Scientific
Research
Publishing

The Geography as a Regulator of Genetic Flow and Genetic Structure in Andorra

Elena Campoy, Antonio González-Martín*

Department of Zoology and Physical Anthropology, Faculty of Biology, Complutense University, Madrid, Spain
Email: *antonio@ucm.es

Abstract

The Principality of Andorra is one of the smallest states in Europe. Traditionally its economy has been based on the agriculture and cattle ranch although in the last decades it has developed an important tourist activity that has become a collector of emigrants. The origin of marriage consorts has been analyzed in the six parishes that have traditionally constituted the Principality. In total, there have been 10,208 marriages covering a continuous period from 1606 to 1960. From this information, two migration matrices have been constructed, one general and one intra-population. The study of the first one shows the existence of a male migration to long distance, mainly from the nearby Catalan provinces. It also shows that the main parameter that regulates migration is geography and that it indirectly defines the language's performance and political boundaries as the genetic filters. For the study of the intrapopulation matrix, two methods have been used. On the one hand, a tree has been constructed and has been linked to Principal Component Analysis (PCA). Different matrices have also been tested by applying the Mantel test. The results indicate that there is no asymmetry in the mobility of men and women in an intra-population level, a result that is justified by the demographic and social structure of the small population centers of the Pyrenees. It is also shown that geography is the main governing factor travel within the Principality. And finally, we can observe the existence of a genetic substructure in the Andorran population marked again by geography.

1. Introduction

One of the main phenomena that determine the human population genetic composition and structure is mobility (Rosenberg et al., 2002). Its effect produces,

on one hand, an increase of the genetic diversity of the receiver population, and on the other hand, it genetically homogenizes the emitting and receiver populations by means of genetic flow (Jobling, 2012).

The migration, or genetic flow, is a complex phenomenon whose effects depend on numerous factors (Hartl & Clark, 2007). Furthermore, in our species, it is related to cultural barriers (language, religion or political borders). In fact, the potential isolation of a human population is a combination of parameters with different nature: both geographical and cultural barriers have to be considered in our species (Sánchez-Mazas & Barbujani, 2013).

Normally, the migratory movements imply medium and long distances displacements (Castles, 2003), but there exists another type of movement (more difficult to detect) in a small-scale and that is frequent inside the populations. These movements let us to determine the existence of internal subdivisions. These subdivisions determine the genetic structure of a population, increase the rate of homozygosis (Wahlund, 1928) and influence the level of micro-differentiation of communities (Wright, 1951; Harpending & Jenkins, 1973; Nei, 1972, 1973). Another consequence of a population's subdivision is the increase of the differential amount of isolation, thus, increasing the likelihood of a local genetic adaptation (Slatkin, 1987; McCullough, 1989). In addition, knowing the genetic structure of the populations is a basic parameter to reconstruct their biological history (Tishkoff et al., 2000) and to understand the distribution of some diseases (Ziv & Burchard, 2003; Reiner et al., 2005).

However, addressing the intrapopulation movements in anthropology is not easy. Not even the molecular tools allow the detailed reconstruction of the movements produced inside a population and their biological consequences. Nevertheless, the singularity of our species provides anthropological techniques that can be used to detect these movements. In particular, the habit that our ancestors have had to write down the vital events produced inside a population and that are conserved in religious and/or civil files. These files have to accomplish some properties: they have to be continuous in time and geography and they have to contain valuable information about the objectives of their exploitation. The detailed study of these files provides information about the genetic structure of a population, and this information can be used in subsequent molecular studies to enhance their resolution. Therefore, geography is not the only factor influencing the structure of a population (migration and genetic relationship between two populations are smaller when the distance between them is bigger) (Barbujani & Sokal, 1990, 1991) and we have to take into account cultural factors.

The religious files of the Principality of Andorra achieve these qualities and amount of information expectations. Besides, this population has two singularities that are not very common in anthropological studies. First, it is a country, and this allows the detailed reconstruction of the demographical events that have taken place in the state. Second, the information can be used to determine the internal micro-mobility, but also, to value the effect of geographical, political or cultural barriers as biological filters.

In this work, we study, from demographic data, the origin of Andorran spouses in order to determine if political, geographical and cultural boundaries have acted as genetic filters. Additionally, we will study internal mobility patterns of Andorran spouses and their relation to the topography of the country. Finally, from the matrimonial interchange patterns, we will determine if there are any subdivisions and if these are significant enough as to determine the existence of an internal genetic structure.

2. Materials and Methods

The Principality of Andorra. Andorra is one of the three smallest European countries and has an area of 648 km². It is located on the southern slope of the Central Pyrenees (Figure 1), and it is made up of three valleys surrounded by mountains of more than 3000 m altitude.

The History of the Principality of Andorra is atypical within the European context. The first references of the country's existence date back to the 9th Century. Originally the territory belonged to the Counts of Urgell (Catalonia, Spain) but, during the 10th and 11th Centuries, they ceded their rights to the Catholic Church which then adopted a feudal government system. Since those centuries and practically until a few years ago, the Principality has been governed by a system named "Co-Principality". The origin of such a system is not very clear. It consists of a shared government between the President of the Republic of France

Figure 1. Situation of the Principality of Andorra and its population centers. The underlined nuclei represent the municipal capitals.

and the Bishop of Urgell, situated in the Catalan region of Lleida. This circumstance allowed this small state to maintain its independence from the annexation desires of the two neighbouring countries.

Its inhabitants have been devoted to agriculture and migrating cattle raisings. Nowadays, however, most of its population is devoted to trade and tourism (Calvo et al., 1990). Actually, the demographic history of Andorra continues to develop around the exploitation systems that the population has adopted throughout its history.

