



ORIGINAL PAPER

Genetic diversity in Northern Spain (Basque Country and Cantabria): GM and KM variation related to demographic histories

E Esteban¹, JM Dugoujon², E Guitard², MT Sénégas², C Manzano³, C de la Rúa³, N Valveny¹ and P Moral¹

¹Unitat d'Antropologia, Departament de Biologia Animal, Facultat de Biologia, Universitat de Barcelona, Spain

²Centre d'Immunopathologie et de Génétique Humaine, CNRS, CHU de Purpan, Toulouse, France

³Departamento de Biología Animal y Genética, Facultad de Ciencias, Universidad del País Vasco, Bilbao, Spain

Genetic diversity in Northern Spain (SW Europe) was assessed through the analysis of the GM and KM immunoglobulin markers in 505 individuals using a set of 17 allotypes, including the G2M(23) allotype which has been infrequently used before now. The individuals were representative of three anthropologically well-defined populations belonging to two geographically and archaeologically distinct areas in the Basque Country (Guipúzcoa and Alava provinces) and to the mountainous region of Montes de Pas in the province of Cantabria. Gene frequency distributions indicated a high genetic divergence between Montes de Pas and the Basque Country, and a relative degree of heterogeneity between the two Basque regions. The genetic differentiation of Montes de Pas, which is consistent with previous classical polymorphism analyses, suggests a considerable genetic variation range within the Iberian Peninsula, possibly higher than that often polarised around the Basque versus non-Basque variation. Analyses of genetic structure show that the major differentiation of Montes de Pas could be related to the historically documented mixed origin of this population. The moderate genetic distances between regions in the Spanish Basque Country could be explained by differential systematic pressures acting through a stronger gene flow in the South than in the more isolated Northern areas. The comparisons with neighbouring populations from the French Pyrenees suggest that the present genetic variation revealed by Ig polymorphisms in SW Europe can be related to historical demographic processes including gene flow and/or low population sizes.

Keywords: Immunoglobulins; GM; KM; Basques; Cantabria

Correspondence: Pedro Moral, Unitat d'Antropologia, Departament de Biologia Animal, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, 08028-Barcelona, Spain. Tel: 34 3 402 14 61; Fax: 34 3 411 08 87; E-mail: pmoral@porthos.bio.ub.es

Received 13 January 1997; revised 9 December 1997; accepted 23 December 1997

Introduction

GM and KM immunoglobulin (Ig) antigens, along with HLA system, are the classical polymorphisms showing the greatest variation in humans. The GM system is defined by a set of antigens (allotypes) on the Ig heavy chains of IgG1, IgG2 and IgG3 subclasses. Since the genes encoding these subclasses are closely linked on chromosome 14,^{1,2} the different allotypes of each subclass are inherited in fixed combinations or haplotypes. The KM allotypes located on the constant part of Ig kappa light chains are independently inherited by the *Kappa* gene on chromosome 2.³ Whereas the variation in the KM system is only noticeable between the main ethnic groups, GM haplotypes have been shown to be a powerful tool in human population genetics due to their high frequency variation between populations. Furthermore, since several haplotypes are specific to certain groups^{4,5} they allow the possibility to evaluate gene flow and admixtures.

Despite the outstanding potential of GM polymorphism, in practice its usefulness to human population studies have been impoverished by the irregular distribution of available data and by the heterogeneity of allotypes tested in the different groups, which poses a problem for comparison purposes. In Europe, for instance, GM distribution has been extensively analysed, but not all different areas are uniformly represented: most surveys did not test for the interesting G2M(23) allotype that enables discrimination of the most frequent haplotype *GM*3;5**, 26,27 between *GM*3;.,5*,26,27* and *GM*3; 23; 5*,26,27*. Likewise, through the determination of the G3M(6,15,16,24), African haplotypes can be accurately distinguished from those originating in western Asia.

This paper deals with GM and KM diversity in three Spanish populations. The study used a complete panel of allotypes. Although studies of GM in Spanish populations have been reported,⁶⁻⁹ the lack of informative allotype determinations represents a remarkable gap in information. This study analyses three Spanish samples, coming from the Guipúzcoa and Alava provinces of the Basque Country, and from Montes de Pas in the Cantabria region (Figure 1).

The anthropological and geographic features of the populations examined bring additional interest to this study. The first point refers to the geographic position of the Basque samples in the south-western Pyrenees, the geographical and political border which separates the Iberian Peninsula from the rest of mainland Europe. The genetic landscape of the Pyrenees has

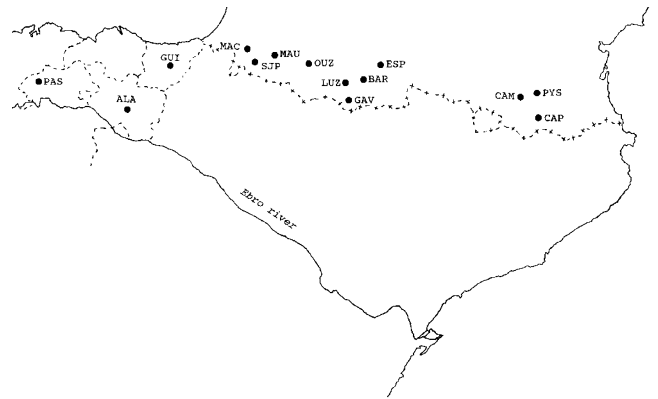


Figure 1 Geographical location of the Spanish and Pyrenean populations. Population abbreviations: PAS: Montes de Pas; ALA: Alava; GUI: Guipúzcoa; CAP: Capcir; PYS: Pays de Sault; CAM: Camurac; ESP: Esparros; BAR: Barèges; GAV: Gavarnie; LUZ: Luz-St Sauveur; OUZ: Vallée de l'Ouzon; MAU: Mauléon; SJP: St Jean de Pied de Port; MAC: Macaye.

