo</b>	-	580 (18.45)	-	479 (19.83)	-	483 (25.26)
<b>Belvi</b>	-	225 (30.66)	-	205 (29.76)	-	215(40.00)
<b>Desulo</b>	-	-	473 (18.60)	463 (17.49)	455 (17.58)	755 (25.16)
<b>Gadoni</b>	-	83 (51.81)	-	206 (27.18)	-	314 (29.94)
<b>Tonara</b>	651 (15.36)	-	-	655 (16.18)	-	729 (24.55)

The isonymic relationship between populations was calculated with the index of “standardized isonymy” proposed by Chen and Cavalli-Sforza (1983):

where  $R_{ij}$  is the coefficient of similarity between two populations  $i$  and  $j$ ,  $x_i$  and  $x_j$  are the frequencies in them of the  $x$ -th surname, and the total is extended to all the surnames. The value of  $R_{ij}$  is zero when the two populations do not have any surname in common; it is 1 when the two populations share the same surnames with equal frequencies, independently of their numbers.

The isonymy index was used to produce a similarity matrix between the populations of the five villages of Barbagia di Belvì, the starting area of transhumance. Non-metric Multidimensional Scaling (nmMDS) was used to produce a synthetic plot of the isonymic relationships among the five Sardinian villages and their variation in time. The results showed that the five villages have a peculiar surname structure, very diversified and self-preserved over time. This allowed us to proceed with a subsequent spatial investigation.

When sampling locations are known, the association between genetic and geographical distances can be tested by spatial autocorrelation or geographically weighted regression methods.

Spatial autocorrelation analysis was applied to the quantitative data summarizing the genetic similarity between populations in relation to their proximity in space. In this specific case, spatial autocorrelation allows us to determine the similarity of values of the frequency of a surname between pairs of localities within arbitrary distance classes. This method allowed estimation of the spatial distribution of the selected surnames in the starting, transitional and arrival areas of transhumance in order to emphasize the specific diffusion processes of the individuals independently of the migration rates of the various localities. Spatial autocorrelation was developed by Moran (1950) and refined by Ripley (1981) and by Cliff and Ord (1981). It was first applied to biological problems by Sokal and Oden (1978a, b). The computation of spatial autocorrelation requires a set of localities represented as points in the plane (Caravello and Tasso, 1999). We used Sardinian municipalities located along transhumance routes and compared the results with the remaining municipalities of the island not involved in transhumance. The coefficient of autocorrelation (Moran's  $I$ ) is calculated as:

where  $x$  represents the surname,  $w_{ij}$  is a proximity weight representing the degree of relatedness between location  $i$  and  $j$  (greater relatedness implies a higher weight);  $n$  is the number of locations;



and  $f_j$  are the frequency of name  $x$  at location  $i$  and  $j$  respectively. Moran's Index has been widely used in surname analyses (Caravello and Tasso, 1999, 2007; Tasso and Caravello, 2010; Román-Busto et al., 2013). The  $I$  coefficient generally varies between -1 (negative autocorrelation) and +1 (positive autocorrelation). In the absence of spatial autocorrelation, the expected value is  $-1/(n - 1)$ , where  $n$  is the size of the sample. A significant negative autocorrelation indicates that at a determined distance the frequencies of a variable are dissimilar, while a significant positive autocorrelation indicates a similarity between populations; a non-significant value shows no relationship between pairs of frequencies of a surname at a certain distance (Sokal and Oden, 1978a).

In this study we selected the 18 most common surnames extracted from the ancient registers of parish censuses of the considered villages. We think that these surnames have a low level of polyphyletism, since they are absent or present at very low frequencies in the other parishes for which it was possible to obtain data from the 19<sup>th</sup> century.

We explored two different proximity matrices, one based on transhumance regions and one based on non-transhumance regions.

In the transhumance matrix, weights were chosen such that

Instead, matrix is such that

Both, and have zeros in diagonal.

Our goal was to check whether some surnames are clustered in transhumance regions. Thus we tested for independence of name  $x$  in transitional regions.

## Results

Table 2 shows the most common surnames originating in Barbagia di Belvì, derived from data of the 19<sup>th</sup> century *Status Animarum*, and the relative frequencies in the historical-geographical areas involved in transhumance. Figure 3 shows where these frequencies are higher (dark areas) or lower (white areas).