Demographically, Andorra maintained a stable population of around 5000 inhabitants up to the end of the 19th century. The introduction of new exploitation strategies—agriculture and livestock have been substituted by tourism and commerce (González-Martín, 2008)—has resulted in an increase of the population due to vegetative growth as well as significant migrations (Adellach & Ganyet, 1977). As such, in the official census of 1947 there were 5385 inhabitants, in 1990 there were 54,507 inhabitants, and in the last census of 2016, Andorra counted with 78,264 inhabitants (Departament d'Estadística, 2017). Nevertheless, up to the middle of the 20th Century, the population maintained a demographic balance principally towards the nearby bordering Catalan territories (González-Martín and Toja, 1996). This situation radically changes the adoption of the new economic strategies resulting, as a repercussion, in a growing demographic explosion.

It is convenient to state clearly that the country maintains a strong cultural homogeneity in spite of its geographical, demographical and economical variations.

The source of information. Data have been obtained from matrimonial archives in the six parishes that have traditionally constituted the Principality of Andorra. The time period represented in the record books is continuous although in each parish the records begin in different years. That is to say, in the parish of Andorra la Vella, records begin in 1650 (with a total of 2885 marriages), in Sant Julià de Lòria they begin in 1700 (1861 marriages), in Encamp they begin in 1627 (1220 marriages), in La Massana they begin in 1673 (1461 marriages), in Ordino they begin in 1729 (1040 marriages) and, lastly, in Canillo records begin in 1606 (1741 marriages). This gives a total of 10,208 marriages for the period 1606 to 1960. In 9749 (95.5%) of these matrimonies the origin of both spouses is known. This number represents the 100% of the marriages of the country since Andorra is a traditional catholic country where all the marriages have had to be registered in their corresponding parishes.

Methodology of work. We have two objectives. On one side, to determine the origins of the migrants that arrive at the Principality to know the parameters that have actuated as genetic filters. On the other hand, determining, through the models of interchange related by the marriage among nuclei of population, the existence of internal subdivisions. It does however, have a disadvantage: the model for matrimonial interchange varies over time as can be observed in Table 1. Nevertheless, the objective of this work is not to make a temporary analysis of

Table 1. Temporal variation of migration taxes: in line origins of migrants. N; number of individuals. Bordering; refers to the Catalan communities that bordered Andorra on the southern slope.

Period	N	Bordering	France	Rest bishopric	Other origins	Total
1601	125	2.40	0.80	0	0	3.20
1626	333	4.50	3.30	0	0	7.81
1651	529	4.54	3.02	0.19	0.57	8.32
1676	816	4.17	1.59	0	0.24	6.01
1701	1147	6.54	2.27	0	0.43	9.24
1726	1256	7.80	1.27	0.24	0.56	9.87
1751	1463	9.30	1.23	0.07	1.23	11.82
1776	1521	9.60	1.05	0.39	0.99	12.03
1801	1496	11.10	0.47	0.53	0.47	12.57
1826	1999	8.95	0.85	0.70	1.40	16.66
1851	2071	9.99	0.48	0.39	1.83	12.70
1876	1989	17.55	0.95	1.86	3.17	23.53
1901	2467	24.12	1.30	8.11	14.35	47.87
1926	1723	21.30	3.13	4.93	23.79	53.16
1951	952	17.75	4.20	4.31	40.55	66.81
Total	19,887	13.57	1.49	2.03	6.72	23.60

the changes that have been produced in these models, but to determine, notwithstanding the temporary swings in the matrimonial structure, an underlying general model in the genetic structure of the population. Also, it must be clarified that the essential premise of this work is to represent the major number of population nuclei in the country which, indirectly, represent the maximum of the geographic variation. From this perspective, the study of the temporary variations in the matrimonial interchange model is not viable because fragmenting the database implies that the size of the sample of each population will considerably decrease, above all, in the oldest time periods.

From this data, an origin matrix has been built (Bodmer & Cavalli-Sforza, 1968) grouping the origin of spouses into ten categories; six corresponding to the parishes in the Principality and four referring to the place of birth of foreign individuals (Table 2). This matrix quantifies the genetic flux produced within a population and among neighboring communities and reports on the distance of gametes union. Additionally, it allows us to determine if matrimonies unite, genetically, from geographically different areas (Coleman, 1977).

On the other hand, the origin of individuals born and married in the Principality has been studied. In the study, all populations representing at least 1.25% of the total of references of native Andorrans married in the same Principality are expressed (Table 3). Small populations that did not meet this condition have been added to the nearest nuclei that met such condition: a matrix of 28 × 28

Table 2. Migration matrix of marriages carried out in Andorra, from 1601 to 1960. Female spouses are represented on rows and male in columns. Bordering; refers to the Catalan communities that bordered Andorra on the southern slope.

	Andorra	Canillo	Encamp	La Massana	Ordino	Sant Juliá	Bordering	France	Rest bishopric	Other origins	Total
Andorra	923	74	73	126	68	114	319	46	29	95	1898
Canillo	96	1032	128	84	111	47	95	37	5	28	1663
Encamp	94	98	431	58	25	43	106	26	8	33	922
La Massana	194	111	60	615	161	74	179	22	8	26	1450
Ordino	57	83	26	138	417	26	82	15	8	34	886
Sant Juliá	122	41	25	40	28	473	304	18	11	61	1123
Bordering	144	36	45	57	41	119	414	17	39	106	1018
France	9	4	6	3	7	11	11	14	1	13	79
Rest bishopric	11	2	2	2	2	3	32	2	92	25	173
Other origins	17	5	6	6	7	15	69	14	29	369	537
Total	1667	1486	802	1129	867	955	1611	211	230	791	9749

Table 3. Locations used for Andorra's internal structure study. Abbreviation of populations, number of male and female individuals addressed on files, total and percent, are expressed.