been for a long time of interest to anthropologists trying to assess the genetic trail of possible east-west and north-south population interchanges. Some studies, using other classical markers such as blood groups, proteins and HLA antigens, have suggested the existence of an east-west gradient with the genetic distinctiveness of the Basque population at the western extreme.¹⁰ However, Ig data reported only for the northern side of the Pyrenees^{11,12} seem to suggest greater genetic proximity between the western and central Pyrenees, as against the easternmost part. These first data about the southern side of the Pyrenees will enable the genetic relationships between Pyrenean populations to be analysed better.

Concerning the Basque samples, Guipúzcoa, Alava and Vizcaya are the component provinces of the autonomous region of the Basque Country in northern Spain. There are also Basques in Navarra (Spain) and in the Département des Pyrénées-Atlantiques in south-western France. The Basques are ethnically distinct from surrounding peoples. Many studies have established their genetic and linguistic peculiarities, considering the present-day Basques a relic of the ancient population of Europe [see as a review Cavalli-Sforza *et al*¹³]. This peculiarity has been preserved more fully in Spanish Basques than in French Basques, who have probably had greater influence from abroad. In spite of the numerous surveys on the genetic characteristics of the Spanish Basques, no data on Ig markers have yet been published. Another question is the homogeneity or heterogeneity of the Spanish Basque population.

Several studies^{14,15} have suggested a certain degree of genetic heterogeneity, the main subdivision being between North (Guipúzcoa and Vizcaya provinces) and South (Alava province). This differentiation agrees with the geographical subdivision established by the mountain range that constitutes the watershed, and is consistent with other archaeological,¹⁶ ethnological¹⁷ and skeletal¹⁸ differentials. The Ig genetic markers of clearly separated geographical areas of the Spanish Basque Country will enable the hypothesis of Manzano *et al*⁹ which distinguishes the 'Atlantic' Spanish Basques in the North (Guipúzcoa and Vizcaya provinces) from the southern group with more 'Mediterranean' influence (Alava province) to be tested.

The study of a particular population, geographically near the Basque Country but non-Basque, from the Montes de Pas can serve not only to analyse the genetic relationships of Basques with surrounding populations, but also to explore the unique characteristics of this population. Montes de Pas is a rough region encompassing several small and deep-set valleys in the south of the Cantabrian Mountains, between Cantabria and Burgos provinces (North Spain). The human presence in this region could be very ancient due to the existence of natural passes between these two provinces. The oldest historical evidence supports the passage of Romans, and later Visigoths, in order to conquer Cantabria. However, human settlement was probably not constant until the beginning of the 11th century when the award of pastures was historically documented. At this time, repopulation from neighbouring areas took place, perhaps with small contributions from ethnic minorities. Since the 18th century, the population of the region has fluctuated at around 5000 inhabitants. People from Montes de Pas have a different life style from other people in the Iberian Peninsula. The 'Pasiegos' are shepherds who move seasonally between different cow pastures they own along the mountains. They possess a house in each pasture, so farmhouses are scattered over the mountainous slopes. Their extreme adaptation to the environment has produced a relatively isolated population with little interaction with other parts of the province (83.7% of endogamy and 6.1×10^{-3} of consanguinity²⁰); as a result of this, legends about their possibly Moorish or Jewish origins have arisen. In contrast with the Basques, the human population genetics of Montes de Pas are poorly known. At present only data on several protein polymorphisms²¹ are available. The study of Ig genetic variants presented here will provide new information

on the genetic structure of this population, its affinities with other neighbouring groups and the patterns of genetic diversity among Iberian Peninsula populations.

Materials and Methods

A total of 503 individuals coming from three Northern Spanish regions, Guipúzcoa ($n = 112$), Alava ($n = 294$), and Montes de Pas ($n = 100$), were studied. Blood specimens were collected from adult, unrelated, apparently healthy people of both sexes living in rural areas, and only individuals whose four grandparents were natives of the same region were considered. The Basque samples from Guipúzcoa and Alava were taken from autochthonous individuals of the two provinces. Guipúzcoa, is the smallest of the Basque provinces (1997 square km, and 705 933 inhabitants), whose central position has enabled it to preserve its language and ancient traditions more fully than the other two provinces. To the south of the Pyrenees, the province of Alava, with an area of 3074 square km and 260 580 inhabitants, is open to the Ebro basin. The Montes de Pas sample comes mainly from hamlets scattered in the main valleys of this mountainous region of around 5000 inhabitants.

GM and KM Typing

Serum samples were typed for 16 GM allotypes belonging to the three IgG subclasses G1M(1,2,3,17), G2M(23), G3M(5,6,10,11,13,14,15,16,21,24,28) and for the KM(1) allotype, according to the standard agglutination inhibition method.²² The reagents used are listed in Table 1.

