CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

The structure and migration patterns of the population of Uruguay through isonymy

A. Carrieri[1†] M. Sans[2*†] J. E. Dipierri[2], E. Alfaro[2], E. Mamolini[1], M. Sandri[1], A. Rodríguez-Larralde[3], C. Scapoli[1‡] and I. Barrai[1‡]

[1]Department of Life Sciences and Biotechnology, University of Ferrara, Italy, [2]Departamento de Antropologia Biológica, Facultad de Humanidades y Ciencias de la Educación, Universidad de la República, Uruguay and [3]Instituto de Ecorregiones Andinas (UNJu – CONICET), Instituto de Biología de la Altura, Universidad de Jujuy, S. Salvador de Jujuy, Argentina
*Corresponding author. Email: mbsans@gmail.com

Abstract

Surname distribution can be a useful tool for studying the genetic structure of a human population. In South America, the Uruguay population has traditionally been considered to be of European ancestry, despite its trihybrid origin, as proved through genetics. The aim of this study was to investigate the structure of the Uruguayan population, resulting from population movements and surname drift in the country. The distribution of the surnames of 2,501,774 people on the electoral register was studied in the nineteen departments of Uruguay. Multivariate approaches were used to estimate isonymic parameters. Isolation by Distance was measured by correlating isonymic and geographic distances. In the study sample, the most frequent surnames were consistently Spanish, reflecting the fact that the first immigration waves occurred before Uruguayan independence. Only a few surnames of Native origin were recorded. The effective surname number (α) for the entire country was 302, and the average for departments was 235.8 ± 19. Inbreeding estimates were lower in the south-west of the country and in the densely populated Montevideo area. Isonymic distances between departments were significantly correlated with linear geographic distance ($p < 0.001$) indicating continuously increasing surname distances up to 400 km. Surnames form clusters related to geographic regions affected by different historical processes. The isonymic structure of Uruguay shows a radiation towards the east and north, with short-range migration playing a major role, while the contribution of drift, considering the small variance of α, appears to be minor.

Keywords: Population Structure; Geographic distribution; Migration

Introduction

Study of the genetic structure of the populations of South American nations has been relatively recent, and most studies have mainly examined the frequencies of traditional DNA markers. This is the case for Uruguay, whose population has been analysed using classical genetic markers as well as nuclear and mitochondrial DNA markers (Sans *et al.*, 1994, 1997, 2002, 2006, 2011; Bravi *et al.*, 1997; Bertoni *et al.*, 2003; Bonilla *et al.*, 2004; Gascue *et al.*, 2005). Genetic data have shown that the continental contribution to the Uruguayan population is around 10–14% Native American, 7–9% sub-Saharan African, and the rest European or Mediterranean (Hidalgo *et al.*, 2005; Bonilla *et al.*, 2015). However, Uruguay lacks data on its genetic structure using surnames, with only one

---

[†]Contributed equally.
[‡]Contributed equally.

previous study including isonymy related to a small population derived from a former Native reduction, Villa Soriano (Barreto, 2011).

The use of surnames has been controversial as it has been shown that they have sometimes been changed or taken from other, non-related individuals. For example, in Uruguay the present population lacks Native-origin names, with a very few exceptions. For example, the descendants of the Charrúa Chief (*cacique*) Sepé initially used the surname Sepé but changed it recently to García (Acosta y Lara, 1981). González-Rissotto and Rodríguez-Varese (1989) stated that Indians from the Jesuit Missions that came to Uruguay changed their habits and their names in an attempt to hide their Native ancestry. Also, African slaves and their descendants took different surnames, such as those of their owners or others, as shown by Rosal (2009) in a study of the people of African and African descent in Buenos Aires, Argentina. However, very few studies have been performed on this in Uruguay, and in general terms, most have focused on particular issues – for example, the Africans or African descendants that took their name of the national hero Artigas (Yarza Rovira, 2009). Moreover, some surnames from Europe, the Mediterranean region or elsewhere have been changed or spelt in different ways (Flores, 2010). Finally, paternal surnames are frequently lost because of illegitimate unions (see for example, Barreto, 2011), adoption and the inversion of the order of maternal and paternal surnames as a consequence of the influence of the Portuguese-Brazilian system of surnames, especially near the Brazilian border. Moreover, since 2004, in Uruguay it has been possible to choose the paternal or maternal surname as the first surname, and children with only one last name receive either the two maternal (or paternal) surnames or are assigned a 'common name' when the child has only one parent, and this parent only has one surname (2004 Uruguay Law 17.823). However, all individuals considered here were born before 1993; that is, before the application of the Law.

The aim of this work was to make a comprehensive study of the present isonymic structure of Uruguay, resulting from population movements and surname drift in the country. For that purpose, the Uruguayan population was investigated to detect its structure through the study of isonymy (Crow & Mange, 1965; Yasuda & Morton, 1967) at the main administrative level of the nation, namely its nineteen departments. Isonymic distances between departments were analysed in relation to geography. Moreover, by studying the geographic heterogeneity of surnames based on the analysis of different isonymy parameters, it was possible to detect the signs of the direction of migrations.

Methods

Data

The República Oriental del Uruguay is one of the smallest countries in South America, surrounded by Argentina and Brazil, and with an Atlantic Ocean coastline. At its maximum, it is about 700 km long and 500 km wide, with an area of slightly over 176,000 km^2. According to the 2011 National Census (INE, 2011) it is inhabited by approximately 3.3 million persons. The country is divided into nineteen departments.

In Uruguay, surnames originated and have been established generally in the same way as in most South American countries. However, Uruguay and Argentina make an exception to the Spanish biparental surname system, with the maternal surname not generally being mentioned, although it appears on official documents.

In 2012, the authors JED and EA obtained surname data from the Corte Electoral of Uruguay suitable for describing the isonymy structure of the country. The data comprised the electoral list for the 2011 presidential general election – a total of 2.5 million individuals over the age of 18 years, distributed in the nineteen departments of the country.

The 2011 National Census found that 7.8% of those interviewed reported African descent, while the percentage with Amerindian descent was 4.9%; 90.8% stated European descent and only 0.5% recognized an Asian descent, while 3.6% gave no response (INE, 2012). These percentages take into account multiple descents. Since ethnic origin is only weakly related to surnames, in the

Figure 1. Map of Uruguay showing the nineteen departments.

present analysis ethnicity was ignored and departments (sub-national level) were taken as the statistical units. The geography of Uruguay's departments is well defined, and all individuals in the sample were classified according to the department in which they were enrolled to vote. This location was usually the place of residence, and did not always coincide with their place of birth in Uruguay. For the analysis, 2,501,774 paternal surnames were available.