Nuclei of population	No. ♂	No. ♀	Total	%
Sant Julià de Lòria	530	653	1,183	7.80
Aixovall	143	163	306	2.02
Aubinyà	143	140	283	1.86
Aixirivall	157	177	334	2.20
Andorra la Vella	653	808	1,461	9.63
Les escaldes	508	559	1,067	7.03
Engordany	394	471	865	5.70
Santa Coloma	147	166	313	2.06
Encamp	279	344	623	4.11
El tremat	239	216	455	3.00
La mosquera	303	383	686	4.52
La Massana	340	436	776	5.11
Anyós	113	169	282	1.86
Arinsal	143	195	338	2.23
Erts	112	142	254	1.67
Escás	105	138	243	1.60
Pal	118	134	252	1.66
Sispony	205	257	462	3.04
Canillo	819	933	1,752	11.55
Meritxell	121	148	269	1.77
Prats	172	182	354	2.33
Ransol	188	182	370	2.44
Soldeu	131	129	260	1.71
Aldosa	88	119	207	1.36
Ordino	269	313	582	3.84
La cortinada	234	206	440	2.90
Llorts	213	215	428	2.82
Sornàs	160	166	326	2.15
Total	7027	8144	15,171	100

was obtained. **Figure 1** represents the location of these populations. The distances covered between populations express the net of relations established by socioeconomic (Harrison et al., 1974; Susanne, 1983), religious (Segalen & Jacquard, 1973), cultural (Boyce et al., 1967; Kücheman et al., 1967), administrative means and geographical activities.

In order to facilitate statistical applications, the internal mobility matrix has been modified. This correction consisted of independently taking into account spouses mobility from that of gender, that is, transforming a square matrix into a hemi-matrix expressing in every box the sum of matrimonies corresponding to ♂ i − ♀ j and ♂ j − ♀ i. This last grouping would be incorrect if male mobility patterns differ from female migration patterns, thus, their similarity must be proven previously. Therefore, the two hemi-matrixes from the initial square matrix were compared through Mantel's test (Mantel, 1967): there is a positive and significant correlation between both (0.83; $p = 0.0005$), that is to say, male mobility structure displacing from one location to another in order to marry is the same as for women. There is a flux, although not necessarily with the same intensity, proportional and balanced according to gender for the interchange of spouses between different towns.

This matrix was again contrasted through Mantel's test with two geographical matrixes: the first one expresses the distance, in straight line and measured in kilometers, between population nuclei. The second one expresses in kilometers the distance between population nuclei through the roads and pathways most frequently used.

The internal subdivision of the Andorran population and the relation between populations are estimated by the application of the "squared Euclidean distances" (Bisquerra, 1989). The data used are the intensity of the interchanges between population nuclei in the Principality of Andorra. This system used in taxonomic studies of human populations, as well as other multi-variant methods, adapts easily to data employed in bio-demographic studies (Smith & Hudson, 1984; Smith et al., 1984; Fuster, 1985; Pollitzer et al., 1988; Luchetti & Soliani, 1989; Calafell & Hernández, 1993; Esparza et al., 2006; Mikerezi et al., 2013). Finally, distances matrixes have been represented by neighbor-joining representation (Saitou & Nei, 1987; Felsenstein, 1989; Li & Graur, 1991). Node strength in the groups has been evaluated by the application of a "bootstrap" (Efron, 1982; Felsenstein, 1985; Cavalli-Sforza et al., 1994). The resulting graphic will express the existing relation between population nuclei based on a matrimonial interchange pattern and will describe areas where endogamy is at a maximum. Finally, a principal component analysis (PCA) has been built using the size of the interchanges between population nucleus.

3. Results

Matrimony implies, in almost every case, displacement of at least one of the spouses. In fact, displacement expresses gamete origin and, therefore, it is one of the essential parameters in explaining and understanding the genetic structure of

a population.

In spouse's origin matrix (Table 2) can be appreciated the high representation of the main diagonal, above all, spouses from the parish of Canillo, from La Massana, from the boundary region of the principality and other from origins. The main diagonal expresses 49.03% of the total data. In almost 50% of matrimonies in Andorra, both spouses come from the same place.

A second important fact is the asymmetry given in the partial totals of columns and rows, that is, between the origin of male and female spouses. Furthermore, it is an incomplete asymmetry, because, when Andorran parishes are taken into account, the higher partials correspond to female individuals. However, when the rest of the origins are taken into account, male individuals are more represented. This observation reveals the existence of a general and differential migration pattern related to spouse's origin and gender. These results do not disagree with those obtained when comparing, through a Mantel test, the two hemi-matrixes, since calculations have been carried out, exclusively and in much detail on the internal matrix and not for total data. Mantel's test expresses a proportionality relation between the two hemi-matrixes of the internal matrix, that is to say, between the movement of men and of women within the parishes of the Principality. Also, this statistic does not value the numerical differences but rather the proportional relation between both hemi-matrixes. In this case the asymmetry includes all of the individuals who married in the Principality and shows the phenomenon of male hypermobility over large distances, a common situation in Pyrenean populations (Toja, 1987). In fact, one of the phenomena detected in the Principality has been the emission of excess demographics due to migrations, also this collective has been preferentially constituted by male individuals (González-Martín & Toja, 2002).

Another outstanding aspect is the little representation of individuals with French origin, as compared to those coming from Catalan regions. Actually, this is an anticipated result given the geographic differences that exist between both borders, access being very steep in the north and very easy in the sloping or Catalonian south. Some other cultural aspects have to be added to this geographical parameter: Catalan is an official language in the Catalonian neighbor regions as well as in Andorra. It is hard to assess the degree of responsibility that geography and culture have as controllers of the Spanish and French genetic flow but it is reasonable to argue that the combined effect of both has made the population movements from the meridional regions more intense.

