The Basques in Europe: a genetic analysis.

A. PIAZZA
N. CAPPELLO
E. OLIVETTI
S. RENDINE

SUMMARY

Our knowledge of genetic differentiation in human populations of Europe at present, with special emphasis on Basques and their Mediterranean neighbours, is presented and discussed.

The preliminary results obtained by data from 15 genetic systems (ABO, RH, MNS, KELL, P1, HLA-A, HLA-B. GM, KM, BF, GC, GLO1, HP, TF, PI) and 65 alleles sampled from 52 ethnic groups are shown. Some tentative interpretations of the analysis are given:

(a) Modern Basques are more closely related (in genetic terms) to European rather than non-European populations.

(b) If genetic differentiation of modern Basques from the neighbouring Béarn sample occurred at the latest during the Romanization time, therefore it seems plausible that Basques differentiated from European populations before the Celtic invasion. This confirms the linguistic evidence that Basques have nothing to share with Celts.

(c) If Basques share a common genetic ancestry with Caucasian speaking and North-African populations, this occurred before the introduction of Neolithic farming. This may support the idea that the Basques are the descendants of a Paleolithic population.

1. INTRODUCTION

Almost half a century ago it was suggested (BOSCH-GIMPERA, 1943) that the Basques are the descendants of the populations who lived in Western Europe during the late paleolithic period. Their withdrawal to the area of the Pyrenees Mountains, probably caused by different waves of invasion, left the Basques untouched by the Eastern European invasions of the Iron Age. Moreover Basques speak a language which is quite different from Indo-European languages spoken in the rest of Europe and linguistic records confirm that they had no or very few contacts with the Celts and with the Iberians.

If the language of the Basques bears any testimony of their ethnic genealogy and of the preservation of the cultural inheritance of their ancestors, another matter could be their biological features as compared to other European populations. In his study of the geographic distribution of Rh blood groups, A.E. MOURANT (1948) pointed out that, as originally observed by Etcheverry in a paper hitherto unknown to the European workers, the Rh negative gene, which is found almost exclusively in Caucasoids, has its highest frequency among the Basques. MOURANT hypothesized that modern Basques may consist of

a paleolithic population with an extremely high Rh negative gene, who later mixed with people from the Mediterranean area. Thirty years later Cavalli-Sforza and his co-workers (AMMERMANN and CAVALLI-SFORZA, 1984) suggested that the immigrants might have been the neolithic farmers who came from the Near East and whose descendants today have an Rh negative gene with very low frequency. In this perspective, the Basques could be the last relict of a mesolithic population.

Our analysis summarizes tentatively our knowledge of genetic differentiation in human populations of Europe at present, with special emphasis on Basques and their Mediterranean neighbours.

2. THE BASQUES ARE AN OLD AND LONG-ISOLATED POPULATION

Three lines of evidence are generally given (for a short review see ALLIÈRES 1977):

a) **Linguistic evidence**
   — The language does not belong to the Indo-European family.
   — Toponymy shows Basque place-names extending to a large area surrounding the present Basque country.

b) **Archaelogical evidence**
   — No discontinuity between paleolithic and neolithic settlements.
   — Anthropometric features similar to those of the late paleolithic inhabitants of Europe.

University of Torino
Dipartimento di Genetica, Biologia e Chimica Medica
Via Santena 19, 10126 Torino. Italy.

— Late domestication of cattle: records of domestication before the Iron Age are scanty.

c) **Genetic evidence**

— Lowest frequency of blood group B in Europe (MOURANT, KOPEC, DOMANYEWSKA-SOBC-ZAK 1976).

— Highest frequency of blood groups Rhesus (d), HLA-B18, BfF1 in Europe. For HLA see ARNAIZ-VILLENA, RODRÍGUEZ DE CÓRDOBA, VELA, PASCUAL, CERVERÓ, BOOTELLO 1981.

— Correlation between blood group isogenic lines and toponyms in the Basque country (RUFFIÉ and BERNARD 1974).