The study area covered the entire nation. The nineteen departments differ in their position, area and population (Fig. 1), and most border the three main rivers of Río de la Plata, Uruguay and Río Negro. Río Negro divides the country north and south, and this affects the country's population distribution: while the south was mainly populated from the harbour of Montevideo, the north (at least the north-east) was mainly populated from Brazil (Pi Hugarte & Vidart, 1969).

Analysis

The surname distribution for the whole Uruguay sample was studied by fitting a regression line to the log-log transformation of the number of surnames (S) represented k times (Fox & Lasker, 1983) in populations I and in population J. The main statistics derived from surname distributions were isonymy and random isonymy. Isonymy within a group J is given by:

$$I_{jj} = \sum_k p_{kj}^2$$

where p_{kj} is the relative frequency of surname k in group J, and the sums comprise all surnames. Random isonymy between groups I and J is estimated as:

$$I_{ij} = \sum_k p_{ki} p_{kj}$$

where p_{ki} and p_{kj} are the relative frequencies of surname k in groups I and J, respectively, and the sums comprise all surnames. In addition, Fisher's $\alpha = 1/I_{jj}$ and Karlin-McGregor's $\nu = \alpha/(N-\alpha)$ were calculated, where N is the surname's number.

To detect isolation by distance, the linear correlation of surname distances (Lasker's, Euclidean and Nei's distances) between localities I and J with their geographic distances was calculated.

Lasker's distance (Rodríguez-Larralde *et al.*, 1998) is defined as:

$$L = -\log(I_{ij})$$

Euclidean distance (Cavalli-Sforza & Edwards, 1967) is defined as:

$$E = \sqrt{1 - \sum_k \sqrt{p_{ki}p_{kj}}}$$

where the summation is over all surnames. Nei's distance (Nei, 1973) is given by:

$$N_d = -\log\left(\frac{I_{ij}}{\sqrt{(I_{ii}I_{jj})}}\right)$$

The definitions of these statistics and their meaning in the study of the microevolution in human groups have been described previously by Barrai *et al.* (1996, 2000), Rodríguez-Larralde *et al.* (2011) and Dipierri *et al.* (2005); but see also Herrera Paz *et al.* (2014) for a description of Fisher's α corresponding to the effective surname number; for an exhaustive review see Relethford (1988).

The random kinship $\Phi_{ij}(x)$ between any two localities I and J at distance x is given by:

$$\Phi_{ij}(x) = K \exp(-Bx)$$

where K is the average kinship at geographic distance $x = 0$, say average F_{ST} (measured as $I_{jj}/4$) and B is a function of the average mutation rate and the variance of x (Malécot, 1955; Kimura, 1960). The value of $\Phi_{ij}(x)$ is always positive and is expected to decrease exponentially to 0 with increasing distance. Random kinship between groups I and J was estimated as:

$$\Phi_{ij}(x) = I_{ij}(x)/4$$

where average F_{ST} is the average kinship at distance $x = 0$ (Barrai *et al.*, 2012). Linear geographic distances were obtained using ArcGis® (ESRI) software.

The significance of correlations was assessed with Mantel's test using 1000 permutations (Mantel, 1967; Smouse *et al.*, 1986). For a graphic representation of the relationship between surname distributions at different locations, these were mapped on the first and second dimensions of the Multidimensional Scaling (MDS) of Nei's distance matrix. To this purpose, the R® software package was used (R Development Core Team, 2012). To detect the direction of surname diffusion, following Menozzi *et al.* (1978), the first three components from the Principal Component Analysis (PCA) of the same matrix were also projected individually on the Uruguay map, again with the ArcGis software package.

To clarify and complement the clusters obtained with the MDS, dendrograms of departments were built, obtained from the matrices of isonymic distances between these administrative sections. The dendrograms were only an additional option for the clustering, and do not imply that their results were generated by subsequent splits of pre-existing clusters.

Results and Discussion

Most frequent surnames

The distribution of study sample surnames by department is given in Table 1, along with the main parameters derived from isonymy theory. The distribution of the logarithm of the number of surnames over the logarithm of the number of times they occur (Fox & Lasker 1983) was fairly linear. Some convexity in the distribution was apparent, and a possible explanation for this is the immigration of independent family groups with uncommon surnames, which has been observed in Uruguay.

Table 1. Distribution of study sample surnames and main parameters derived from isonymy theory by Uruguay department

Department	N	S	α	ν	I	F_{ST}	Long	Lat
Montevideo	10,30,800	53,858	346.2	0.00034	0.00288	0.000720	−56.227	−34.826
Canelones	361,446	22,909	250.6	0.00069	0.00398	0.000994	−55.959	−34.519
Maldonado	121,442	10,814	213.7	0.00176	0.00467	0.001167	−54.862	−34.547
Rocha	58,386	3964	121.8	0.00208	0.00816	0.002040	−54.023	−33.943
Treinta y Tres	39,743	2702	181.5	0.00455	0.00550	0.001376	−54.305	−33.001
Cerro Largo	67,956	3697	154.8	0.00227	0.00643	0.001608	−54.375	−32.402
Rivera	82,983	5113	161.1	0.00194	0.00618	0.001545	−55.034	−31.647
Artigas	58,761	3653	172.0	0.00292	0.00580	0.001449	−56.951	−30.547
Salto	98,384	5840	237.6	0.00241	0.00420	0.001050	−57.049	−31.237
Paysandú	90,667	6588	393.5	0.00432	0.00255	0.000637	−57.379	−32.033
Río Negro	42,702	4255	329.4	0.00766	0.00305	0.000762	−57.466	−32.717
Soriano	69,446	4950	376.5	0.00539	0.00266	0.000666	−57.764	−33.478
Colonia	100,687	6797	361.5	0.00358	0.00277	0.000692	−57.679	−34.112
San José	78,072	6232	198.2	0.00253	0.00503	0.001258	−56.746	−34.274
Flores	21,715	2310	208.9	0.00953	0.00481	0.001203	−56.936	−33.510
Florida	55,357	4165	188.6	0.00340	0.00529	0.001323	−55.902	−33.703
Durazno	46,297	3638	222.0	0.00477	0.00451	0.001126	−56.108	−32.935
Lavalleja	50,396	3528	194.9	0.00385	0.00512	0.001281	−55.009	−33.868
Tacuarembó	26,526	2503	168.6	0.00631	0.00590	0.001483	−55.807	−32.040

Long = longitude; Lat = latitude; N = number of surnames; S = number of different surnames; α = Fisher's; ν = Karlin-McGregor's ν; I = isonymy; and F_{ST} = random component of inbreeding.