Statistical Analysis

Haplotype frequencies were estimated using the maximum likelihood method.²³ Goodness of fit to the Hardy-Weinberg equilibrium and allele frequency differences between populations were evaluated by means of the chi-square test. Heterozygosities were estimated as $1 - \sum x_i^2$, where x_i is the estimated frequency of the i -allele in each system.

For comparative purposes, data on other Pyrenean populations were selected from the literature.¹¹ Their geographic locations are shown in Figure 1. To quantify the degree of genetic diversity among populations, F_{ST} statistic²⁴ was used as a measure of the proportion of genetic variance attributable to population subdivision. The significance of F_{ST} was determined by Workman and Niswander's²⁵ method. Genetic distances between pairs of populations were calculated by GENDIST program in PHYLIP 3.5c package²⁶ according to the methods of Cavalli-Sforza and Edwards,²⁷ Reynolds *et al*²⁸ and Nei.²⁹ Reynolds' distances were selected because these measurements are proportional to the time of population separation if the main differentiation agent is genetic drift, as seems to be the case for human populations at a microgeographical level. Standard errors of the distances were computed by the bootstrap resampling method.³⁰ A graphical representation was carried out through principal-component analysis of a variance-covariance matrix from original data

Table 1 Reagents used for GM and KM allotype determination

System	Chain	Allotype	Anti-allotype	Coating antigen (anti-D)	
GM	1	G1M(1)	Don	Gan	
		G1M(2)	Sub	Gan	
		G1M(3)	Rum	Bei	
	2	G1M(17)	Seg	Gan	
		G2M(23)	SH-21 ¹	Del	
		3	G3M(5)	Pag	Eyc
			G3M(6)	Ser	2781 ²
			G3M(10)	Des	Eyc
	3	G3M(11)	Feu	Eyc	
		G3M(13)	Bet	Eyc	
		G3M(14)	Bla	Eyc	
		G3M(15)	2624 ²	4095 ²	
		G3M(16)	2639 ²	4095 ²	
		G3M(21)	Des	Ant	
		G3M(24)	D2018	2781 ²	
		G3M(28)	Lla	Ant	
	KM		KM(1)	Por	Bei

Anti-allotypes and Anti-D were provided by the Centre Régional de Transfusion Sanguine, d'Hématologie, Immunologie et Génétique Humaine, CHU de Purpan, Toulouse (France), except for: ¹provided by Sigma-Aldrich Quimica SA; ²provided by the CLB Central Laboratory of the Netherlands Red Cross Blood Transfusion Service, Amsterdam (Holland). Gift from Dr G de Lange.

(allele frequencies and populations) according to Harpending & Jenkins.³¹

Results

Phenotype and Haplotype Distributions

GM and KM phenotype distributions are shown in Table 2. Some uncommon phenotypes, all of them previously reported^{32-34,11} [GM(1,2,3,17,23;5*), GM(3;23;5*,21,28) and GM(1,3;...;5*,21,28) in Alava; GM(1,3,17;23;21,28) and GM(1,3,17;23;10,11,13,15,16) in Montes de Pas], were excluded from haplotype calculations, except for the carriers of haplotype GM*1,17,23;21,28, one of the most frequent among the rare haplotypes. In all cases, phenotype distributions showed no significant deviations from the Hardy-Weimberg equilibrium. Table 3 reports the GM and KM haplotype frequencies with the estimated confidence intervals.

Heterogeneity for the distribution of GM was observed among the three populations examined ($\chi^2_8 = 17.458$, $P = 0.025$). Pairwise comparisons also revealed significant differences (Table 4) for several GM haplotypes. Unlike for GM, similar frequencies for

KM alleles were found among the three populations (Table 3).

Spanish Haplotype Frequencies within the European Variation

In relation to GM and KM variation in Europe (Table 5), the most remarkable findings were the high GM*3;23;5* frequency of Guipúzcoa and the low GM*3;5* (GM*3;...;5* and GM*3;23;5* combined) value in Montes de Pas. Also the GM*1,17;...;21,28 frequency in Montes de Pas was relatively high in comparison with both Europe and North Africa.

The GM*1,17;...;5*, characteristic of black Africa⁵ and also frequent in North Africa, is present in Alava and Montes de Pas but absent in Guipúzcoa. The GM*1,17;...;10,11,13,15,16, particularly common in Asian populations⁵ but rare or absent in most of North and Central Europe, was found in the three populations examined, the Guipúzcoa and Montes de Pas values lying among the highest so far described in Europe and North Africa (Table 5).

Table 2 GM and KM phenotypes distribution, Hardy-Weimberg equilibrium and heterozygosities in the Guipúzcoa, Alava and Montes de Pas populations

Phenotypes	Guipúzcoa	Alava	Mt. de Pas
<i>GM</i>			
3;23;5*	46	98	27
3;...;5*	2	10	3
1,17;...;21,28	8	24	9
1,2,17;...;21,28	1	10	2
1,17;...;5* 21,28	0	3	1
1,3,17;23;5* 21,28	29	66	22
1,3,17;...;5* 21,28	12	38	17
1,2,3,17;23;5* 21,28	5	20	3
1,2,3,17;...;5* 21,28	4	12	2
1,3,17;23;5*	0	3	3
1,3,17;...;5*	0	2	0
1,2,17;23;21,28	0	1	0
1,17;23;21,28	2	1	0
1,3,17;23;5* 15,16	0	2	1
1,17;...;10,11,13,15,16,21,28	0	1	3
1,17;...;5*	0	2	0
1,3;23;5*	0	1	0
1,3,17;...;5* 15,16	2	0	5
N	111	294	98
$\frac{df}{P}$	0.1101 ¹⁰ _{0.999}	3.3013 ¹¹ _{0.986}	1.5063 ⁹ _{0.997}
Heterozygosity	0.669	0.710	0.723
<i>KM</i>			
KM(1)	12	36	14
KM(-1)	99	258	86
N	111	294	100

5* = 5,10,11,13,14.

df degree of freedom.