Comparisons between the matrix of internal migration and geographic migrations expressed the relation between the topographic distribution of population nuclei and the mobility and genetic interchange of their residents. By contrasting the marital distance matrix with the distances by road, a correlation of −0.52 ($p < 0.001$) was obtained; the same comparison with straight line distances gave a correlation of −0.48 (also with $p < 0.001$). Therefore, there is a clear correlation, inverse and linear, between nuclei situation and genetic flux intensity within the Principality. An outstanding fact is that a great deal of correlation occurs be-

tween the internal matrix of mobility and the real distance between two nuclei. Such detail demonstrates that not only is the distance between nuclei significant in the shaping of marriages, but also the taken route.

The study of interchange intensity between population nuclei has been centred on the graphic representation of the squared Euclidean distances matrix obtained from the distances matrix of internal migration. This new matrix represents the differences produced between the marriage interchange among populations: two populations with no gene interchange will present high Euclidean distance values, and thus, they will be different from a genetic point of view. On the contrary, those which have kept a similar interchange pattern will be up to a minor distance and will find themselves, genetically closer. Starting from the homogenization process, amongst other phenomena, provoked by gene interchange, the nuclei which have maintained high frequencies in individual interchange will be genetically homogeneous compared with those nuclei that have remained demographically isolated and without interchange.

The matrix of squared Euclidean distances has been represented by Neighbor-Joining grouping methods (**Figure 2**). This tree hierarchically groups the
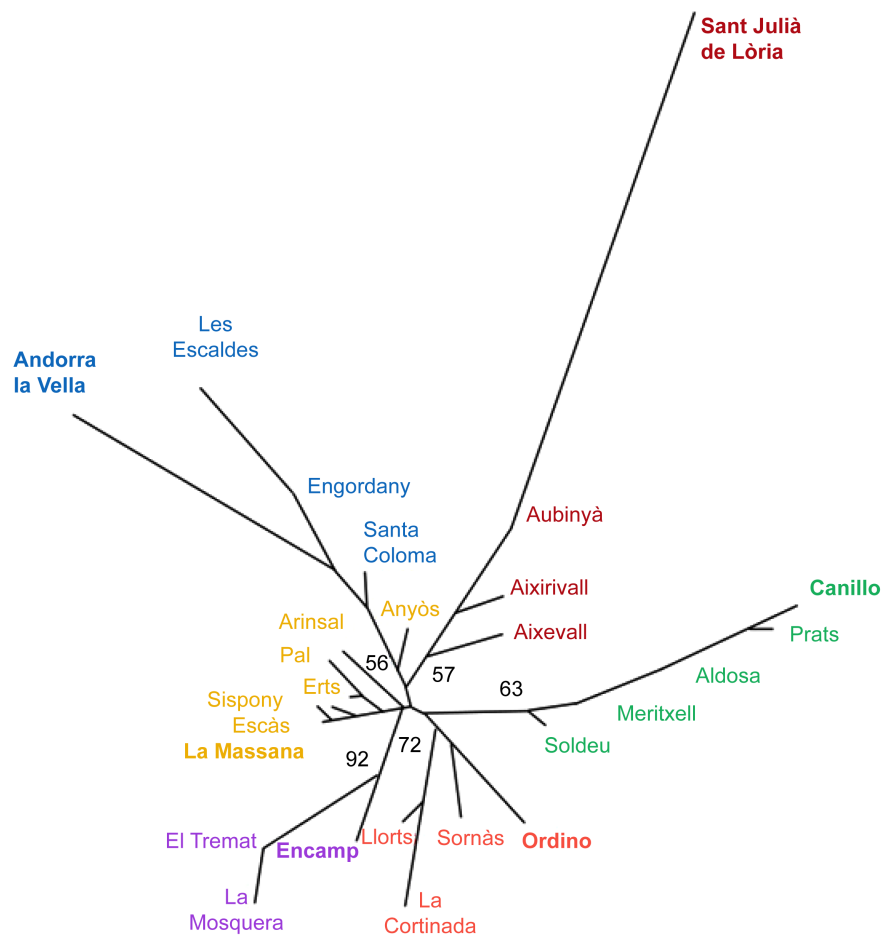

Figure 2. Representation of the Euclidean distance squared by a Neighbor-Joining. Numerical values express the robustness of the stripes calculations in the 1000 repetition bootstrap method. The municipal capitals are in bold.

population nuclei, always linking those presenting less Euclidean distance. To validate this topography, a 1,000 iteration "bootstrap" has been applied. If short distances between two nuclei are based on a wide range of coincidences in various variables, the method will tend to maintain short distances. Now, if affinity between two nuclei is due to great differences, yet in few variables, the "bootstrap" will increase distances between towns. In this way, an estimation of strength and reliability on the trees will be obtained, expressing in percentages the amount of times an association, in a random sampling of data, is repeated.

The main branches have little strength as shown by the low node values: the association suggested by the main branches may have a random origin. However, the derived branches have a strength that reaches in some cases the 90%. Affinities between nuclei constituting the same parish are much more important than branches expressing relations between parishes. In 92% of resampling, the three nuclei constituting Encamp's parish appear to be related to one another. These values are maintained at a high level for the most northerly parishes; 72% for the nuclei belonging to the Ordino parish and 63% for those belonging to the parish of Canillo. The rest represent lesser strengths; 57%, 56% and 49% respectively for the municipalities of Santa Julia, Andorra la Vella and La Massana. It is interesting to see how the nuclei, without exception, of a specific parish gather strongly among themselves. This information is most interesting if it is taken into account that a geographic proximity is not necessary between nuclei from a same parish.