The average number of instances (persons) with the same paternal surname, that is (Sample Size)/(Number of Surnames) ratio (SS/S), the so called 'type-token' ratio of glottologists (Adamic & Huberman, 2002), was 36 (see Table 2) – the lowest ratio that was observed in previous studies covering six countries and more than 77 million surnames from Spanish-speaking populations. However, this value is similar to those from Spain, and comparable to those of other European countries.

The low type-token ratio, as well as the high incidence of uncommon surnames, in Uruguay is coherent with the several migration waves that make up the Uruguayan population, together with the slave trade and Native background. Unlike in Latin American countries more related to the time of first conquest, Uruguay was populated more recently by Europeans, with the first settlements being founded at the end of the 17th century by Iberians (Spanish and Portuguese), who brought some African or African-descendant slaves. In 1662, Spain founded Villa Soriano, initially a Native Indian reduction, and in 1680, the Portuguese founded Colonia del Sacramento. Later, in 1724–27, the Spanish (mostly from the Canary Islands) founded the capital city of Montevideo. After the installation of the Republic (1830), in the 19th and 20th centuries waves of immigrants came from different European and Near Eastern countries. In order these were French (mostly Basques), Spanish, Italian, Slaves, Armenians, Jews and Syrian-Lebanese (Pi Hugarte & Vidart, 1969). These migrations lasted until the 1960s decade and have resumed recently.

The 50 most frequent surnames were studied in some detail. The most frequent surnames were Rodríguez with 75,039 occurrences, González with 50,573, Martínez with 37,637, Fernández with

Table 2. Comparison of isonymy parameters in nine European countries, five South American countries, USA and Texas and in Yakutia, Siberia

Country	Sample size (SS, in millions)	S	Average α	Isolation by Distance	Type-token (SS/S) ratio (%)	Source
Europe						
Austria	1	140,766	854	0.59	7.1	Scapoli et al. (2007)
Albania	3	37,184	123	0.71	82	Mikerezi et al. (2013)
Belgium	1.1	137,442	997	0.74	8	Scapoli et al. (2007)
France	6	495,104	1615	0.69	12.1	Scapoli et al. (2007)
Germany	5.2	462,526	1596	0.51	11.2	Scapoli et al. (2007)
Netherland	2.4	126,485	787	0.46	19	Scapoli et al. (2007)
Italy	5.1	215,623	1236	0.61	23.7	Scapoli et al. (2007)
Spain	3.6					
Paternal		94,886	134	0.21	38	Rodríguez-Larralde et al. (2003)
Maternal		110,034	144	0.26	33	Rodríguez-Larralde et al. (2003)
Switzerland	1.7	166,116	891	0.72	10.2	Scapoli et al. (2007)
Asia						
Yakutia	0.5	44,625	107	0.69	11.1	Tarskaia et al. (2009)
North America						
Texas	3.6	235,740	734	0.42	15.3	Rodríguez-Larralde et al. (2007)
USA	18	899,585	1366	0.24	20	Barrai et al. (2001)
South America						
Argentina	22.6	414,441	422	0.47	54.5	Dipierri et al. (2005)
Bolivia	23.2	174,922	122	0.5	144.6	Rodríguez-Larralde et al. (2011)
Chile	16.4	72,667	209	0.63	224.0	Barrai et al. (2012)
Paraguay	4.8	39,047	108	0.42	122.9	Dipierri et al. (2011)
Uruguay	2.5	70,395	301	0.56	35.5	Present study
Venezuela	3.9	68,665	122	0.78	56.8	Rodríguez-Larralde et al. (2000)
Central America						
Honduras	4.3	14,665	113	0.43	296.5	Herrera Paz et al. (2014)

31,769, Pérez with 28,511 and García with 27,331. After these, were Silva (27,070), López (21,959), Pereira (21,337) and in the tenth place Sosa (20,257). Overall, the first ten surnames comprised 341,483 individuals, or 13.6% of the total elector sample.

The most frequent surnames were consistently Spanish. The main difference is represented by Portuguese surnames, frequent in the North East of the country near the Brazilian border. So, surnames like Ferreira (15,624), Pereira (21,137) and Silva (27,070) were highly represented. These surnames, however, are frequent also in Spain. Iberian surnames reflect the conquerors and colonizers that have entered the country since the 17th and 18th centuries. Moreover, care should be taken as Iberian surnames also applied to Natives and Africans, who took their names from conquerors, owners or others during the Spanish/Portuguese domination, with some

exceptions such as Brown and Rondeau, taken from Admiral George Brown and General Jose Rondeau – military leaders who participated in the Independence war against Spain (Lucas, 1992).

Italian surnames were less represented. The most frequent was Ferrari (2206), followed by Rossi (1746) and Bianchi (1176), the three ubiquitous in Italy, then Parodi (1390), typical of the Genoa area, and Sanguinetti (1249) from Eastern Ligury. Two Basque surnames – Duarte (43rd, with 6300 representations) and Larrosa (49th, with 5632) – were among the most frequent 50 surnames. Italian migration was mainly related to the first waves after Independence, following the French (mostly Basque) and Spanish (Pi Hugarte & Vidart, 1969).

These 50 most common surnames did not include anyone related to the latest immigration waves that took place at the end of the 19th and beginning of the first half of the 20th centuries, from different regions, including Central–Eastern Europe and the Near East. However, it is possible to identify some that denote those different origins, such as Muller/Moller (German, 627 representations), Armand Ugon (Waldensian, 317), Miller (English, 310), Schmidt (German, 216), Cohen (Jewish, 154) and Garabedián (Armenian, 59). Moreover, as spellings usually have several forms, it is possible to consider terminations: 'ián', usually Armenian, had 5800 appearances; 'sky/ski' (Slavic, mostly Polish), 2958; 'berg' (German/German Jewish), 913; and 'skas' (Lithuanian), 276. At least two Native Indian surnames were also detected, originally from the Jesuitic Missions: Yasuiré, with 111 representations, and Barité, with 64. No African-origin surnames were identified, despite special attention being given to the possible use of place names, places of birth and ethnic groups, as described by Cuba Manrique (2002) in her study about the surnames of African slaves in Perú. A detailed analysis of the origin of surnames and their mutations will be done in a future article.

Isonymy by department

Fisher's α is an estimate of the 'effective surname number' – the number of surnames which, having the same frequency, would result in the same isonymy as that actually observed. Uruguay department α was found to be correlated ($r = 0.736$, $p < 0.001$) with longitude, but not with latitude ($r = 0.24$, non-significant). So, high α and low F_{ST} were clustered in the south-western departments, while lower α and higher inbreeding were found in the north-eastern departments. This could be explained by immigration and regional history: the north-east region was populated mainly from Brazil, while the south-west received European (as well as Iberians, Italian and Basques: Germans, Waldesians, British, Russians and Armenians) and Near East immigration waves (Turkish, Syrian, Lebanese), some of them funding colonies in the west and south-west of the country (Pi Hugarte & Vidart, 1969; Vidart & Pi Hugarte, 1969; Barrán & Nahum, 1971).