*Table 2 – Surnames originating in Barbagia di Belvì and their current relative frequencies in the historical-geographical zones involved in transhumance*

Tirso	0.323	0.000	0.017	0.000	0.527	0.883	0.000	0.000	0.000
Camp. Oristano	0.079	0.007	0.094	0.158	0.475	0.457	0.007	0.004	0.004
Arborea	0.000	0.000	0.306	0.000	0.866	0.739	0.000	0.000	0.000
Mandrolisai	0.000	0.000	0.281	0.000	2.532	0.633	0.070	0.000	0.000
Barbagia Belvì	0.481	1.002	0.601	0.280	2.324	2.043	1.923	2.003	0.521
Marmilla	0.013	0.000	0.281	0.067	0.548	0.735	0.013	0.000	0.000
Sarcidano	0.091	0.036	0.018	0.073	0.365	0.201	0.055	0.000	0.000
Seulo	0.995	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Iglesiente	0.118	0.004	0.181	0.038	0.337	0.573	0.004	0.034	0.000
Trexenta	0.078	0.011	0.145	0.089	0.413	0.983	0.000	0.011	0.000
Sulcis	0.130	0.032	0.433	0.007	0.364	0.512	0.004	0.018	0.000
Camp. Cagliari	0.006	0.000	0.206	0.006	0.467	0.309	0.000	0.067	0.000

	PORRU	SAU	SECCI	TODDE	TORE	URRU	VACCA	ZANDA	ZUCCA
Tirso	0.017	0.000	0.085	0.000	0.068	0.662	0.119	0.051	0.357
	0.155	0.036	0.148	0.054	0.040	0.151	0.507	0.058	0.446
Arborea	0.306	0.000	0.790	0.153	0.000	0.612	0.178	0.000	1.504
	0.000	0.211	0.141	0.492	0.211	1.828	0.281	0.070	0.000
Barbagia Belvì	1.362	2.564	1.322	1.122	0.761	0.681	1.042	1.683	0.441
Marmilla	0.053	0.000	0.067	0.013	0.000	0.000	1.096	0.027	0.214
Sarcidano	0.310	0.018	0.966	0.018	0.000	0.365	0.547	0.018	0.729
Seulo	0.306	0.000	0.230	0.842	0.000	0.000	0.153	0.000	0.000
Iglesiente	0.093	0.004	0.274	0.097	0.017	0.109	0.770	0.316	0.198
Trexenta	0.782	0.000	0.324	0.011	0.000	0.067	0.514	0.045	0.101
Sulcis	0.108	0.014	0.368	0.065	0.025	0.123	0.649	0.054	0.133
Camp. Cagliari	0.133	0.018	0.582	0.097	0.006	0.036	0.430	0.115	0.424

Figure 3

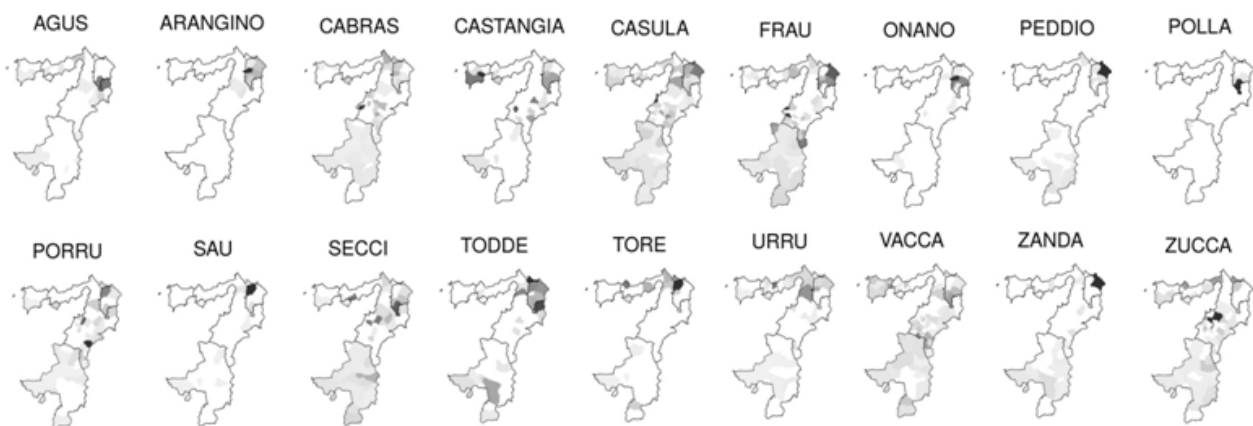
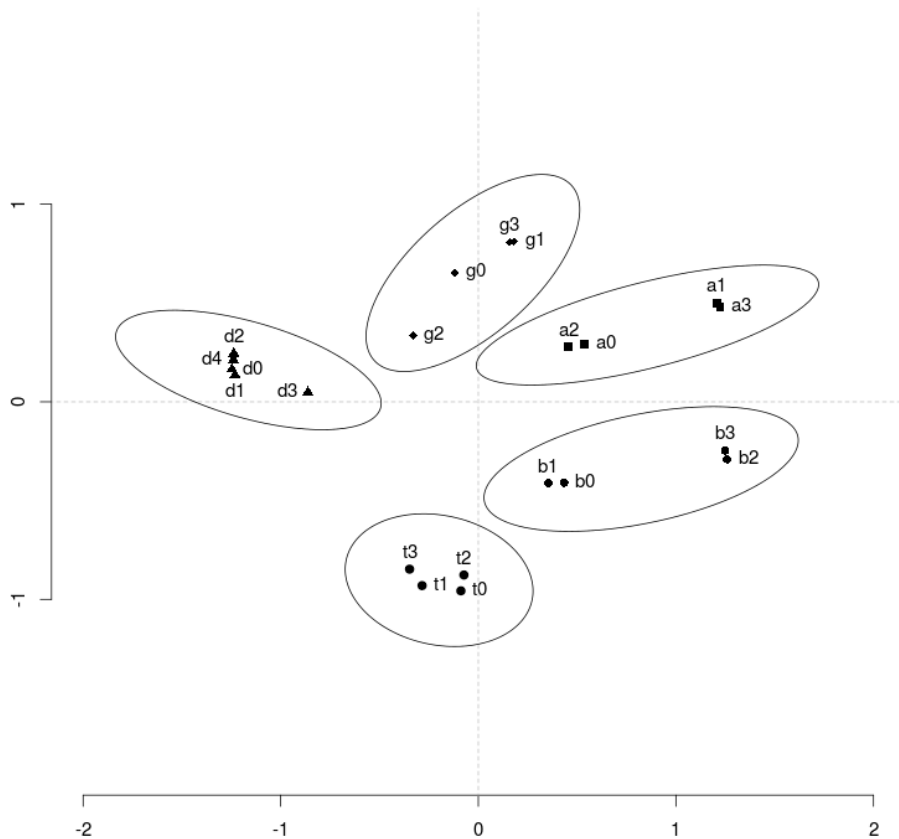