P probability.

Table 3 GM haplotype and KM frequencies (confidence intervals in parentheses)

GM haplotypes	Guipúzcoa	Alava	Montes de Pas
1,17;...;21,28	0.2567 (0.196–0.326)	0.2840 (0.243–0.323)	0.3227 (0.252–0.392)
1,2,17;...;21,28	0.0446 (0.022–0.078)	0.0749 (0.054–0.104)	0.0359 (0.015–0.070)
3;...;5*	0.1867 (0.126–0.256)	0.2072 (0.167–0.257)	0.2298 (0.159–0.309)
3;23;5*	0.4770 (0.396–0.556)	0.3996 (0.349–0.449)	0.3409 (0.270–0.420)
1,17;...;5*	–	0.0254 (0.014–0.041)	0.0253 (0.008–0.055)
1,17;...;10,11,13,15,16	0.0177 (0.005–0.042)	0.0017 (0.0002–0.0078)	0.0455 (0.021–0.082)
1,17;23;21,28	0.0173 (0.002–0.052)	0.0056 (0.001–0.018)	–
1,3;±23;5*	–	0.0017 (0.0002–0.0078)	–
KM			
KM*1+KM*1,2	0.0556	0.0632	0.0726
KM*3	0.9444	0.9368	0.9274

5* = 5,10,11,13,14.

Table 4 Significant differences for several GM haplotypes in the three populations studied

Haplotype	Population comparisons	
Common GM haplotypes	Montes de Pas–Guipúzcoa	$\chi^2_4 = 9.884, P = 0.042$
	Montes de Pas–Alava	$\chi^2_4 = 10.497, P = 0.032$
GM*1,17;...;5*	Guipúzcoa–Alava	$P = 0.007$
	Guipúzcoa–Montes de Pas	$P = 0.022$
GM*1,17;...;10,11,13,15,16	Alava–Guipúzcoa	$P = 0.032$
	Alava–Montes de Pas	$P < 0.001$

degree of freedom.

In italics, Fisher exact probability.

Table 5 Variation ranges for GM and KM systems in European and North African populations

GM haplotypes	Europe			North Africa		
	<i>n</i> ¹	Range	Extreme samples	<i>n</i>	Range	Extreme samples
1,17;...;21,28	34	0.088–0.349	Croatia/FrenchPyr.	7	0.206–0.308	Kesra B. ³ /Douiret B.
1,2,17;...;21,28	34	0.007–0.147	Sardinia/Norway	7	0.000–0.051	Several B./Tunisia G. ⁴
3;5*	34	0.554–0.854	FrenchPyr./Sardinia	7	0.400–0.718	Takrouna B./Kesra B.
3;...;5*	11 ²	0.167–0.308	Camurac/P. Sault	6	0.085–0.303	Tunisia G./Douiret B.
3;23.5*	11	0.295–0.474	V. Ouzon/Barèges	6	0.274–0.535	Takrouna B./Kesra B.
1,17;...;5*	34	0.000–0.041	Several/French Pyr.	7	0.063–0.266	Kesra B./Takrouna B.
1,17;...;10,11,13,15,16	34	0.000–0.005	Several/French Pyr.	7	0.000–0.007	Several B./Tunisia G.
1,17;23;21,28	11	0.000–0.022	Several/LuzSt. Sav.	6	0.000–0.007	Several B./Tunisia G.
KM system						
KM*1+KM*1,2	34	0.033–0.100	Austria/Sardinia	7	0.121–0.205	Kesra B./Douiret B.

¹*n* is the number of studies reviewed. The sources of data are the general collections of Steinberg & Cook [5], completed, for particular cases, with data from Blanc & Ducos [32], Dougoujon *et al.* [11], Pizza *et al.* [33], Lefranc *et al.* [34], Chaabani *et al.* [35], Chaabani & Cox [36], Helal *et al.* [37] and Dugoujon *et al.* [38].

²Only data from French Pyrenean populations were available.

³B, Berber population from Tunisia.

⁴G, General non-Berber population from Tunisia.

As for KM allotypes, the observed frequencies $KM^*1 + KM^*1,2$ in the Spanish samples range within the variation described in Europe and are noticeably lower than in North Africa (Table 5).

Comparisons between Populations from Pyrenees

The genetic relationships between the three Spanish populations were assessed in comparison with other historic and geographically related groups from the French Pyrenees^{11,12}. A first approach to interpopulation heterogeneity was tested by contingency chi-square test and Wright's F statistics. The chi-square analysis indicates remarkable divergences between the three Spanish samples and the eleven French Pyrenean groups, Montes de Pas being the most differentiated (all comparisons were significant), followed by Guipúzcoa (5 out of 11 values were significant) and Alava (4 out of 11 differences were significant). For KM, none of the 33 comparisons was statistically significant. The total number of significant differences (20 out of 66 comparisons) is clearly higher than that expected by pure chance (3.3, at a significant level of 0.05).