The effective surname number, α, for Uruguay was estimated to be 302 for the country as a whole. The average for the nineteen departments was 235.8 ± 19.1. Differences between administrative levels, and the country as a whole, were observed when different subdivisions of the same area and population were considered. This is the Prefecture Effect, called F_{ST} by Nei and Imaizumi (1966) in Japan, and so named by Scapoli *et al.* (2007). Nei and Imaizumi observed that, for the same area and population, small subdivisions have larger F_{ST} values, and larger subdivisions have smaller F_{ST} values. In their study, the effect was seen in towns and in the Japanese prefectures where the towns were located; hence the name. In Uruguay, departments are analogous to Japanese prefectures, so the difference in α between the country as a whole (302) and the departmental average (235.8) is a standard Prefecture Effect.

Values of the inbreeding coefficients and of α for departments are given in Table 1. The highest value of α (393.4), corresponding to the lowest random inbreeding, F_{ST} (0.000637), were observed in Paysandú, followed by Soriano and (376.5, $p < 0.001$ respectively) and Colonia (361.5, $p < 0.001$). These three departments are on the Argentinean border, along the Uruguay River. The opposite values correspond to the north-east, being lowest in Rocha (121.8) and highest
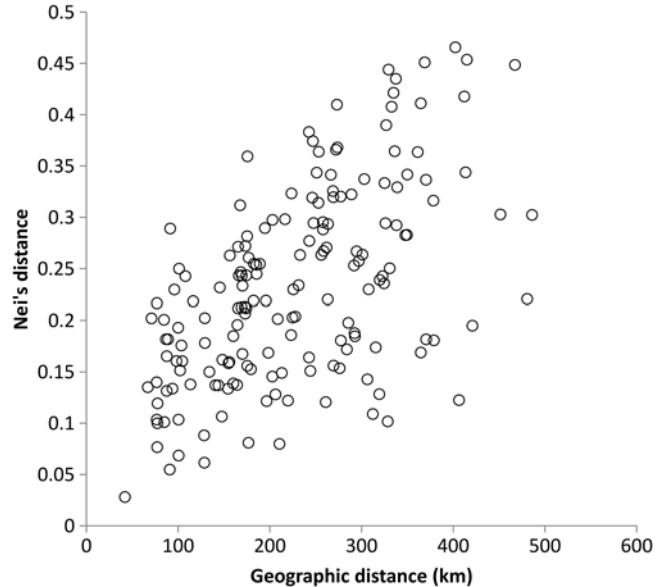
Figure 2. Variation of Nei's distance between departments with geographic linear distance.

F_{ST} (0.002040) followed by Cerro Largo (154.8, $p < 0.002$ respectively) and Rivera (161.1, $p < 0.002$). All these departments are on the Brazilian border. In Montevideo, the department with the highest number of electors, α, was 346.1, while at the other extreme, the department of Flores, with about fifty times fewer electors than the capital, had an α of 208.8, indicating a minor effect of sample size on α.

Isolation by Distance

Isolation by Distance was studied through the correlation between geographic and surname distances at the department level. Nei's, Euclidean and Lasker's distances between the nineteen departments were significantly correlated with linear geographic distances, being highest for Nei's ($r = 0.57 \pm 0.073$), intermediate for Euclidean ($r = 0.46 \pm 0.079$) and lowest for Lasker's ($r = 0.17 \pm 0.097$) distances. As an example, the scatter diagram of Nei's distance between departments over geographic distance is given in Fig. 2. Nei's distance had the largest correlation.

The signal extracted from the scatter diagram of Nei's distance for departments is given in Fig. 3. A clear tendency towards an asymptote is not observed, as it was in Spain, Bolivia, Chile and Honduras (Rodríguez-Larralde *et al.*, 2003, 2011; Barrai *et al.*, 2012; Herrera Paz *et al.*, 2014). In these countries the relation between isonymic and geographic distance flattened after 100 km. There was a continuous increase in surname distance up to 400 km, indicating the presence of increasing isolation and drift up to that distance. In total 171 pairs of isonymic–geographic distances were obtained – a number too small for strong considerations, particularly in comparison with other South American countries. These data can be related to two facts: 1) the lack of natural barriers inside the country, and 2) the fact that, during historic times, the country was mainly populated from the Montevideo area, radiating from the south to the rest of the territory.

Kinship

Random kinship between departments was plotted as a function of geographic distance (Fig. 4). Kinship tended to decrease with distance, as predicted by Malécot (1955) (see also Kimura, 1960).

However, given the very small number of points relating isonymic and geographic distances, the kinship decay could not be described as exponential. Specifically, the exponential decay should be characteristic of structures that are more linear than Uruguay; for example, as observed in Chile (Barrai *et al.*, 2012), Honduras (Herrera Paz *et al.*, 2014) and in Albania (Mikerezi *et al.*, 2013). In the case of Uruguay, kinship was found to decrease with distance, but it was not possible to determine the shape of the decay function.
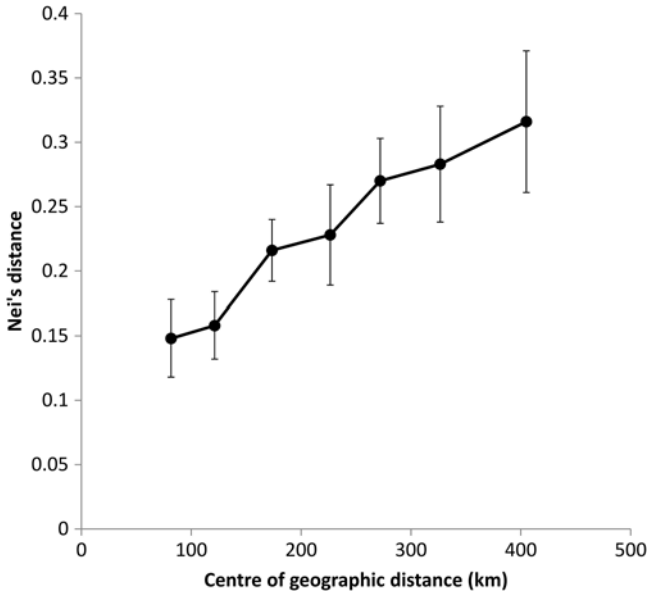
Figure 3. Signal extraction from the variation of Nei's distance (±SD) between departments over geographic distances.
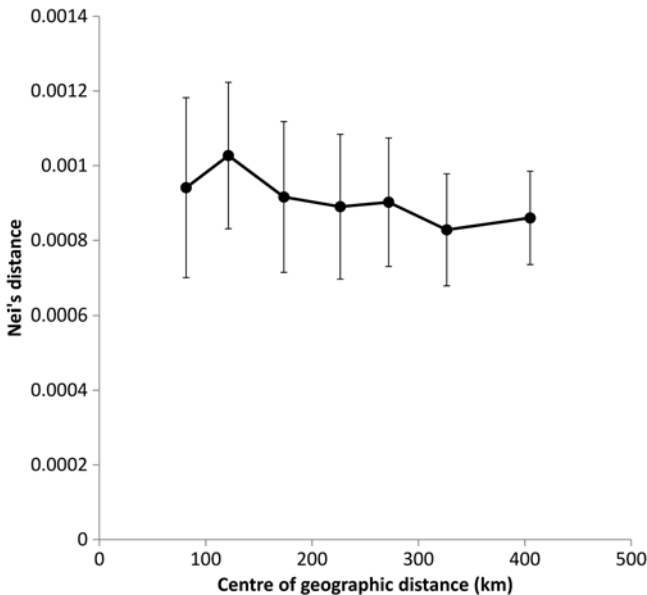
Figure 4. Decay of random kinship (±SD) in Uruguay over geographic distance and pairwise distances between departments.