Figure 4



<i>Villages</i>	<i>code</i>	<i>Villages</i>	<i>code</i>	<i>Villages</i>	<i>code</i>	<i>Villages</i>	<i>code</i>
Aritzo 1815	a0	Belvi 1840	b1	Desulo 1844	d2	Gadoni 1993	g2
Aritzo 1840	a1	Belvi 1993	b2	Desulo 1994	d3	Tonara 1811	t0
Aritzo 1994	a2	Desulo 1816	d0	Gadoni 1815	g0	Tonara 1840	t1
Belvi 1815	b0	Desulo 1840	d1	Gadoni 1840	g1	Tonara 1994	t2

Figure 4 illustrates the relationships among the villages of Barbagia di Belvì and their variations in time. The populations are clearly differentiated on the scatter plot and each of them maintains its specific identity in time. The first dimension separates Aritzo and Belvì from Desulo and Tonara, while Gadoni occupies a central, intermediate position. Desulo and Tonara form a closed cluster without differentiation in time; they are the villages with the most homogeneous structure of surnames. Population growth in the first half of the 20<sup>th</sup> century seems not to have changed the surname structure of the two villages. This was probably because of their greater geographical isolation due to their higher altitude (*ca.* 900 m a.s.l.). The other three villages exhibit two small distinct clusters, one including the 19<sup>th</sup> century populations and the other the 20<sup>th</sup> century

populations. This indicates that changes have occurred in the interval between the surveys. Nevertheless, these villages maintained their specificity of surname structure, since they are differentiated from each other on the basis of both surveys (19<sup>th</sup> and 20<sup>th</sup> centuries). The temporal differentiation observed in the other three villages indicates that the isolation of Desulo and Tonara continued while the other three villages tended to change, moving away from the other ones that represent the most isolated part of the Barbagia di Belvì.

Table 3 shows the values of Moran's  $I$  with . Note that target surnames such as "Arangino", "Castangia", "Casula", "Onano", "Peddio", "Polla", "Sau", "Todde" and "Zucca" have significant Moran's  $I$  positive autocorrelation. These results indicate that those surnames are clustered in the transhumance regions. In particular "Arangino" and "Onano" have high values of the index, clearly clustered in the transhumance path, which is consistent with Figure 3.

Table 3 – Moran's  $I$  evaluated with

	AGUS	ARANGINO	CABRAS	CASTANGIA	CASULA	FRAU	ONANO	PEDDIO		
	-0.0158	0.3620	0.0066	0.0578	0.0525	0.0344	0.2239	0.0480		
	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027		
St. Err.	0.0254	0.0166	0.0264	0.0264	0.0270	0.0271	0.0184	0.0113		
$p$ -value	0.6046	<b>0.0000</b>	0.7247	<b>0.0221</b>	<b>0.0411</b>	0.1709	<b>0.0000</b>	<b>0.0000</b>		

	POLLA	PORRU	SAU	SECCI	TODDE	TORE	URRU	VACCA	ZANDA	ZUCCA
	0.0212	0.0285	0.0263	0.0093	0.1040	0.0388	0.0266	0.0350	-0.0170	0.0418
	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027
St. Err.	0.0103	0.0254	0.0127	0.0268	0.0265	0.0242	0.0256	0.0263	0.0235	0.0264
$p$ -value	<b>0.0209</b>	0.2199	<b>0.0220</b>	0.6549	<b>0.0001</b>	0.0872	0.2539	0.1524	0.5424	<b>0.0930</b>

Table 4 shows the values of Moran's  $I$  with . It is noteworthy that each surname has a very low negative Moran's  $I$ , even the ones with a significant index. This indicates that all the examined surnames are spatially uncorrelated or very slightly negatively correlated in the regions outside of the transhumance routes.