3. MAIN QUESTIONS

These can be summarized as follows, the references being the oldest we noticed:

a) **Does the extreme isolation shown by the Basque language also reflect a similar genetic isolation?**

It is well known that the use of a language by a population is a cultural trait which can be imposed by few politically powerful people without any substantial effect on the genetic structure of the population itself. However when it is associated to a specific genetic identity, then language and population may share a common origin.

b) **How old are the Basques?**

What support can genetic analysis provide to date the time of the original Basque settlements?

c) **To what extent may the Basque population be considered genetically related to the following population?**

— the «VASCONES» who are thought to have formed the ethnic group who inhabited a greater part of Spanish Navarre and the French Gascogne (Aquitania as it was called in the time of J. Caesar) and to have split afterwards into a Romance (GASCON) and a non-Romance (BASQUE) speaking population. (LUCHAIRE 1877).

— the IBERIANS (TOVAR 1950) and their putative ancestors the BERBERS from North Africa (GÈZE 1883).

— some MEDITERRANEAN populations whose ancestors spoke non Indo-European languages or Indo-European languages with non Indo-European substrata: they are supposed to have been settled in the Near East, North Africa, Sardinia, Liguria, Iberia. (HUBSCHMID 1960).

— the CAUCASIAN speaking populations (TROMBETTI 1926, LAFON 1933).

4. GENETIC DATA

Blood types detected by immunological techniques, *electrophoretic variants* reflecting variations in electrophoretic mobility of enzymes or proteins, and *anthropometric traits* such as morphological measurements, skin and hair colour, body shape, etc., are the major source of data for measuring variations in extant human populations. Blood types and electrophoretic variants are genetically controlled markers, i.e. their transmission from one generation to the following one and their distribution in different populations can be predicted by probabilistic laws. These are usually expressed in terms of quantities called *gene frequencies:* a gene is a segment of DNA in an individual's chromosome which controls the expression of a genetic trait, for instance a blood type, and the frequencies of their variants in a population (called *alleles)* define the gene (or allele) frequencies. They usually vary in different POpulations and within the same population — over time. The evolution of human populations can therefore be studied in terms of their differences in gene frequencies. In order to gain some insight into the nature of these differences, one has to know the genetic mechanisms which can influence their variation in time. Four such mechanisms are commonly described: *mutation* which creates new genetic variants and then new genes at a very slow pace; *migration* and admixture between populations that have been and are remarkable factors of genetic change in our history; natural selection which allows the survival of genes conferring to the individual the highest probability of adapting to the environment where he lives; *genetic drift* whose effect is a random change of gene frequencies due to the finite size of the population, the smaller the population is, the more the change becomes apparent.

It is intuitively clear that the differences of gene frequencies in different populations give us information on the evolutionary history of the population themselves. But it is also clear that they cannot tell us which of the previous mechanisms were playing a rôle in it. Naturally all existing genes originate through mutations, but the frequency with which they occurr seems to be rarely if ever above 1 per 10,000 per generation per gene. In the genetic differentiation of our species, *Homo sapiens sapiens,* which occupied at most 50,000-100,000 years, mutations rates alone are likely to play a secondary role, if any. Natural selection has the distinction of being not only the unique adaptive mechanism of evolution, but also the fastest, well in the span of life of

our species and perfectly compatible with a process of migration giving the same amount of differentiation in the same time. Clearly natural selection is going to affect every gene in a different way and differently in different environments, while migratin affects all genes in a similar way.

The last consideration would lead to methodologies of data analysis allowing to distinguish genetic differentiation by migration from genetic differentiation by natural selection, could random genetic drift not play any substantial role in human evolution. However —speciallyin certain periods of our history and in special geographic areas— populations are structured in samples of relatively small sizes as in the case of Basques, and random fluctuations of their gene frequencies taking place at every generation are indeed the rule rather than the exception. Also random genetic drift affects all the genes equally, being entirely dependent on the sample size of the population and not on the single gene. A way to discriminate genetic drift from migration is by observing that in a population the genetic variability usually increases with migration from outside and decrases with genetic drift because random fluctuations may drive gene frequencies to values of zero. In fact a sort of equilibrium between migration and genetic drift may occurr that could explain why different populations presumably exposed to the same environmental conditions show differences in gene frequencies kept stable in time.