Relations between departments of Uruguay

In order to obtain a general idea on the movements of population groups in Uruguay, MDSs and PCAs were performed on the matrix of Nei's isonymic distances between departments. The PCA projection on the first two axes of Nei's matrix between departments (Fig. 5) indicated one south-centre cluster formed by Colonia, Soriano, Flores, San José, Florida and Lavalleja, and another cluster in the south including Montevideo, Canelones and Maldonado. The other departments were projected less closely.
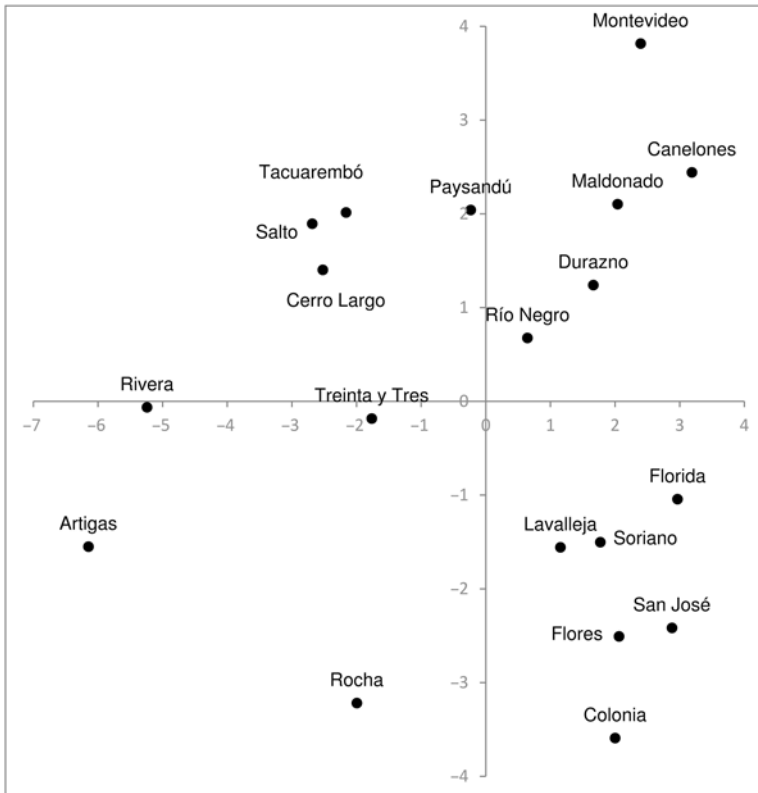
Figure 5. Projection of Nei's distance matrix on the first two components of the PCA. The first component removes 44.36% of variability, and the second component 26.03%.

Figure 6 shows the dendrogram obtained from Nei's matrix. From right to left, a first cluster is composed of two subclusters: the most southern one (Maldonado, Canelones, Montevideo and Lavalleja) and the most eastern one (Cerro Largo, Treinta y Tres, associated with coastal Rocha). A second cluster, also with two subclusters, is formed by the west-central departments of Durazno, Soriano, Río Negro and Paysandú on the right bank of the Uruguay River, on the one hand, and by south-central Flores, Florida, San José and Colonia, on the other. Finally, the last departments to join the dendrogram are the most north-western ones: Salto and Artigas bordering with Rivera and Tacuarembó. All these clusters are related to historical events, and despite the lack of data, seem to be related also to molecular information (Sans *et al.*, 1997, 2006, 2011, 2015; Bonilla *et al.*, 2004; Gascue *et al.*, 2005; Hidalgo *et al.*, 2014). In this sense, in the first subcluster, Montevideo has only a 1% Native contribution using classical markers and 21% considering mitochondrial DNA (mtDNA); in the second subcluster, Cerro Largo has 8% Native contribution using classical
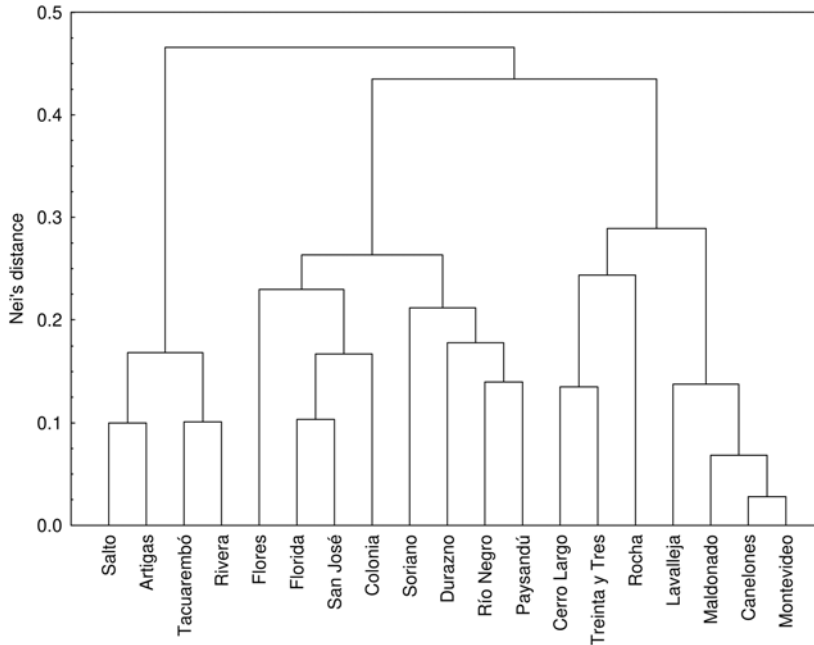
Figure 6. Dendrogram of Uruguay departments showing Nei's distance, based on complete-linkage clustering method of cluster fusion.

markers and 32% with mtDNA; and in the last cluster, Tacuarembó has 20% and 62% respectively (Sans *et al.*, 1997, 2006; Bonilla *et al.*, 2004; Gascue *et al.*, 2005).