Table 4 – Moran's  $I$  evaluated with

	AGUS	ARANGINO	CABRAS	CASTANGIA	CASULA	FRAU	ONANO	PEDDIO		
	-0.0023	-0.0067	-0.0027	-0.0034	-0.0027	-0.0030	-0.0052	-0.0033		
	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027		
St. Err.	0.0006	0.0003	0.0006	0.0006	0.0006	0.0006	0.0004	0.0002		
$p$ -value	0.5779	<b>0.0000</b>	0.9325	0.2084	0.9147	0.5288	<b>0.0000</b>	<b>0.0025</b>		

	POLLA	PORRU	SAU	SECCI	TODDE	TORE	URRU	VACCA	ZANDA	ZUCCA
	-0.0030	-0.0051	-0.0030	-0.0023	-0.0031	-0.0041	-0.0038	-0.0045	-0.0030	-0.0087
	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027	-0.0027

St. Err.	0.0002	0.0006	0.0002	0.0006	0.0006	0.0005	0.0006	0.0006	0.0005	0.0006
<i>p</i> -value	0.0610	<b>0.0000</b>	0.1182	0.5681	0.4130	<b>0.0069</b>	<b>0.0418</b>	<b>0.0013</b>	0.5591	<b>0.0000</b>

## Discussion

Data on classical genetic markers (Cavalli-Sforza et al., 1994; Vona, 1997), the HLA system (Contu et al., 1992; Lampis et al., 2000; Grimaldi et al., 2001), autosomal markers (Moral et al., 1994; Calò et al., 1998; Battaglia et al., 2003), mtDNA (Barbujani et al., 1995; Morelli et al., 2000; Malaspina et al., 2000; Richards et al., 2000; Fraumene et al., 2003; Falchi et al., 2006) and Y-chromosome polymorphisms (Semino et al., 2000; Scozzari et al., 2001; Ghiani and Vona, 2002; Francalacci et al., 2003; Quintana-Murci et al., 2003; Rootsi et al., 2004; Capelli et al., 2006) show that the Sardinian population (or its subpopulations) has a peculiar genetic structure that clearly differentiates it from the continental European and circum-Mediterranean populations, as well as from the other populations that currently constitute Italy. Moreover, Sardinia is characterized by specific autochthonous surnames (Lucchetti et al., 1996).

The genetic differences between Sardinians and the other Mediterranean and European populations are very likely due to the upper Paleolithic settlement of the island (Chikhi et al., 2002; Vona and Calò, 2006; D'Amore et al., 2010), its geographical isolation and, despite the many colonizations, the marginal position of the island with respect to mainstream economic development (Le Lannou, 1941; Sanna, 2006). Therefore, the Sardinian people, considered as a whole or as sub-populations, are very distinct among European populations. Nevertheless, the Sardinian population exhibits microgeographical heterogeneity (Vona and Calò, 2006). In fact, various authors have found marked variability within Sardinia (Piazza et al., 1985; Porcella and Vona, 1987; Ulizzi et al., 1988; Cavalli-Sforza et al., 1994; Vona, 1997; Cappello et al., 1996; Calò et al., 1998; Vona and Calò, 2006), leading to a mosaic genetic structure of the Sardinian population, often with different patterns according to the considered genetic markers. Regarding the debate about the heterogeneity of Sardinian populations, it is interesting that the Y-chromosome haplogroup that particularly characterizes the Sardinians with respect to other European and circum-Mediterranean populations (I2a2, M26) is distributed almost uniformly throughout the island (Zei et al., 2003). However, Zei et al. (2003) found that this genetic marker showed high heterogeneity among the three zones that they considered (northern, central-eastern and south-western) when the samples were formed on the basis of the origin of the surnames rather than the zone of residence. Hence, surname analysis is an excellent tool to validate the genetic relationships among populations.

Therefore, Sardinians are heterogeneous in relation to the specific history of the communities living in the island's different historical-geographical zones and also to the isolation of the mountain zones

from those of the plains. In particular, Barbagia di Belvì exhibits isonymic variability in relation to the geographical position of its villages.

Comparison of the single villages indicates that they are different from each other and the differentiation persists over time, so that the same village maintains its identity over the centuries: evidently, the geographical dimension and the local peculiarities are predominant. This is especially true for the two most isolated villages (Desulo and Tonara) but it also holds for the other villages, perhaps more open to migratory exchanges (Aritzo, Belvì, Gadoni), although they became differentiated over time while maintaining their diversified characterization.