F_{ST} values for interpopulation differentiation are given in Table 6. The F_{ST} for GM reveals a significant (at the 0.001 level) degree of variation across the 14 populations considered, and also across the more strictly Pyrenean groups, if the geographically separate population of Montes de Pas is excluded. These values are similar in magnitude to values found among Mediterranean regions for 15 genetic markers.³⁹ The F_{ST} values obtained from KM were substantially smaller than those for GM. On trying to find a geographic pattern of gene variation, a hierarchical F_{ST} analysis of the Pyrenean populations grouped into three main regions (Western, Central and Eastern Pyrenees, Table 6) indicated that there was no differentiation among regions, the diversity within regions

being equivalent to the total interpopulation divergence (Table 6). F_{ST} values for separate regions showed significant genetic heterogeneity for Central and Eastern regions, in contrast to the homogeneity present in the Western group composed of 5 Basque populations (three from France and two from Spain).

Pairwise genetic distances, based on GM and KM gene frequencies, between the three Spanish samples and 11 French Pyrenean populations are shown in Table 7, along with their bootstrap errors indicating that most of the distance estimates (83.5%) are significantly different from zero. Among Spanish groups, the largest differentiation corresponded to the Montes de Pas-Guipúzcoa comparison (genetic distance = 16.3×10^{-3}), whereas Alava Basques were equidistant from those populations with a three-fold smaller distance. An analysis of the intra-Spanish distances inside the overall Pyrenean variation [distance range: 1.1–34.1 ($\times 10^{-3}$)] shows that: i) 67% of the values in Table 7 are low (lesser than 10×10^{-3}), pointing to overall genetic homogeneity; ii) within the remaining largest distances, most of them (87%) correspond to the comparisons involving Vallée de l'Ouzon, Pays de Sault and Montes de Pas, which appear as clear outliers in a neighbour-joining tree (data not shown).

Figure 2 displays a principal-component analysis map from both populations and alleles. The first three scaled eigenvectors account for 32.9%, 23.3%, and 18.3%, respectively (74.4% accumulated) of the total variance. In the population plot on the first two axes (Figure 2a), the first component differentiates Montes de Pays from Pyrenean populations, in line with its geographic location and its particular origin. The second axis reflects a differentiation within the Pyrenees with a peripheral position of several groups such as Vallée de l'Ouzon and Pays de Sault on the one hand, and Guipúzcoa Basques and Luz-St Sauveur on the other.

Table 6 Gene diversity analysis in Pyrenean populations

Locus	Hierarchical F -statistics			Non-hierarchical F -statistics			
	¹ F_{PR}	² F_{RT}	³ F_{PT}	West ⁺ (n = 5)	Central ⁺ (n = 5)	East ⁺ (n = 3)	All populat. ⁺ (n = 14)
GM	0.007	0.000	0.007*	0.000	0.008*	0.007*	0.007*
KM	0.002	0.000	0.001	0.003	0.000	0.000	0.001
Average	0.006	0.000	0.006*	0.001	0.007*	0.006*	0.006

¹ F_{PR} : gene diversity among populations (P) within regions (R).

² F_{RT} : gene diversity among regions.

³ F_{PT} : total gene diversity among populations.

*Significance with $P < 0.001$.

⁺Populations: West (Macaye, Mauleon, St Jean de Pied de Port, Alava, Guipúzcoa); Central (Esparros, Vallée de l'Ouzon, Luz-St Sauveur); East (Capcir, Pays de Sault, Camurac). In All populations, Montes de Pas has also been included.

Table 7 Reynolds genetic distance (1000) below diagonal. Standard errors (1000) above diagonal. Population abbreviations defined as in Figure 1

	CAP	PYS	CAM	ESP	OUZ	LUZ	GAV	BAR	MAC	MAU	SJP	GUI	ALA	PAS
CAP		3.6	1.1	1.2	5.4	4.7	1.0	2.7	2.0	1.3	2.3	3.6	1.3	4.1
PYS	9.7		10.0	4.7	2.6	9.1	8.6	10.1	6.1	3.8	5.8	9.0	4.9	3.0
CAM	3.6	20.2		1.0	10.2	2.8	0.7	1.5	3.9	2.3	1.2	2.1	1.3	4.7
ESP	1.6	10.1	2.5		5.8	3.5	1.5	1.0	1.7	0.8	1.1	2.2	0.4	2.3
OUZ	15.3	8.5	24.6	16.1		4.8	4.5	12.2	12.8	8.4	10.0	11.8	3.8	2.2
LUZ	9.8	17.7	7.9	7.4	10.5		2.6	2.9	8.2	5.4	3.2	3.5	1.6	2.6
GAV	2.7	15.1	2.2	2.9	17.5	5.9		3.6	2.8	2.0	3.0	4.0	1.4	4.4
BAR	7.7	21.9	1.6	3.6	27.6	8.2	5.2		3.9	1.7	1.4	1.1	2.3	7.2
MAC	8.2	14.2	9.5	7.0	34.1	23.8	11.4	9.4		0.4	4.0	2.5	2.5	6.6
MAU	2.9	9.1	4.5	1.9	21.7	13.3	5.0	5.3	1.8		2.0	1.2	1.1	3.7
SJP	4.0	16.5	2.3	2.5	22.6	9.3	6.4	2.9	8.2	4.0		2.0	1.4	4.8
GUI	8.0	20.7	2.9	4.7	30.2	11.1	6.6	1.3	6.3	4.0	2.9		2.3	6.9
ALA	2.7	8.4	3.7	1.1	11.3	3.9	2.0	4.6	9.8	3.3	4.5	5.7		1.8
PAS	11.1	8.1	15.5	8.7	4.9	5.9	11.7	15.2	23.0	13.0	13.4	16.3	5.4	