Mapping of the first three components of Nei's matrix

The structures revealed by the MDSs and the dendrograms are only partially indicative of the possible movements of the population. Therefore, to get a general idea of the direction of any population movements in Uruguay (following Menozzi *et al.*, 1978), the first three components of Nei's distance matrix, obtained from the PCA and MDS, were mapped (following Menozzi *et al.*, 1978). The PCA components were provided because the relative importance of each component was given by the corresponding eigenvalue, while the MDS provides the value of the stress for a judgement of the overall fitting on the dimensions. The resulting maps are given in Fig. 7.

The intensity of colour in each map is proportional to the deviation of the department on the respective axis. The map variation of the first component of the PCA, which accounts for almost half the variability (44.36%) in the north–south direction, indicates a sense of movement from the south of the country towards the north and the east along the Río de la Plata. This might mean that the main immigration passed through Montevideo, in the extreme south of the country, as seen when analysing Isolation by Distance. The second and third components (26.03% and 13.86%) give a somewhat similar indication, although with minor intensity, because of their size. Overall, the three components account for more than 84% of the surname variation as obtained from Nei's distance matrix. Then, the sense of movement may be postulated as being from the south towards the north, and from west towards the east. This might be inferred also from looking at the PCAs in Fig. 7.

The mappings of the first three dimensions of the MDS are compatible with those obtained from the PCA. The indication of possible movement towards the east and north seems clear enough for the first and second dimension, but less so for the third. So, the isonymic structure of Uruguay seems to be mainly due to migration from the Plata regions, with radiation towards
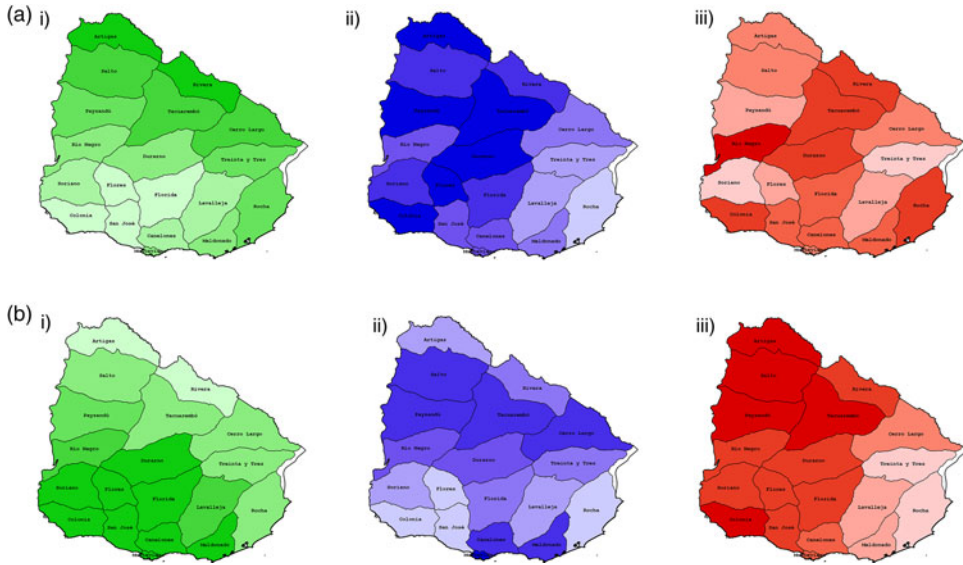
Figure 7. Projection of Nei's matrix of surname distances on districts of Uruguay by mapping (a) the first three PCA factors (i: Factor 1 = 44.36.5%; ii: Factor 2 = 26.03%; iii: Factor 3 = 13.86%); (b) the first three MDS dimensions (i: Dimension 1; ii: Dimension 2; iii: Dimension 3. Stress = 10.47%).

the east and north, and with subsequent isolation and drift, with short-range migration playing a major (and drift a minor) role in the generation of the present geographical variation in surnames.

At present, most internal migration seems to take place towards the capital and other main towns upstream and on the left bank of the River Uruguay. Although these movements have not been documented, recent internal migrations show that Montevideo is no longer the main centre of internal population movement and, moreover, part of its population has migrated to the neighbouring area of Canelones (Pellegrino, 2003). These results are an addition to the available knowledge that has been acquired over the course of time to meet the challenges of planning and organization of the territory in Uruguay (Yagüe & Díaz Puente, 2008).

The methodology described in this paper has been used to analyse the isonymic structure of several South American countries (Rodríguez-Larralde *et al.,* 2000, 2011; Dipierri *et al.,* 2005, 2011; Barrai *et al.,* 2012) and in Central American Honduras (Herrera Paz *et al.,* 2014). In these countries, 4 (Venezuela), 24 (Argentina), 23 (Bolivia), 4.5 (Paraguay), 4.5 (Honduras) and 16.5 (Chile) million surnames from censuses and electoral registers were used, similar to the Uruguay study. In contrast, in European countries and in the USA surnames have been taken from telephone directories (Barrai *et al.,* 2001; Scapoli *et al.,* 2005, 2007; Rodríguez-Larralde *et al.,* 2007).

The average value of α for different population unities (cities, states, districts, provinces or departments depending on the country, divisions that are relatively comparable) and Isolation by Distance, as measured by the correlation between isonymic and geographic distances, are given in Table 2 for the countries studied up to now. Several features emerge from this comparison. First, there is a general similarity among European nations and the USA in the profusion of surnames as measured by α, with the exception of Albania and Spain. Uruguay has an intermediate position between these two countries and the rest of Europe/USA, and is different to South American countries, with the exception of Argentina. Second, for Isolation by Distance as measured by the linear correlation, Spain and USA have the lowest values, and Italy, France and Venezuela, the highest, with Uruguay being close to the average. Finally, the type-token ratio (SS/S) is lower in Uruguay relative to other South or Central American countries, and is similar to those of Spain and little higher than those of most of the rest of Europe.

Conclusions

In conclusion, surnames can be used to show social patterns related to geographic spread and distance, ancient or recent migration movements, population dynamics, ethnic or geographic origins, ethnicity or other historical facts (Colantonio *et al.*, 2003; Darlu *et al.*, 2012; Cheshire, 2014). Despite the lack of names of Native or African origin, illegitimate unions, as well as different spellings or surname changes, isonymy studies can work with a large amount of data allowing comprehensive analysis of present and past populations. This study's analysis in Uruguay has several interesting finding, including the general similarity among European nations in the profusion of surnames as measured by α and for Isolation by Distance as measured by the linear correlation. Moreover, in Uruguay, the average number of persons with the same surname is small (36) for a South American country.