Therefore, our results indicate that transhumance from the mountainous center to the southern lowlands and the recent internal migrations (Sanna, 2006) resulted in an appreciable flow of surnames toward the villages of the southern plains. Likewise, there has been a substantial gene flow between the mountain villages and those of the plains. In addition, the historical and geographical areas of the mountains have retained their specific surname identities, which strongly differentiate them from each other.

Nevertheless, we must consider the possible effect on the results of so-called virilocal marriages, i.e. the custom of celebrating the marriage in the village of the bride and then living in the village of the bridegroom (Sanna et al., 2004, 2006). Hence the children of parents who have followed the practice of virilocal marriage will bear the surname of the father, who remains to live in his native village. The maiden surnames of brides who migrate from the villages of Barbagia leave no traces in their native villages. It is also true, however, that marriages are not only exchanges of surnames and therefore of genes, but also occasions of cultural exchanges.

The genetic structure simulated by surnames indicates continuity between the populations of the villages of the southern lowlands and the populations of the mountain villages of the central area of Sardinia. Moreover, the latter differ according to the historical-geographical zones to which they belong; in view of their positions on the nmMDS plot, the villages of Barbagia di Belvì are differentiated from each other.

The results suggest that there has been appreciable admixture between the Sardinian populations of the central mountains and those of the southern lowlands.

In view of the similarities found between the surnames of the populations of the central area of the island and those of the lowlands, future research should address the genetic relationships among these Sardinian individuals. For this purpose, studies should include individuals whose origin is known and for whom a genealogical reconstruction is possible.

### **Compliance with Ethical Standards:**

The authors declare that they have no conflict of interest.

## **Acknowledgements**

The authors thank the ecclesiastical authorities and the staff of the Archivio della Curia Arcivescovile di Oristano for their assistance with the data collection.

The authors would like to thank Christie P.W. for the translation.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## **References**

Barbujani G, Bertorelle G, Capitani G, Scozzari R. 1995. Geographical structuring in the mtDNA of Italians. *Proc. Natl. Acad. Sci. USA* 92:9171-9175.

Battaglia C, Ruscitto D, Destro-Bisol G, Vacca L, Calò C, Vona G. 2003. Frequencies at CD4, FES, and F13A1 microsatellite loci in central-southern Sardinia (Italy). *J. Forensic Sci.* 48:442-445.

Boattini A, Lisa A, Fiorani O, Zei G, Pettener D, Manni F. 2012. General method to unravel ancient population structures through surnames, final validation on italian data. *Human Biology.* 84: Iss. 3, Article 2. <http://digitalcommons.wayne.edu/humbiol/vol84/iss3/2>.

Bowden GR, Balaesque P, King TE, Hansen Z, Lee AC, Pergl-Wilson G, Hurley E, Roberts SJ, Waite P, Jesch J, Jones AL, Thomas MG, Harding SE, Jobling MA. 2008. Excavating past population structures by surname-based sampling: the genetic legacy of the Vikings in Northwest England. *Mol. Biol. Evol.* 25:301-309.

Calò CM, Autuori L, Di Gaetano C, Latini V, Mameli GE, Memmi M, Varesi L, Vona G. 1998. The polymorphism of the APOB3' VNTR in the populations of the three largest islands of the western Mediterranean. *Anthrop. Anz.* 56:1-12.

Capelli C, Redhead N, Romano V, Cali F, Lefranc G, Delague V, Megarbane A, Felice AE, Pascali VL, Neophytou PI, Poulli Z, Novelletto A, Malaspina P, Terrenato L, Berebbi A, Fellous M, Thomas MG, Goldstein DB. 2006. Population structure in the Mediterranean basin: a Y chromosome perspective. *Ann. Hum. Genet.* 70:207-225.