This second component also reveals the genetic divergence between the two Spanish Basque groups, Alava lying closer to the average population variation. The particular contribution of each allele to the variation along the first two axes (Figure 2b) identifies $GM^*1,17; \dots; 10,11,13,15,16$ and $GM^*1,17; \dots; 21,28$ alleles as highly associated with Montes de Pas divergence,

and $GM^*1,17; 23; 21,28$, $GM^*3; 23; 5^*$ and $GM^*3; \dots; 5^*$ as mainly correlated with within-Pyrenees differentiation.

Discussion

The analysis of the Ig haplotypes discussed below is based on the most complete allotype information tested to date on Spanish populations, particularly with the determination of the G2M(23) allotype that enables the frequent European haplotype $GM^*3; 5^*$ to be subdivided. This subdivision may be useful for detecting underlying genetic variation at a local level, as demonstrated by the absence of significance in the Montes de Pas–Guipúzcoa comparison when their haplotypes were joined ($\chi^2_3 = 5.75$, $P = 0.124$) and the important role of these haplotypes for detecting genetic differentiation among Pyrenean populations, as seen in Figure 2.

Concerning the Ig gene distribution in Northern Spain, the major findings of this study are the high genetic divergence between Montes de Pas and Basques and a relative degree of differentiation between the two Basque groups. At first glance, the major differentiation of Montes de Pas could be thought to be due to the genetic peculiarity of Basque people, but comparisons with other Spanish groups (unpublished data), the Pyrenees and other continental populations emphasise that the outlying features of Montes de Pas, rather than Basque peculiarity, are responsible for such divergence. In fact, Montes de Pas has more extreme GM frequencies than Basques for $GM^*1,17; 21,28$ (near to maximum), $GM^*3; 5^*$ (near to minimum) and $GM^*1,17; \dots; 10,11,13,15,16$ (maximum)

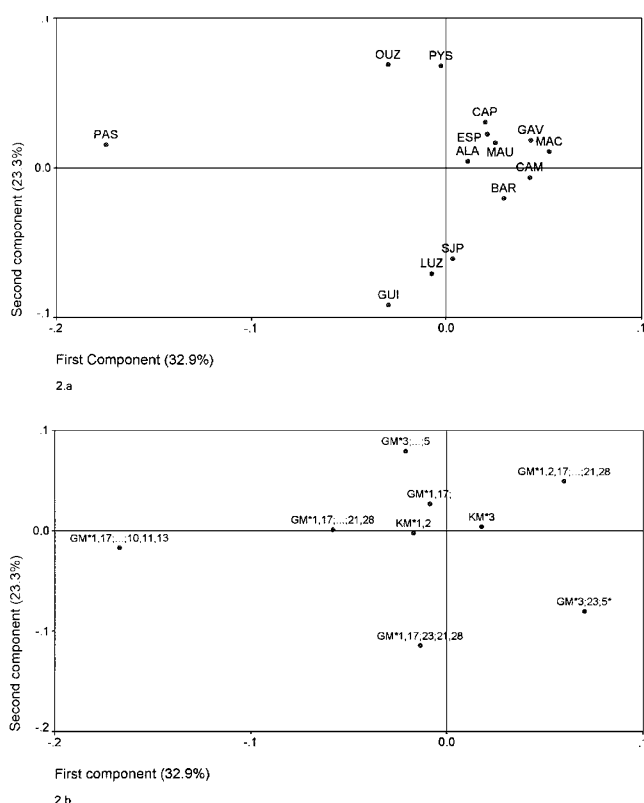


Figure 2 Populations (2a) and haplotypes (2b) plotted on first two scaled eigenvectors.

within European haplotype ranges (see Table 5). This local variation is discussed in relation to the above issues in the Introduction.

This new information on Ig markers fails to demonstrate that the significant genetic variation in the Pyrenean area could be associated with any particular geographic pattern. The higher intra- than inter-region GM genetic diversity observed is not consistent with the subdivision into three geographic regions. Besides, the facts that the gene differentiation within the Central and the Eastern Pyrenees is as high as in the whole area, and that the lower between-region average distance corresponds to the comparison between the two extremes (West–East), do not support any W–E gradient for GM haplotype Pyrenean variation. Through a different analysis our results agree with those of Dugoujon *et al*¹¹ and Hazout *et al*,¹² identifying an overall genetic homogeneity among Pyrenean populations with several outliers irregularly distributed. These particular groups, represented mainly by Vallée de l'Ouzon and to a lesser degree by Pays de Sault and Luz-St Sauveur, account for most overall diversity because when they were discarded no heterogeneity was apparent ($F_{ST} = 0.001$, for both GM and KM systems). It seems as if the Ig variation pattern in the Pyrenees should be related to local microdifferentiation processes reflecting unique historical events rather than ancient general settlement patterns. Since there is no reason to suspect heterogeneity in selection pressures in this restricted area, genetic drift and gene flow are the presumed agents responsible for the observed GM haplotype differentiation.