Different regions can be defined in Uruguay on the basis of longitude. The random inbreeding estimate (F_{ST}) was lower and the effective quantity of surnames (α) higher in the more densely populated south-west area, and in Montevideo, with Paysandú, Río Negro, Soriano and Colonia having the lowest levels of inbreeding. It can be concluded that the current population structure in Uruguay is the result of the action of short-range directional migration, particularly along the River Uruguay and along the Río de la Plata. The north-east (Rivera, Tacuarembó, Cerro Largo) has the highest level of inbreeding F_{ST} and the lowest α, indicating greater isolation. In contrast to regions that were populated by a series of migration waves of different origins, the north-east had a historic penetration from the Brazilian border and the inheritance of land (Carvalho Neto, 1965; Rama, 1967).

It should be mentioned that in Uruguay electors are registered according with their place of residence when they were around 18 years old; if they move, even to neighbouring countries, they do not always change their addresses. So, internal migration cannot be always be seen in the available electoral registers. Nevertheless, from the current analysis it appears that migration is a major contributor to surname differentiation, while the contribution of drift, considering the small variance in α, seems to be minor.

Author ORCIDs. M. Sans, 0000-0002-1564-8761

References

Acosta y Lara E (1981) Un linaje Charrúa en Tacuarembó. *Revista Facultad de Humanidades y Ciencias (Montevideo), serie Ciencias Antropológicas* **1**, 1–30.

Adamic LA and Huberman BA (2002) Zipf law and the internet. *Glottometrics* **3**, 143–150.

Barrai I, Rodríguez-Larralde A, Dipierri J, Alfaro E, Acevedo N, Mamolini E et al. (2012) Surnames in Chile. A study of the population of Chile through isonymy. *American Journal of Physical Anthropology* **147**, 380–388.

Barrai I, Rodríguez-Larralde A, Mamolini E, Manni F and Scapoli C (2000) Elements of the surname structure of Austria. *Annals of Human Biology* **27**, 607–622.

Barrai I, Rodríguez-Larralde A, Mamolini E, Manni F and Scapoli C (2001) Isonymy structure of USA population. *American Journal of Physical Anthropology* **114**, 109–123.

Barrai I, Scapoli C, Beretta M, Nesti C, Mamolini E and Rodríguez-Larralde A (1996) Isonymy and the genetic structure of Switzerland. I: the distributions of surnames. *Annals of Human Biology* **23**, 431–455.

Barrán JP and Nahum B (1971) *Historia rural del Uruguay moderno.* Banda Oriental, Montevideo.

Barreto I (2011) *Estudio biodemográfico de la población de Villa Soriano.* Biblioteca Plural, Montevideo.

Bertoni B, Budowle B, Sans M, Barton SA and Chakraborty R (2003) Admixture in Hispanics distribution of ancestral population contributions in the Continental United States. *Human Biology* **75**, 1–11.

Bonilla C, Bertoni B, González S, Cardoso H., Brum-Zorrilla N and Sans M (2004) Substantial Native American female contribution to the population of Tacuarembó, Uruguay, reveals past episodes of sex-biased gene flow. *American Journal of Human Biology* **16**, 289–297.

Bonilla C, Bertoni B, Hidalgo PC, Artagaveytia N, Ackermann E, Barreto I et al. (2015) Breast cancer risk and genetic ancestry: a case-control study in Uruguay. *BMC Cancer* **15**, 434.

Bravi CM, Sans M, Bailliet G, Martínez-Marignac VL, Portas M, Barreto I et al. (1997) Characterization of mitochondrial DNA and Y-chromosome haplotypes in a Uruguayan population of African ancestry. *Human Biology* **69**, 641–652.

Carvalho Neto P de (1965) *El negro uruguayo: hasta la abolición* Editorial Universitaria, Montevideo

Cavalli-Sforza LL and Edwards AWF (1967) Phylogenetic analysis models and estimation procedures. *American Journal of Human Genetics* **19**, 233–257.

Cheshire J (2014) Analysing surnames as geographic data. *Journal of Anthropological Sciences* **92**, 99–117.

Colantonio SE, Lasker GW, Kaplan BA and Fuster V (2003) Use of surname models in human population biology: a review of recent developments. *Human Biology* **75**, 785–807.

Crow JF and Mange A (1965) Measurements of inbreeding from the frequency of marriages between persons of the same surname. *Eugenics Quarterly* **12**, 199–203.

Cuba Manrique MC (2002) Antroponimia e identidad de los negros esclavos en el Perú. *Escritura y Pensamiento* **5**, 123–134.

Darlu P, Bloothooft G, Boattini A, Brouwer L, Brouwer M, Brunet G et al. (2012) The family name as socio-cultural feature and genetic metaphor: from concepts to methods. *Human Biology* **84**, 169–214.

Dipierri JE, Alfaro EL, Scapoli C, Mamolini E, Rodríguez-Larralde A and Barrai I (2005) Surnames in Argentina. A population study through isonymy. *American Journal of Physical Anthropology*, **128**, 199–209.

Dipierri JE, Rodríguez-Larralde A, Alfaro EL, Scapoli C, Mamolini E, Salvatorelli G et al. (2011) Surnames in Paraguay: a study of the population of Paraguay through isonymy. *Annals of Human Genetics* **75**, 678–687.

Flores RD (2010) Familias británicas en la sociedad rural argentina, 1866–1912. *Epocas (Revista de Historia, USAL)* **3**, 95–132

Fox WR and Lasker GW (1983) The distribution of surname frequencies. *International Statistical Review* **51**, 81–87.

Gascue C, Mimbacas A, Sans M, Gallino JP, Bertoni B, Hidalgo P and Cardoso H (2005) Frequencies of the four major Amerindian mtDNA haplogroups in the population of Montevideo, Uruguay. *Human Biology* **77**, 873–878.

González-Rissotto R and Rodríguez-Varese S (1989) La importancia de las Misiones Jesuíticas en la formación de la sociedad uruguaya. *Estudos Ibero-Americanos (PUCRS)* **15**, 191–214.

Herrera Paz EF, Scapoli C, Mamolini E, Sandri M, Carrieri A, Rodríguez-Larralde A and Barrai I (2014) Surnames in Honduras: a study of the population of Honduras through isonymy. *Annals of Human Genetcis* **78**, 165–177.

Hidalgo PC, Bengochea M, Abilleira D, Cabrera A and Alvarez I (2005) Genetic admixture estimate in the Uruguayan population based on the loci LDLR, GYPA, HBGG, GC and D7S8. *International Journal of Human Genetics* **5**, 217–222.