- Cappello N, Rendine S, Griffo R, Mameli GE, Succa V, Vona G. 1996. Genetic analysis of Sardinia: I. Data on 12 polymorphisms in 21 linguistic domains. *Ann. Hum. Genet.* 60:125-141.
- Caravello G, Tasso M. 1999. An analysis of the spatial distribution of surnames in the Lecco area (Lombardy, Italy). *Am. J. Hum. Biol.* 11:305-315.
- Caravello GU, Tasso M. 2007. Surnames as alleles: spatial distribution of surnames in a province of Italian Alps. *J. Biosoc. Sci.* 39:409-419.
- Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. *The history and geography of human genes*. Princeton, NJ: Princeton University Press.
- Chen KH, Cavalli-Sforza LL. 1983. Surnames in Taiwan: interpretations based on geography and history. *Hum. Biol.* 55:367-374.
- Cheshire J. 2014. Analysing surnames as geographic data. *J. Anthropol Sci.* 92:99-117.
- Chikhi L, Nichols RA, Barbujani G, Beaumont MA. 2002. Y genetic data support the Neolithic demic diffusion model. *Proc. Natl. Acad. Sci. USA* 99:11008-11013.
- Cliff AD, Ord JK. 1981. *Spatial processes: Models and applications*. London: Pion.
- Colantonio SE, Lasker GW, Kaplan BA, Fuster V. 2003. Use of surname models in human population biology: a review of recent developments. *Hum. Biol.* 75:785-807.
- Contu L, Arras M, Carcassi C, La Nasa G, Mulargia M. 1992. HLA structure of the Sardinian population: a haplotype study of 551 families. *Tissue Antigens* 40:165-174.
- Crow JF. 1983. Surnames as biological markers. *Hum. Biol.* 55:383-397.
- D'Amore G, Di Marco S, Floris G, Pacciani E, Sanna E. 2010. Craniofacial morphometric variation and the biological history of the peopling of Sardinia. *Homo - Journal of Comparative Human Biology* 61: 385-412.



Darlu, P, Bloothoft, G, Boattini A et al. 2012. The family name as socio-cultural feature and genetic metaphor: From concepts to methods. *Hum. Biol.* 84:169-214.

Falchi A, Giovannoni L, Calò CM, Piras 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.* 5: 9-14.

Francalacci P, Morelli L, Underhill PA, Lillie AS, Passarino G, Useli A, Madeddu R, Paoli G, Tofanelli S, Calò 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. Anthropol.* 121:270-279.

Fraumene C, Petretto E, Angius A, Pirastu M. 2003. Striking differentiation of sub-populations within a genetically homogeneous isolate (Ogliastra) in Sardinia as revealed by mtDNA analysis. *Hum. Genet.* 114:1-10

Ghiani ME, Vona G. 2002. Y-chromosome-specific microsatellite variation in a population sample from Sardinia (Italy). *Coll. Antropol.* 26:387-401.

Ghiani Moi P. 1964. *Sardegna d'oggi*. Roma, Italy: La Poliedrica.

Graf OM, Zlojutro M, Rubicz R et al. 2010. Surname distributions and their association with Y-chromosome markers in the Aleutian Islands. *Hum. Biol.* 82:745-757.

Grimaldi MC, Crouau-Roy B, Amoros JP, Cambon-Thomsen A, Carcassi C, Orru S, Viader C, Contu L. 2001. West Mediterranean islands (Corsica, Balearic islands, Sardinia) and the Basque population: contribution of HLA class I molecular markers to their evolutionary history. *Tissue Antigens.* 58:281-292.

Hill EW, Jobling MA, Bradley DG. 2000. Y-chromosome variation and Irish origins. *Nature* 404:351-352.

Jobling MA. 2001. In the name of the father: Surnames and genetics. *Trends Genet.* 17:353-357.

Jobling MA, Tyler-Smith C. 2003. The human Y chromosome: An evolutionary marker comes of age. *Nat. Rev. Genet.* 4:598-612.

King TE, Ballereau SJ, Schurer KE, Jobling MA. 2006. Genetic signatures of co-ancestry within surnames. *Curr. Biol.* 16:384-388.

King TE, Jobling MA. 2009a. Founders, drift, and infidelity: The relationship between Y chromosome diversity and patrilineal surnames. *Mol. Biol. Evol.* 26:1093-1102.

King TE, and Jobling MA. 2009b. What's in a name? Y chromosomes, surnames and the genetic genealogy revolution. *Trends Genet.* 25:351-360.

Lampis R, Morelli L, De Virgiliis S, Congia M, Cucca F. 2000. The distribution of HLA haplotypes reveals that the Sardinian population is genetically differentiated from the other Caucasian populations. *Tissue Antigens* 56:515-521.

Larmuseau MHD, Vanoverbeke J, Gielis G, Vanderheyden N, Larmuseau HFM, Decorte R. 2012. In the name of the migrant father - Analysis of surname origins identifies genetic admixture events undetectable from genealogical records. *Heredity.* 109:90-95.

Lasker GW. 1985. *Surnames and genetic structure*. Cambridge, UK: Cambridge University Press.

Le Lannou M. 1941. *Patres et paysans de la Sardaigne*. Tours, France: Arrault.

Lucchetti E, Pettener D, De Iasio S, Martuzzi Veronesi F, Conterio F. 1987. Analisi e confronti tra la struttura biologica e la struttura per cognomi nella popolazione della provincia di Sassari. *Antropologia Contemporanea* 10 (Suppl. 1-2):237-245.

Lucchetti E, Siri E, Soliani L. 1996. Internal migrations and relationships between Italian regional populations: An analysis based on surname distributions. *Rivista di Antropologia-JASs* 74:171-183.