A consequence of the previous considerations is that when Basque genetic data are compared with those of other populations, we are faced by the problem of distinguishing a similarity by common environment or common culture. An empirical strategy to solve the dilemma is to evaluate differences based not on single genes (for instance one blood type only) but shared by as many as possible genes. As a swallow does not make a summer, so a gene in common does not prove a common genetic origin.

It is the cumulative genetic difference between populations which is expected to summarize their evolutionary history, being proportional to the time of their separation and inversely related to the migration between them. In line with this reasoning the analysis which follows deals with as many genes as we could find in the literature and with statistical tools allowing the simultaneous consideration of many variables.

5. GENES AND POPULATIONS

Data were obtained for a total of 15 genetic systems and 65 alleles. The alleles are variants of the same genetic system (or locus) and the frequencies of each of them are called gene or allele frequencies. The names of the genetic systems with the corresponding alleles used in the following analysis are listed in TABLE I.

A very large data bank including the gene frequencies collected from all over the world will form the basis of a systematic *Atlas of Human Gene Frequencies* to be published by Cavalli-Sforza, Menozzi and myself. A subsample of this collection has been taken so as to supply a substantial set of gene frequency data for the populations given in TABLE II. The gene frequencies we considered for our analysis are averages over all samples included in our data base with a specific population name, geographical area and linguistic group. Names in capital letters (last column) refer to populations, while names in lower-case letters refer to geographical areas where the samples of the specified linguistic group were drawn from.

Sample sizes are not given in this table, because they vary depending on the genetic systems. Moreover sample sizes less than 100 individuals were not taken into consideration. Gene frequencies and sample sizes are available on request.

GENETIC SYSTEM	ALLELES							
ABO	A	B	O					
RHESUS	CDE	CDe	cDE	cDe	CdE	Cde	cdE	cde
MNS	MS	Ms	NS	Ns				
KELL	K	k						
P1	P	p						
HLA–A	1	2	3	9	10	11	28	29
	30	32						
HLA–B	5	7	8	12	13	14	15	16
	17	18	21	22	27	35	40	
GM	zag	zaxg	zabst	fb				
KM	1	3						
BF	S	F	S0.7	F1				
GC	1	2						
GLO1	1	2						
HP	1	2						
TF	C	B	D					
PI	M	S						

Table 1.
Genetic systems and alleles.

POPULATION	LINGUISTIC GROUP	SAMPLES
BASQUE	NON INDO-EUROPEAN	France, Spain
FRANCE - Central	INDO-EUROPEAN	
FRANCE - North	"	
FRANCE - Béarn	"	
FRANCE - Catalogne	"	
FRANCE - Corse	"	
SPAIN	"	
ITALY - (excl. Lig Sar)	"	
ITALY - Sardinia	"	
ITALY - Liguria	"	
ENGLAND	"	
GREECE	"	
EUROPE - North	INDO-EUROPEAN (excl.*)	BELGIAN, DANE, DUTCH, ICELANDER, IRISH, FINN*, NORWEGIAN, SCOTT, SWEDEN, WELSH
EUROPE - Central	" "	AUSTRIAN, CZECK, GERMAN, HUNGARIAN*, POLE, RUSSIAN, SWISS
AFRICA - North	AFRO-ASIATIC	Algeria, Egypt, Lybia, Morocco, Tunisia
BERBER	"	Algeria, Morocco, Tunisia
AFRICA - East	"	Djibouti, Ethiopia, Kenya, Somalia, Sudan
NEAR EAST	"	BEDOUIN, Iraq, Iran, Jordan Kuwait, Lebanon, Saudi Arabia, Syria, Yemen
TURKIC	ALTAIC	Altai, DOLGAN, TURK, TURKOMAN, TUVAN, UZBEK, YAKUT
CAUCASIAN	CAUCASIAN	Caucasus, Dagestan, Georgia

Table 2.
Populations and samples.

6. METHODS OF ANALYSIS AND MAIN RESULTS

An approach for studying the genetic differentiation of populations has made use of coefficients of *genetic distance* between population pairs. Without going into details on the great deal of literature and controversy generated over them (see LALOUEL 1980 for a review), the genetic distance can be simply defined as a measure of similarity between two populations. Originally proposed by CAVALLI-SFORZA and EDWARDS (1963) more than twenty years ago, this parameter is a useful tool for interpreting genetic differences, whenever we are aware of the (genetic) assumptions which it is based on. Of the many formulations this measure has taken in the literature, we adopted the coancestry coefficient proposed by REYNOLDS, WEIR and COCKERHAM (1983) as the most reliable one in the common case of samples with different sizes. This measure of genetic distance applies specifically to short-term evolution and whenever the divergence between populations sharing a common ancestral population may be regarded as being due mainly to genetic drift, it is proportional to the time of divergence.