The length of the branches can be interpreted as an estimator of the differences in the mobility pattern in each parish. The length of the branches relative to the nuclei of Sant Julia de Lòria, Andorra la Vella and Canillo is very noticeable. In fact, the behaviour of these nuclei can be interpreted from a demographic point of view: in the first two, the ease of connection with communities outside the Principality and their importance as social and economic centres, have turned these nuclei into immigrant collectors, where mixed links—according to spouse's origin—are very often. The parish of Canillo differs from the rest due to a distinct reason; it is the parish which has maintained its isolation until present times and the individual interchange with other nuclei has been relatively low, maintaining a slightly different pattern of spouse's distribution. It must be remembered that these three populations are those that provide the most individuals in the total data, a situation which can accentuate the distances and length of their representative branches.

This information is also evident in the ACP graphic representation (Figure 3). We can appreciate the existence of clusters that group together the nuclei belonging to the same parish. This detail is clearly reflected in the Encamp, Ordino and San Julià de Lòria parishes. The other three parishes also show a big connection but with some exceptions. For example, in the surroundings of la Massana we can find all the nuclei belonging to this parish but Escàs and Anvòs. It is interesting to observe that Escàs tends to displace towards the Ordino cluster in the APC. It is logical if we observe the situation of this nucleus inside the valley
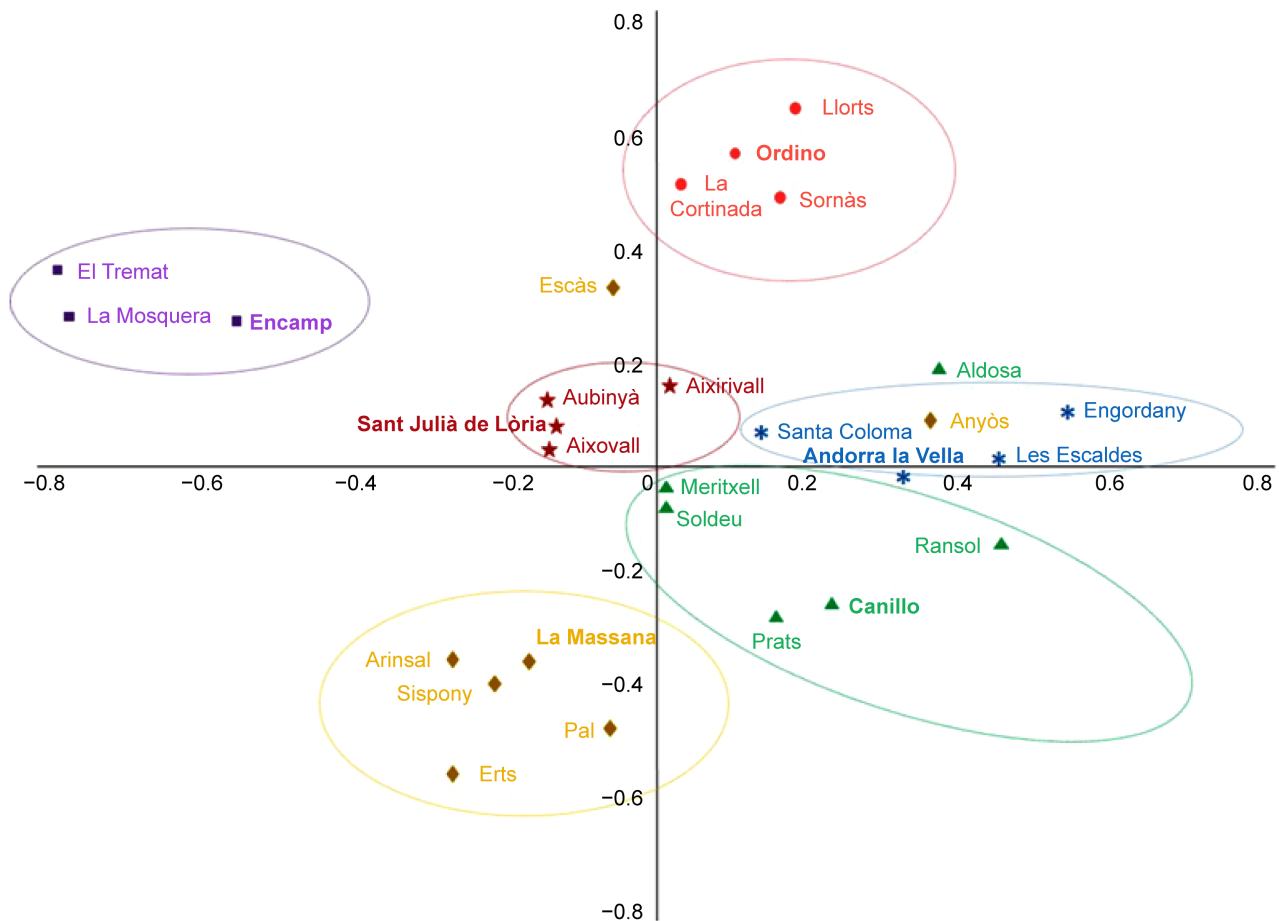
Figure 3. Principal Component Analysis (PCA) based on the intensity of the interchanges between population nuclei in the Principality of Andorra.

regarding the next valley and we take into account the natural crossing between both valleys that is near this nucleus. We can make the same reflection with the position of Anyós, that is very geographically close to the big population nucleus of the Principality, Andorra la Vella. The case of Aldosa is very similar to the Escàs one: it belongs to the Canillo valley but it clusters near the Ordino group. These two valleys are geographically limited by a mountain range but with a natural path very close to Aldosa. To sum up, the clusters exceptions seem to be related with the natural paths that connect the internal valleys of the Principality.