Lucchetti E, Tasso M, Pizzetti P, De Iasio S, Caravello GU. 2008. Similarities in the surnames of island and continental populations of the north-western Mediterranean area. *J. Biosoc. Sci.* 40:359-

377, 2007 Cambridge University Press doi:10.1017/S0021932007002465 First publisher online 24 Oct 2007.

Malaspina P, Cruciani F, Santolamazza P, Torroni A, Pangrazio A, Akar N, Bakalli V, Brdicka R, Jaruzelska J, Kozlov A, Malyarchuk B, Mehdi SQ, Michalodimitrakis E, Varesi L, Memmi MM, Vona G, Villierms R, Parik J, Romano V, Stefan M, Stenico M, Terrenato L, Novelletto A, Scozzari R. 2000. Patterns of male-specific inter-population divergence in Europe, West Asia and North Africa. *Ann. Hum. Genet.* 64:395-412.

Manni F, Toupance B, Sabbagh A, Heyer E. 2005. New method for surname studies of ancient patrilineal population structures, and possible application to improvement of Y-chromosome sampling. *Am. J. Phys. Anthropol.* 126:214-228.

Martínez-Gonzalez LJ, Martínez-Espín E, Alvarez JC et al. 2012. Surname and Y chromosome in Southern Europe: A case study with Colom/Colombo. *Eur. J. Hum. Genet.* 20:211-216.

McEvoy B, Bradley DG. 2006. Y-chromosomes and the extent of patrilineal ancestry in Irish surnames. *Hum. Genet.* 119:212-219.

Meloni B. 1982. La transumanza. In *Contadini e pastori nella Sardegna tradizionale*. Cinisello Balsamo, Italy: CRS-Pizzi, 157-159.

Moral P, Marogna G, Salis M, Succa V, Vona G. 1994. Genetic data on Alghero population (Sardinia): contrast between biological and cultural evidence. *Am. J. Phys. Anthropol.* 93:441-453.

Moran PAP. 1950. Notes on continuous stochastic phenomena. *Biometrika.* 37:17-23.

Morelli L, Grosso MG, Vona G, Varesi L, Torroni A, Francalacci P. 2000. Frequency distribution of mitochondrial DNA haplogroups in Corsica and Sardinia. *Hum. Biol.* 72:585-595.

Orrù A, Sanna E, Lucchetti E. 2008. Transumanza e distribuzione dei cognomi in Sardegna: mobilità dei pastori ed isolamento delle popolazioni. *International Journal of Anthropology s.n.*: 368-375.

- Ortu GG. 1988. La transumanza nella storia della Sardegna. *Quaderni bolotanesi* 14:159-174.
- Piazza A, Mayr WR, Contu L, Amoroso A, Borelli I, Cartoni ES, Marcello C, Moroni A, Olivetti E, Richiardi P, Ceppellini R. 1985. Genetic and population structure of four Sardinian villages. *Ann. Hum. Genet.* 49:47-63.
- Porcella P, Vona G. 1987. Distribution of the AB0 system and Rh factor in Sardinia. *Anthrop. Anz.* 45:309-321.
- Quintana-Murci L, Veitia R, Fellous M, Semino O, Poloni ES. 2003. Genetic structure of Mediterranean populations revealed by Y-chromosome haplotype analysis. *Am. J. Phys. Anthropol.* 121:157-171.
- 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, Golge 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, Norby 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.
- Ripley BD. 1981. *Spatial statistics*. New York: John Wiley & Sons.
- Risso D, Taglioli L, De Iasio S, Guerresi P, Alfani G, Nelli S, Rossi P, Paoli G, Tofanelli S. 2015. Estimating sampling selection bias in human genetics: a phenomenological approach. *Plos One*. doi: 10.1371/journal.pone.0140146
- Román-Busto J, Tasso M, Caravello G, Fuster V, Zuluaga P. 2013. Genetics of population exchange along the historical Portuguese–Spanish border. *J. Biosoc. Sci.* 45:7993. doi:10.1017/S0021932012000363
- Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, Kutuev I, Barac L, Pericic M, Balanovsky O, Pshenicov A, Dion D, Grobei M, Zhivotovsky LA, Battaglia V, Achilli A, Al-Zahery N, Parik J, King R, Cinnioglu C, Khusnutdinova E, Rudan P, Balanovska E, Scheffrahn W, Simonescu M, Brehm A, Goncalves R, Rosa A, Moisan JP, Chaventre A, Ferak V, Füredi S,