In spite of the general homogeneity of the Basque groups in the context of Pyrenean populations, they show a certain degree of differentiation that can be related to geography and population dynamics. The GM haplotype distribution in Guipúzcoa denotes a relatively specific position for this population (see Figure 2) which appears characterised by the highest $GM^*3;23;5^*$, $GM^*1,17;.;10,11,13,15,16$ and $GM^*1,17;23;21,28$ frequencies and the lowest value for $GM^*3;.;5^*$ among Basque groups. This relative divergence of Guipúzcoa could be related to a higher degree of isolation than the Alava population and is consistent with the information from other markers showing a certain genetic differentiation among Spanish Basque regions according to geographical, cultural and anthropometric variation.^{15,19} The most striking differences from Alava are a significant absence of $GM^*1,17;.;5^*$ and a higher presence of $GM^*1,17;.;10,11,13,15,16$

haplotype. If the presence of African $GM^*1,17;.;5^*$ haplotype in the Iberian Peninsula is presumably related to the Moslem invasion in 8th century, then its absence in Guipúzcoa is consistent with historical records indicating that Muslims did not conquer this region, as well as a general isolation from southern population influences. More difficult to explain is the frequency of $GM^*1,17;.;10,11,13,15,16$ which has an assumed west Asian origin and could have followed various paths into European populations. Its presence in Guipúzcoa may reflect a northern influence, perhaps very ancient. As for the recombinant haplotype $GM^*1,17;23;21,28$ described as associated to a genetic anomaly, it can be found scattered in French populations and no relevant information can be gained from its presence in the two Spanish Basque samples. But the greater number GM haplotypes found in Alava support more external influences, in line with its geographical location near the Ebro river valley, a communication path for Mediterranean peoples since prehistoric times.

The most unusual distribution in the Ig analysis corresponds to the Montes de Pas population, which shows extreme frequencies for some common and uncommon GM haplotypes within the European variation. In particular, concerning other south-west European groups, the Ig haplotype distribution of Montes de Pas is characterised by a relatively high incidence of $GM^*1,17;.;5^*$ and $GM^*1,17;.;10,11,13,15,16$ haplotypes. According to historical information, the origin of the present population could be traced back to the 11th century with the resettlement of people from different regions, including Moslem slaves, giving rise later to an endogamous isolated population. Although more robust conclusions remain to be drawn from consideration of many more loci, our analysis of the GM and KM genetic regions shows that instances such as the Montes de Pas population increase the range of genetic variation in the Iberian Peninsula and illustrates how current local genetic variation can be related, at least in some cases, to quite recent historic and demographic processes.

Acknowledgements

This work was supported in part by D.G.I.C.Y.T. PB95-0277-C02-02 and Agencia Española de Cooperación Internacional grants, the Centre National de la Recherche

Scientifique, and by a Commissionat per a Universitats i Recerca de la Generalitat de Catalunya grants to E.E. and N.V. Thanks to John Clayton and André Sevin (CIGH, CNRS) for their help in the estimation of haplotype frequencies.