Table 3. Genetic distances between basques and some selected populations.

POPULATION	GENETIC DISTANCE (x10,000)
FRANCE BÉARN	70
SPAIN	102
FRANCE CATALOGNE	107
FRANCE CENTRAL	113
ENGLAND	132
FRANCE NORTH	153
ITALY	157
EUROPE CENTRAL	163
ITALY LIGURIA	173
EUROPE NORTH	180
GREECE	194
FRANCE CORSE	224
CAUCASIAN	241
NEAR EAST	242
ITALY SARDINIA	312
AFRICA NORTH	355
BERBER	392
TURKIC	664
AFRICA EAST	948

An example of their calculation is displayed in TABLE III. The genetic distances between Basques and the populations of TABLE II are shown.

A simple inspection of TABLE III suggests some interesting considerations which can be summarized in the following points:

RESULT 1

The sample from FRANCE-BÉARN shows the greatest genetic similarity with Basques. This is not unexpected if one thinks of French Gascogne as the geographical area where the «VASCON» ethnic group could have been the last one to have differentiated from the Basques. Evidences from recent and not yet published genetic analyses on the French Provinces support this result. An impressive analogy between gene frequency and dialect geographical distribution have been found: the BÉARN region in particular, when compared with the other French regions, shows clear genetic differences.

RESULT 2

With the interesting exception of Sardinia, a Mediterranean island whose genetic peculiarity is even greater than that of Basques, the Indo-European speaking populations are genetically less distant from the Basques than are the non Indo-European speaking populations.

It is only fear to add that the genetic distance corresponding to the Caucasian speaking populations is much less reliable than the other ones because these populations were not tested for the genetic system HLA (see TABLE I) and this could unbalance the estimate in an unpredictable way.

The *principal component* multivariate technique was also employed. This well-known procedure of displaying multivariate observations by reducing their dimensionality so as to minimize the amount of information lost in such reduction, allows to plot the genetic distances among all the studied populations on a plane. The coordinates of this plane, technically called «principal component» axes, are found by minimizing the distortion of the multidimensional genetic distances when they are projected on the plane.