4. Discussion and Conclusion

Andorra is one of the smallest countries in the world. It is surrounded by three closed valleys and its inhabitants have been devoted, until recently, to agricultural exploitation. Its unique government system, nowadays substituted by a democratic regime, has kept these three valleys independent from the neighbouring countries of Spain and France. Thus, it consists of an interesting human population limited by geographical, cultural and political barriers.

Geographically, its southern sloping border is more attenuated than its northern one, and in terms of culture, the official language of Andorra is Catalonian.

Thus, Andorra has a culture a socioeconomically relation with the neighboring Catalonian regions relation. These particularities have been fundamental for determining contact with other populations: the number of individuals coming from the southern slope is almost nine times higher than that the number of individuals coming from France (González-Martín & Toja, 1998). Despite the political barriers, the migration throughout history has been very important (González-Martín & Toja, 1996) and it has occurred, preferentially, through the more accessible geographical terrain and with the most culturally similar communities. It has been reported that in Switzerland the mountains act as barriers to separate the different language groups (Rodriguez-Larralde et al., 1998). Nevertheless, in Belgium, in spite of the inexistence of geographic barriers, the isolation between the population of the same country is due to languages, emphasizing the importance of languages as genetic filters (Barrai et al., 2004). In Andorra, there exists an intermediate situation between Switzerland and Belgium, where geography and languages have an effect on genetic structure but we cannot quantify each.

Political boundaries in Andorra's case have not acted as a genetic interchange barrier but geographically and culturally imposed restrictions have influenced the direction and intensity of genetic fluxes. Furthermore, these have been significant enough as to dilute boundaries and, as a consequence, to cause a significant genetic homogeneity throughout the geographic region, above all, in the South Pyrenees valleys. The case of Andorra is not the only one where political boundaries have not acted restricting genetic flow and stopping migration: in Olivenza (Spanish city that used to be part of Portugal) migration from Portugal has also been found (Román-Busto & Fuster, 2015).

From the migration matrix structure, other common phenomena can be corroborated. For example, the custom to celebrate the wedding in the bride's parish (Toja, 1987) can directly influence on the asymmetry detected on the matrix. Nevertheless, it seems difficult to assume that the whole asymmetry can be justified by this fact, especially if male hypermobility phenomenon is taken into account. This last phenomenon is based on the greater mobility, mostly over long distances, of male individuals as compared to females. Furthermore, it is not necessarily related to marriage and it is possibly responsible for a great deal of the asymmetry detected on marital distances matrix. Asymmetry in the movements of male and female individuals has also been detected in La Cabrera (Spain) (Boattini et al., 2007) and in the Ebro delta (Esparza et al., 2006). Furthermore, in the last study, a similar percentage of marriages in which both spouses have the same origin (53.08%) as the one we obtained (49.03%) was found, and, a greater male mobility with medium and large distances was detected as well. Nevertheless, we have to take into account that the orography in Andorra and in the Ebro delta is not comparable: Andorra is situated in the Pyrenees and the Ebro delta is a plain without physical barriers.

Study of the marriage interchange patterns within the Principality reflects interesting results. Firstly, interchanges between gender and population nuclei are

produced on a balanced and bidirectional manner, although not necessarily in proportion. To explain this phenomenon, three considerations must be detailed: the greater part of the nuclei are constituted by small populations; all of them are subjected to the same ecological conditions; and that gender proportion is balanced. Thus, offer and availability of individuals for marriage are limited. Interchanges between population nuclei are compensated by two-way movements of different gender to maintain a proportional balance.

The model of matrimonial interchange, the migration models and the endogamy rates have varied over time (as certainly have the relationship between the nuclei of the populations which constitute the Principality of Andorra). Despite these variations there persists an interchange model that presents itself logically, geographically and administratively. That is to say, we can talk of a general model of genetic interchange which can be translated as a well-marked and defined genetic structure.

On the other hand, genetic interchanges between population nuclei are not produced randomly, rather, they are subjected to geographical conditions. Spouse selection is related to the distance between the places of origin of both spouses: the greater the distance, the lesser the probabilities of agreeing to a marriage. What is outstanding is that the greater distance related with matrimonial structure, and by extrapolation with genetic affinity, is the distance by road and not the actual linear distance. These roads and highways are built with a topographic logic, respecting geographical unevenness of terrain, therefore, orography is a subtle parameter regulating displacements and, indirectly, shaping marriages. Thus, it will be orographically decisive when explaining genetic fluxes and biological affinities within a population. The crucial effect of geography (and with lesser importance, orography) has also been detected in another areas of Spain like Fuentes Carrionas (Palencia) where geography was identified as the main factor of isolation among populations (Rodríguez Díaz and Blanco Villegas, 2010); in la Cabrera where the isolation (due to geography) of some parishes is the main factor that influence the genetic structure (Boattini et al., 2007) or in the Ebro delta region where the river itself acts as the main factor of isolation of the populations (Esparza et al., 2006). In the first two studies, the values obtained for the Mantel's Test relating migration matrices with geographic distance were similar to our results (0.52 for road distances and 0.48 for straight-line distances). In the study of La Cabrera (Boattini et al., 2007) the values obtained were 0.50 and 0.52 in a Mantel's test between marital migration and straight-line geographic distance. In the case of Fuentes Carrionas (also a mountainous region), Rodríguez-Díaz and Blanco-Villegas (2010) found the next values between progenitor-descendent origin and road and straight-line distances: 0.567 and 0.545 respectively. This last case concords with our results: there is a bigger relation with the road than with the straight-line distance in the case of mountainous regions. In the case of the Ebro delta region (with a different orography in respect to Andorra) the value of the Mantel's Test between the data of the marriages and the road distance was 0.848.