References

- 1 Croce CM, Shander M, Martinis J *et al*: Chromosomal location of the genes for human immunoglobulin heavy chains. *Proc Natl Acad Sci USA* 1979; **76**: 3416–3419.
- 2 Kirsch IR, Morton CC, Nakahara K, Leder P: Human immunoglobulin heavy chain genes map to a region of translocations in malignant B lymphocytes. *Science* 1982; **216**: 301–303.
- 3 McBride OW, Hieter PA, Hollis GF, Swan D, Otey MC, Leder P: Chromosomal location of human kappa and lambda immunoglobulin light chain constant region genes. *J Exp Med* 1982; **155**: 1480–1490.
- 4 Schanfield MS, van Loghem E: Human immunoglobulin allotypes. In: Weir DM (ed) *Genetics and Molecular Immunology. Handbook of Experimental Immunology*, Blackwell Scientific Publications, 1986, vol. 3, pp 1–18.
- 5 Steinberg AG, Cook CE: *The Distribution of Human Immunoglobulin Allotypes*. Oxford University Press, 1981.
- 6 Schanfield MS, Baylerian R, Maiquez J, Carbonell F: Immunoglobulin allotypes in European populations. IV. Gm, Am and Km allotypic markers in Valencia, Spain. *J Immunogenet* 1981; **8**: 529–532.
- 7 Hernández M: Some genetic markers in Catalonia. *Gene Geography* 1987; **1**: 189–191.
- 8 Almuzara I, Martínez-Jarreta B, Castellano M: Distribution of G1m(1), G1m(2) and G3m(5) allotypes in Aragón. *Hum Hered* 1991; **41**: 409–410.
- 9 Moreno P, Matsumoto H: Immunoglobulin (Gm and Km) allotypes in two populations of Catalonia (Spain). *Am J Phys Anthropol* 1991; **85**: 385–338.
- 10 Calafell F, Bertranpetit J: Mountains and genes: Population history of the Pyrenees. *Hum Biol* 1994; **66**: 823–842.
- 11 Dugoujon JM, Clayton J, Sevin A, Constans J, Loirat F, Hazout S: Immunoglobulin (Gm and Km) allotypes in some Pyrenean populations of France. *Coll Antropol* 1989; **13**: 43–50.
- 12 Hazout S, Dugoujon JM, Loirat F, Constans J: Genetic similarity maps and immunoglobulin allotypes of eleven populations from the Pyrenees (France). *Ann Hum Genet* 1991; **55**: 161–174.
- 13 Cavalli-Sforza LL, Menozzi P, Piazza A: *History and Geography of Human Genes*. Princeton University Press, Princeton, NJ, 1994.
- 14 Aguirre AI, Vicario A, Mazon LI *et al*: AK1, PGD, GC and HP frequencies in the Basque population: a review. *Gene Geography* 1989; **3**: 41–45.
- 15 Manzano C, Aguirre AI, Iriondo M, Martín M, Osaba L, de la Rúa C: Genetic polymorphisms of the Basques from Gipuzkoa: genetic heterogeneity of the Basque population. *Ann Hum Biol* 1996; **23**: 285–296.
- 16 Cava A: El neolítico en al País Vasco. *Munibe* 1990; **42**: 97–106.
- 17 Urbeltz JA: Reflexiones sobre la danza. Euskal dantza tradizionala. In: J. Intxausti (ed.): *Euskal Herria: Historia eta Gizartea*. Caja Laboral Popular-Lan Kide Aurrezkia, 1985, pp437–488.
- 18 De la Rúa C: Los estudios de paleoantropología en el País Vasco. *Munibe (Antropología-Arqueología)* 1990; **42**: 199–219.
- 19 Manzano C, Orue JM, de la Ruac: The 'Basqueness' of the Basques of Alava: A reappraisal from a multidisciplinary perspective. *Am J Phys Anthropol* 1996; **99**: 249–258.
- 20 Arminio ML, Gómez P: Niveles de endogamia y consanguinidad en la población de los Montes de Pas (Cantabria, España).. *V Congreso Español de Antropología Biológica*, León, 1987, pp275–284. Ed. Universidad de León, León (España).
- 21 De Grado M, Esteban E, Moral P: Polimorfismos genéticos en la población de los Montes de Pas. In: Nieto JL and Moreno L (eds) *Avances en Antropología ecológica y genética*, Universidad de Zaragoza, 1995, pp 315–320.
- 22 Field LL, Dugoujon JM: Immunoglobulin allotyping (Gm and Km) of GAW5 families. *Genet Epidemiol* 1989; **6**: 31–34.
- 23 Edwards AWF: *Likelihood*. Cambridge University Press, Cambridge, 1984.
- 24 Wright S: *Evolution and the genetics of populations*. Vol. 4. Variability within and among natural populations University Chicago Press, Chicago, 1978.
- 25 Workman PL, Nisvander JD: Population studies on southwestern Indian Tribes. II. Local genetic differentiation in the Papago. *Am J Hum Genet* 1970; **22**: 24–49.
- 26 Felsenstein J: PHYLIP. Phylogeny Inference Package (version 3.2). *Cladistics* 1989; **5**: 164–166.
- 27 Cavalli-Sforza LL, Edwards WF: Phylogenetic analysis: models and estimation procedures. *Am J Hum Genet* 1967; **19**: 233–257.
- 28 Reynolds J, Weir BS, Cockerham CC: Estimation on the coancestry coefficient: basis of a short-term genetic distance. *Genetics* 1983; **105**: 757–779.
- 29 Nei M, Roychoudhury AK: Gene differences between Caucasian, Negro and Japanese populations. *Science* 1972; **177**: 434–435.
- 30 Efron B: *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1982.
- 31 Harpending HC, Jenkins T: Genetic distance among southern African populations. In: Crawford MH and Workman PL (eds) *Methods and Theories of Anthropological Genetics*, University of New Mexico Press, Albuquerque, 1973, pp177–199.
- 32 Blanc M, Ducos J: Les allotypes des systèmes GM et KM dans les provinces françaises. In: Ohayon E and Cambon-Thomsen A (eds) *Génétique des populations humaines*, Coll INSERM, 1986, vol. 142, pp 207–226.
- 33 Piazza A, van Loghem E, de Lange G, Curtoni ES, Ulizzi L, Terrenato L: Immunoglobulin allotypes in Sardinia. *Am J Hum Genet* 1976; **28**: 77–86.
- 34 Lefranc G, de Lange G, Rivat L *et al*: Gm, Am and Km immunoglobulin allotypes of two populations of Tunisia. *Hum Genet* 1979; **50**: 199–211.
- 35 Chaabani H, Helal AN, van Loghem E *et al*: Genetic study of Tunisian Berbers. I. Gm, Am and Km immunoglobulin allotypes and AB0 groups. *J Immunogen* 1984; **11**: 107–113.



- 36 Chaabani H, Cox DW: Genetic characterization of Tunisian Berbers. *Hum Hered* 1988; **38**: 308-316.
- 37 Helal AN, Rivat L, van Loghem E, de Lange G, Langaney A, Lefranc G: Further contribution of common GM*-Am* haplotypes and Km* alleles on the characterization of the Tunisian population. *Expl Clin Immunogenet* 1988; **5**: 1-14.
- 38 Dugoujon JM, Borot N, Chaventré A, Rudan P: Gm and Km immunoglobulin allotypes in the Olib and Silba islands (Northern Dalmatia, Yugoslavia). *Coll Antrop* 1989; **13**: 291-297.
- 39 Moral P, Marogna G, Salis M, Succa V, Vona G: Genetic data on Alghero population (Sardinia): Contrast between biological and cultural evidence. *Am J Phys Anthropol* 1994; **93**: 441-453.