Hidalgo PC, Mut P, Ackermann E, Figueiro G and Sans M (2014) Questioning the "melting pot": analysis of Alu inserts in three population samples from Uruguay. *Human Biology* **86**, 83–92.

INE (2011) *Censos 2011.* URL: http://www.ine.gub.uy/censos2011/index.html (accessed 6th November 2017).

INE (2012) *Uruguay en cifras 2012.* Instituto Nacional de Estadísticas, Montevideo. URL: http://www.ine.gub.uy/documents/10181/39317/Uruguay+en+cifras+2012.pdf/8a922fc6-242a-4ecc-a145-c334825c8dbd (accessed 6th June 2018).

Kimura M (1960) *Outline of Population Genetics.* Baifukan, Tokyo.

Lucas K (1992) *Rebeliones indígenas y negras en América Latina: entre viento y fuego.* Abya Yala, Quito.

Malécot G (1955) Decrease of relationship with distance. *Cold Spring Harbour Symposium* **20**, 52–53.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209–220.

Menozzi P, Piazza A and Cavalli-Sforza LL (1978) Synthetic maps of human gene frequencies in Europeans. *Science* **201**, 786–792.

Mikerezi I, Xhina E, Scapoli C, Barbujani G, Mamolini E, Sandri M et al. (2013) Surnames in Albania. A study of the population of Albania through isonymy. *Annals of Human Genetics* **77**, 232–243.

Nei M and Imaizumi J (1966) Genetic structure of human populations. I. Local differentiation of blood groups gene frequencies in Japan. *Heredity* **21**, 9–36.

Nei M (1973) The theory and estimation of genetic distance. In Morton NE (ed.) *Genetic Structure of Populations.* Hawaii University Press, Honolulu, 45–54.

Pellegrino A (2003) *Caracterización demográfica del Uruguay.* UNFPA-Facultad de Ciencias Sociales, Universidad de la República, Montevideo.

Pi Hugarte R and Vidart D (1969) *El legado de los inmigrantes, I.* Nuestra Tierra, **29**. Banda Oriental, Montevideo.

R Development Core Team (2012) *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna.

Rama C (1967) *Los Afrouruguayos*. El Siglo Ilustrado, Montevideo.

Relethford JH (1988) Estimation of kinship and genetic distance from surnames. *Human Biology* **60**, 475–492.

Rodríguez-Larralde A, Dipierri J, Gómez EA, Scapoli C, Mamolini E, Salvatorelli G et al. (2011) Surnames in Bolivia: a population study through isonymy. *American Journal of Physical Anthropology* **144**, 177–184.

Rodríguez-Larralde A, González-Martín A, Scapoli C and Barrai I (2003) The names of Spain: a study of the isonymy structure of Spain. *American Journal of Physical Anthropology* **121**, 280–292.

Rodríguez-Larralde A, Morales J and Barrai I (2000) Surname frequency and the isonymy structure of Venezuela. *American Journal of Human Biology* **12**, 352–362.

Rodríguez-Larralde A, Scapoli C, Beretta M, Nesti C, Mamolini E and Barrai I (1998) Isonymy and the genetic structure of Switzerland. II. Isolation by distance. *Annals of Human Biology* **25**, 533–540.

Rodríguez-Larralde A, Scapoli C, Mamolini E and Barrai I (2007) Surnames in Texas: a population study through isonymy. *Human Biology* **79**, 215–239.

Rosal M A (2009) *Africanos y afrodescendientes en el Río de la Plata, Siglos XVIII–XIX*. Dunken,Buenos Aires.

Sans M, Alvarez I, Callegari-Jacques SM and Salzano FM (1994) Genetic similarity and mate selection in Uruguay. *Journal of Biosocial Science* **26**, 285–289.

Sans M, Figueiro G, Ackermann E, Barreto I, Egaña A, Bertoni B et al. (2011) Mitochondrial DNA in Basque descendants from the city of Trinidad, Uruguay: Uruguayan- or Basque-like population? *Human Biology* **83**, 55–70.

Sans M, Merriwether DA, Hidalgo PC, Bentancor N, Weimer TA, Franco MH et al. (2006) Population structure and admixture in Cerro Largo, Uruguay, based on blood markers and mitochondrial DNA polymorphisms. *American Journal of Human Biology* **18**, 513–524.

Sans M, Mones P, Figueiro G, Barreto I, Motti JM, Coble MD, Bravi CM and Hidalgo PC (2015) The mitochondrial DNA history of a former native American village in northern Uruguay. *Amercian Journal of Human Biology* **27**(3), 407–416.

Sans M, Salzano FM and Chakraborty R (1997) Historical genetics in Uruguay: estimates of biological origins and their problems. *Human Biology* **69**, 161–170.

Sans M, Weimer TA, Franco MH, Salzano FM, Bentancor N, Alvarez I et al. (2002) Unequal contributions of male and female gene pools from parental populations in the African descendants of the city of Melo, Uruguay. *American Journal of Physical Anthropology* **118**, 33–44.

Scapoli C, Goebl H, Sobota S, Mamolini E, Rodríguez-Larralde A and Barrai I (2005) Surnames and dialects in France: population structure and cultural evolution. *Journal of Theoretical Biology* **237**, 75–86.

Scapoli C, Mamolini E, Carrieri A, Rodríguez-Larralde A and Barrai I (2007) Surnames in Western Europe: a comparison of the subcontinental populations through isonymy. *Theoretical Population Biology* **71**, 37–48.

Smouse PE, Long JC and Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology* **35**, 627–632.

Tarskaia L, El'chinova GI, Scapoli C, Mamolini E, Carrieri A, Rodríguez-Larralde A and Barrai I (2009) Surnames in Siberia. A study of the population of Yakutia through isonymy. *American Journal of Physical Anthropology* **138**, 190–198.

Vidart D and Pi Hugarte R (1969) *El legado de los inmigrantes, II*. Nuestra Tierra, **39**. Montevideo.

Yagüe JL and Díaz Puente JM (2008) Tres siglos de planificación regional en Uruguay: lecciones de experiencia para afrontar los retos de desarrollo en el siglo XXI. *Estudios Geográficos* **LXIX/264**, 247–280.

Yarza Rovira EJ (2009). Los Artigas de Casupá. La estirpe de Pedro Mónico Artigas. *Revista del Instituto de Estudios Genealógicos del Uruguay* **32**, 420–453.

Yasuda N and Morton NE (1967) Studies on human population structure. In Crow JF and Neel JV (eds) *Proceedings of the 3rd Congress of Human Genetics, Chicago, 1966*. University of Chicago Press, Chicago, pp. 249–265.