Associations described between population nuclei through interchange patterns, respect, with almost no exception, the administrative groupings of the Principality as shown in the Neighbor-Joining tree and in the ACP. The genetic structure is the reflex of the administrative organization of the country expressing sub-divisions where endogamy is maximum, and where micro-evolutionary phenomena will act with greater intensity.

A general view of Andorra's geography sets forth the hypothesis that administrative divisions have been set up according to the orographic pattern of the country. In fact, the country consists of three valleys; a southern one and two openings from north to south. It is permissible to suppose that geography has really determined the administrative structure to facilitate management and communication and, as such, it is truly responsible for Andorra's genetic structure and population.

References

Adellach, B., & Ganyet, R. (1977). *Estadistiques de les Valls d'Andorra*. Andorra la Vella: Consell General de les Valls d'Andorra.

Barbujani, G., & Sokal, R. R. (1990). Zones of Sharp Genetic Change in Europe Are Also Linguistic Boundaries. *Proceedings of the National Academy of Sciences of the United States of America, 87,* 1816-1819. https://doi.org/10.1073/pnas.87.5.1816

Barbujani, G., & Sokal, R. R. (1991). Genetic Population Structure of Italy. II. Physical and Cultural Barriers to Gene Flow. *American Journal of Human Genetics, 48,* 398-411.

Barrai, I., Rodriguez-Larralde, A., Manni, F., Ruggiero, V., Tartari, D., & Scapoli, C. (2004). Isolation by Language and Distance in Belgium. *Annals of Human Genetics, 68,* 1-16. https://doi.org/10.1046/j.1529-8817.2003.00044.x

Bisquerra, R. (1989). *Introducción conceptual al Análisis Multivariante.* Barcelona: Promociones y publicaciones Universitarias.

Boattini, A., Villegas, M. J. B., & Pettener, D. (2007). Genetic Structure of La Cabrera, Spain, from Surnames and Migration Matrices. *Human Biology, 79,* 649-666. https://doi.org/10.1353/hub.2008.0009

Bodmer, W. F., & Cavalli-Sforza, L. L. (1968). A Migration Matrix Model for the Study of Random Genetic Drift. *Genetics, 59,* 565-592.

Boyce, A. J., Kücheman, C. F., & Harrison, G. A. (1967). Neighbourhood Knowledge and the Distribution of Marriage Distances. *Annals of Human Genetics, 30,* 335-338. https://doi.org/10.1111/j.1469-1809.1967.tb00035.x

Calafell, F., & Hernández, M. (1993). Multivariate Approach to Matrimonial Mobility in Catalonia. *Human Biology, 65,* 731-742.

Calvo, R., Martínez, G., & Gasch, N. (1990). *RecullEstadístic General de la Població. Andorra 90*. Andorra la Vella: Conselleria de treball i benestar social.

Castles, S. (2003). *The Age of Migration: International Population Movements in the Modern World*. New York: The Guilford Press.

Cavalli-Sforza, L. L., Menozzi, P., & Piazza, A. (1994). *History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.

Coleman, D. A. (1977). The Geography of Marriage in Britain, 1920-1960. *Annals of Human Biology, 4,* 101-132. https://doi.org/10.1080/03014467700002011

Departament d'Estadística (last checked 19 April 2017). Govern d'Andorra.

http://www.estadistica.ad/serveiestudis/web/index.asp

Efron, B. (1982). *The Jackknife, the Bootstrap and the Other Resembling Plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Esparza, M., García-Moro, C., & Hernández, M. (2006). Genetic Relationships between Parishes in the Ebro Delta Region (Spain) as Estimated by Migration Matrix and Surnames. *Human Biology, 78,* 647-662. https://doi.org/10.1353/hub.2007.0010

Felsenstein, J. (1985). Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution, 39,* 783-791. https://doi.org/10.2307/2408678

Felsenstein, J. (1989). PHYLIP-Phylogeny Inference Package (Ver. 3.2). *Cladistics, 5,* 164-166.

Fuster, V. (1985). Isonymic Relationship Regarding Sex and Name Position in Spanish Community. *Collegium Antropologicum, 9,* 155-165.

González-Martín, A. (2008). Ecological and Cultural Pressure on Marriage Seasonality in the Principality of Andorra. *Journal of Biosocial Sciences, 40,* 1-18. https://doi.org/10.1017/S0021932007002234

González-Martín, A., & Toja, D. I. (1996). Geografia, Migració i Matrimoni a la Parroquia d'Ordino. *Annals de l'Institut d'Estudis Andorrans,* 103-118.

González-Martín, A., & Toja, D. I. (1998). El Estado Civil en los Consortes Andorranos: Inferencias Genéticas y Ecológicas del Análisis de la Poligamia Asincrónica. *Revista Española de AntropologíaBiológica, 19,* 29-44

González-Martín, A., & Toja, D. I. (2002). Inbreeding, Isonymy, and Kin-Structured Migration in the Principality of Andorra. *Human Biology, 74,* 587-600. https://doi.org/10.1353/hub.2002.0042

Harpending, H., & Jenkins, T. (1973). Genetic Distance among Southern Africa Populations. In M. W. Crawford, & P. L. Wokman (Eds.), *Methods and Theories and Anthropological Genetics* (pp. 177-179). Albuquerque, NM: University of New Mexico Press.