- Oefner PJ, Shen P, Beckman L, Mikerezi I, Terzić R, Primorac D, Cambon-Thomsen A, Krumina A, Torroni A, Underhill PA, Santachiara-Benerecetti AS, VILLEMS R, Semino O. 2004. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am. J. Hum. Genet.* 75:128-137.
- Sanna E. 2006. *Il popolamento della Sardegna e le origini dei Sardi*. Cagliari, Italy: CUEC.
- Sanna E, Danubio ME. 2008. Seasonality of marriages in Sardinian pastoral and agricultural communities in the nineteenth century. *J. Biosoc. Sci.* 40:577-586.
- Sanna E, Iovine MC, Floris G. 2004. Evolution of marital structure in 20 Sardinian villages from 1800 to 1974. *Anthropol. Anz.* 62:169-184.
- Sanna E, Iovine MC, Melis M, Floris G. 2006. Lasker's coefficient of isonymy between and among 16 Sardinian villages in the periods 1825-1849, 1875-1899, 1925-1949. *Am. J. Hum. Biol.* 18:621-629.
- Sanna E, Melis M, Floris G. 2001. Coefficient of relationship by isonymy between 14 Sardinian villages in the periods 1800-1824 and 1950-1974. *Z. Morph. Anthropol.* 83:117-127.
- 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.* 62:871-844.
- Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedectis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiæ M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA. 2000. The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y-chromosome perspective. *Science* 290:1155-1159.
- Sokal RR, Oden NL. 1978a. Spatial autocorrelation in biology. 1. Methodology. *Biol. J. Linn. Soc.* 10: 199-228.
- Sokal RR, Oden NL. 1978b. Spatial autocorrelation in biology. 2. Some biological implications,

and four applications of evolutionary and ecological interest. *Biol. J. Linn. Soc.* 10: 229-249.

Sykes B, Irven C. 2000. Surnames and the Y chromosome. *Am. J. Hum. Genet.* 66:1417-1419.

Tasso M, Caravello G. 2010. Cognomi come alleli: distribuzioni spaziali dei cognomi nella provincia di Rovigo (Italia). *Antropo.* 21:9-18. [www.didac.ehu.es/antropo](http://www.didac.ehu.es/antropo)

Ulizzi L, Scozzari R, Stefanini R, Di Cori L, Salsano F, Santolamazza C, Santachiara-Benerecetti SA, Beretta M, Bernini L, Van Loghem E, Nijenhuis LE, Modiano G, Terrenato L. 1988. Genetic heterogeneity in Sardinia, 15 polymorphisms examined in 11 isolates. *Gene Geog.* 2:141-157.

Vona G. 1997. The peopling of Sardinia (Italy): history and effects. *Int. J. Anthropol.* 12:71-87.

Vona G, Calò CM. 2006. History of Sardinian population (Italy, Western Mediterranean) as inferred from genetic analysis. In *Human genetic isolates* (C. Calò and G. Vona Eds.) Trivandrum-Kerala, India: Research Signpost, 1-28.

Wijsman E, Zei G, Moroni A, Cavalli-Sforza LL. 1984. Surnames in Sardinia. II. Computation of migration matrices from surname distributions in different periods. *Ann. Hum. Genet.* 48:65-78.

Winney B, Boumertit A, Day T, Davison D, Echeta C, Evseeva I, Hutnik K, Leslie S, Nicodemus K, Royrvik EC, Tonks S, Yang X, Cheshire J, Longley P, Mateos P, Groom A, Relton C, Bishop DT, Black K, Northwood E, Parkinson L, Frayling TM, Steele A, Sampson JR, King T, Dixon R, Middleton D, Jennings B, Bowden R, Donnelly P, Bodmer W. 2012. People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population. *Eur. J. Hum. Genet.* 20: 203-210.

Zei G, Guglielmino Matessi R, Siri E, Moroni A, Cavalli-Sforza LL. 1983a. Surnames in Sardinia. I. Fit of frequency distribution for neutral alleles and genetic population structure. *Ann. Hum. Genet.* 47:329-352.

Zei G, Guglielmino CR, Siri E, Moroni A, Cavalli-Sforza LL. 1983b. Surnames as neutral alleles: observations in Sardinia. *Hum. Biol.* 2:357-408.

Zei G, Lisa A, Fiorani O, Magri C, Quintana-Murci L, Semino O, Santachiara-Benerecetti AS. 2003. From surnames to the history of Y chromosomes: the Sardinian population as a paradigm. *Eur. J. Hum. Genet.* 11:802-807.

Zei G, Piazza A, Moroni A, Cavalli-Sforza LL. 1986. Surnames in Sardinia. III. The spatial distribution of surnames for testing neutrality of genes. *Ann. Hum. Genet.* 50:169-180.

## **Figure Legends**

*Figure 1 – The 34 historical-geographical zones of Sardinia according to Ghiani Moi (1964). Highlighted areas are starting, transitional and arrival points of transhumance*

*Figure 2 - Map of the long-distance transhumance routes*

*Figure 3 – Current distribution of 18 surnames selected in the historical-geographical zones involved in transhumance; darker regions correspond to higher frequencies*

*Figure 4 – Scatter plot of non-metric Multidimensional Scaling of the five villages of Barbagia di Belvì in time*