Harrison, G. A., Kücheman, C. F., Hiorns, R. W., & Carrivich, P. J. (1974). Social, Mobility, Assortative Marriage and Their Interrelationship with Marital Distance and Age in Oxford City. *Annals of Human Biology, 1,* 211-223. https://doi.org/10.1080/03014467400000221

Hartl, D., & Clark, A. G. (2007). *Principles of Population Genetics*. Sunderland, MA: Sinauer Associates.

Jobling, M. A. (2012). The Impact of Recent Events on Human Genetic Diversity. *Philosophical Transactions of the Royal Society B, 367,* 793-799. https://doi.org/10.1098/rstb.2011.0297

Kücheman, C. F., Boyce, A., & Harrison, G. A. (1967). A Demography and Genetic Study of an Oxfordshire Villages. *Human Biology, 39,* 251-276.

Li, W. H., & Graur, D. (1991). *Fundamentals of Molecular Evolution*. Sunderland, MA: Sinauer Associates.

Luchetti, E., & Soliani, L. (1989). Similarità tra Popolazioni esaminate mediante i Cognomi. *Rivista Di Antropologia, 57,* 181-198.

Mantel, N. (1967). The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research, 27,* 209-220.

McCullough, J. M. (1989). Relation of Community Size to Endogamy in a Traditional Society: Pátzcuaro, Mexico, 1903-1932. *American Journal of Human Biology, 1,* 281-287. https://doi.org/10.1002/ajhb.1310010308

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. https://doi.org/10.1111/ahg.12015

Nei, M. (1972). Genetic Distance between Populations. *American Naturalist, 106,* 283-292. https://doi.org/10.1086/282771

Nei, M. (1973). The Theory and Estimation of Genetic Distance. In N. E. Morton (Ed.), *Genetic Structure of Populations* (pp. 45-54). Honolulu, HI: University of Hawaii Press.

Pollitzer, W. S., Smith, M. T., Williams, & W. R. (1988). A Study of Isonymic Relationships in Fylingdales Parish from Marriages Records from 1654 through 1916. *Human Biology, 60,* 363-382.

Reiner, A. P., Ziv, E., Lind, D. L., Nievergelt, C. M., Schork, N. J., Cummings, S. R. et al. (2005). Population Structure, Admixture, and Aging-Related Phenotypes in African American Adults: The Cardiovascular Health Study. *American Journal of Human Genetics, 76,* 463-477. https://doi.org/10.1086/428654

Rodríguez-Díaz, R., & Blanco-Villegas, M. J. (2010). Genetic Structure of a Rural Region in Spain: Distribution of Surnames and Gene Flow. *Human Biology, 82,* 301-314. https://doi.org/10.3378/027.082.0304

Rodriguez-Larralde, A., Scapoli, C., Beretta, M., Nesti, C., Mamolini, E., & Barrai, I. (1998). Isonymy and the Genetic Structure of Switzerland. II. Isolation by Distance. *Annals of Human Biology, 25,* 533-540. https://doi.org/10.1080/03014469800006772

Román-Busto, J., & Fuster, V. (2015). Secular Trends in the Relationship between Surnames in a Population: Study of a Border Town. *HOMO: Journal of Comparative Human Biology, 66,* 158-164.

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., & Feldman, M. W. (2002). Genetic Structure of Human Populations. *Science, 298,* 2381-2385. https://doi.org/10.1126/science.1078311

Saitou, N., & Nei, M. (1987). The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular Biology and Evolution, 4,* 406-425.

Sánchez-Mazas, A., & Barbujabi, G. (2013). European Genetic Diversity through Space and Time. *Human Heredity, 76,* 119-120. https://doi.org/10.1159/000362130

Segalen, M., & Jacquard, A. (1973). Isolement Sociologique et Isolement Genetic. *Population, 3,* 487-498.

Slatkin, M. (1987). Gene Flow and the Geographic Structure of Natural Populations. *Science, 236,* 787-792. https://doi.org/10.1126/science.3576198

Smith, M. T., & Hudson, B. L. (1984). Isonymic Relationships in the Parish of Fylingdales, North Yorkshire, in 1851. *Annals of Human Biology, 11,* 141-148. https://doi.org/10.1080/03014468400006991

Smith, M. T., Smith, B. L., & Williams, W. R. (1984). Changing Isonymic Relationships in Fylingdales Parish, North Yorkshire, 1841-1881. *Annals of Human Biology, 11,* 449-457. https://doi.org/10.1080/03014468400007341

Susanne, C. H. (1983). Neighbourhood Knowledge, Distance at Marriage and Migration. *Garcia de Orta, Série de Antropología, 2,* 41-48.

Tishkoff, S. A., Pakstis, A. J., Stoneking, M., Kidd, J. R., Destro-Bisol, G., Sanjantila, A. et al. (2000). Short Tandem-Repeat Polymorphism/*Alu* Haplotype Variation at the PLAT Locus: Implications for Modern Human Origins. *American Journal of Human Genetics, 67,* 901-925. https://doi.org/10.1086/303068

Toja, D. I. (1987). *Estructura Matrimonial de las Poblaciones de dos Valles del Pirineo.* Ph.D. Thesis, Barcelona: University of Barcelona.

Wahlund, S. (1928). Zusammensetzung von Populationen und Korrelationserscheinun-

gen vom Standpunkt der Vererbungslehre aus Betrachtet. *Hereditas, 11,* 65-106.
https://doi.org/10.1111/j.1601-5223.1928.tb02483.x

Wright, S. (1951). The Genetical Structure of Populations. *Annals of Eugenics, 15,* 322-354.

Ziv, E., & Burchard, E. G. (2003). Human Population Structure and Genetic Association Studies. *Pharmacogenomics, 4,* 431-441. https://doi.org/10.1517/phgs.4.4.431.22758