Figure 1 is a representation of this kind: it gives a synthetic even if lightly distorted view of the genetic position of Basques as compared to the populations described above. The distance between the centers of any pair of circles measures their genetic differences with the least possible distortion. The area of the circles around each point are by construction inversely proportional to such distortion, so that only the distances between points associated to nearly equal areas can be safely compared, while all other are in fact larger than their represented
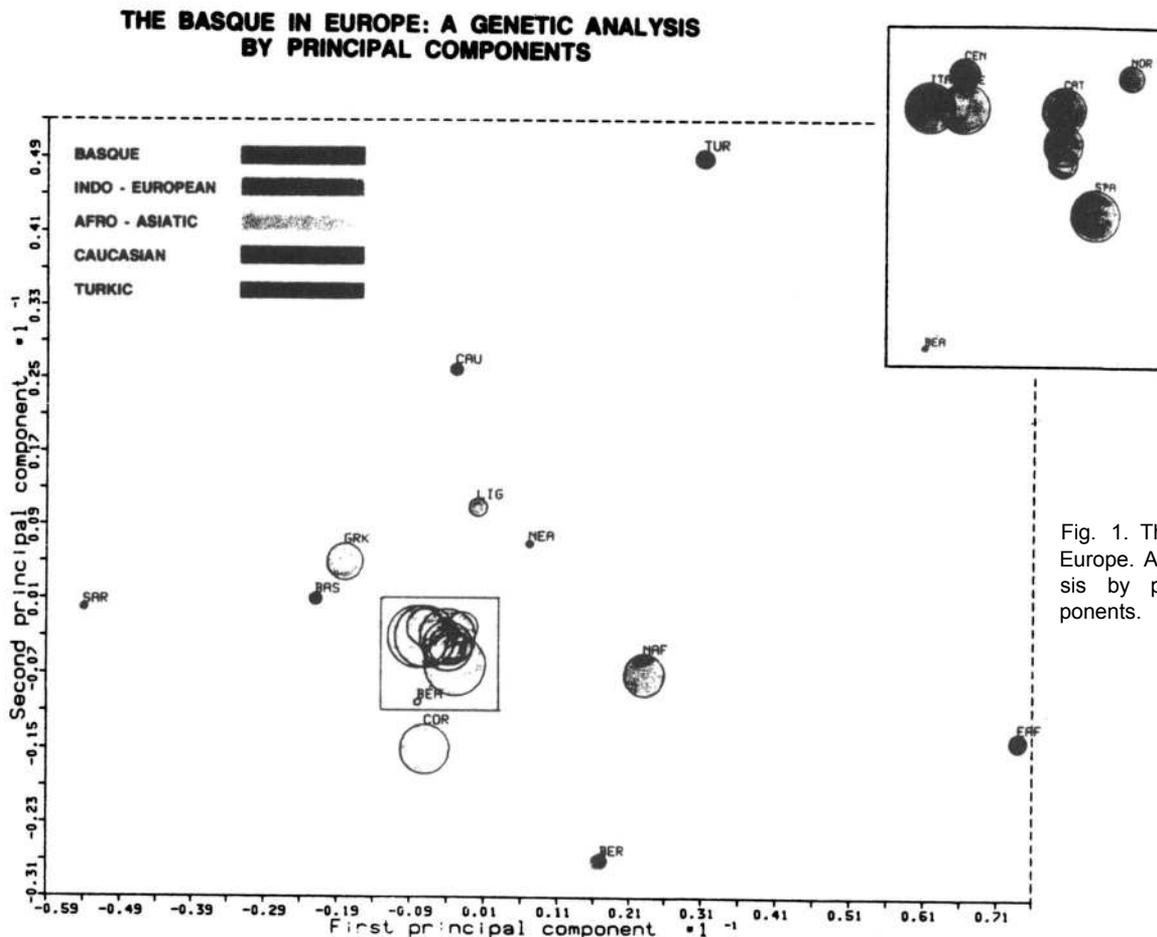
Fig. 1. The Basques in Europe. A genetic analysis by principal components.

projections on the plane. For instance, the Basques and the French from Béarn because of their nearly equal areas are genetically more similar than Basques and Greeks who seem less distant in the two-dimensional projection. The figure complements the information given in TABLE III by adding the reciprocal relationships among populations.

RESULT 3

A remarkable feature of this condensed display is that Sardinia results genetically different from all other Indo-European speaking populations by an amount which is greater than that of Basques but in the same direction (same side of the plot). This observation will be further elaborated below.

Another way to display the information of the previous figure in a more suggestive synthesis is to plot the geographical position of each population and to relate it to a fictitious genetic position which is compatible with its genetic distances from all other populations. Technically speaking this means to fit a set of genetic distances into another set of distances, the geographical ones. The problem can be solved by a nonmetric multidimensional scaling technique (SCHÖNEMANN and CARROL 1970) and the application to our data is shown in *Figure 2.*

Each line in the figure joins the geographical position of the population (smaller point) with a «ge-

netic position» (larger point) consistent with other genetic distances. The length of the line segment is a measure of discrepancy between genetic and geographical points with no departing lines are supposed to have their «genetic position» inside the shaded area located between Sardinia and continental Italy.

Some interesting features of this figure deserve mention:

— The genetic positions of African populations remain in Africa.

— Near East and Caucasian speaking populations seem to get closer to West Europe as if their genes were similar to the European ones than one might expect from their geographical position.

— Basque and Ligurian populations are shifted to the East.

— Genetic isolation of Sardinia is confirmed.

— Greece behaves differently from other Indo-European speaking populations.

A very popular representation of human evolution uses *trees of descent.* Relationships among populations can be always represented by trees, as they are one of the possible modes of showing similarities among taxonomic units. The problem is whether tree-like structures are a good or bad representation of our empirical data. If evolution is assumed to be independent in the various branches of the tree (i.e. there are not migratory exchanges,
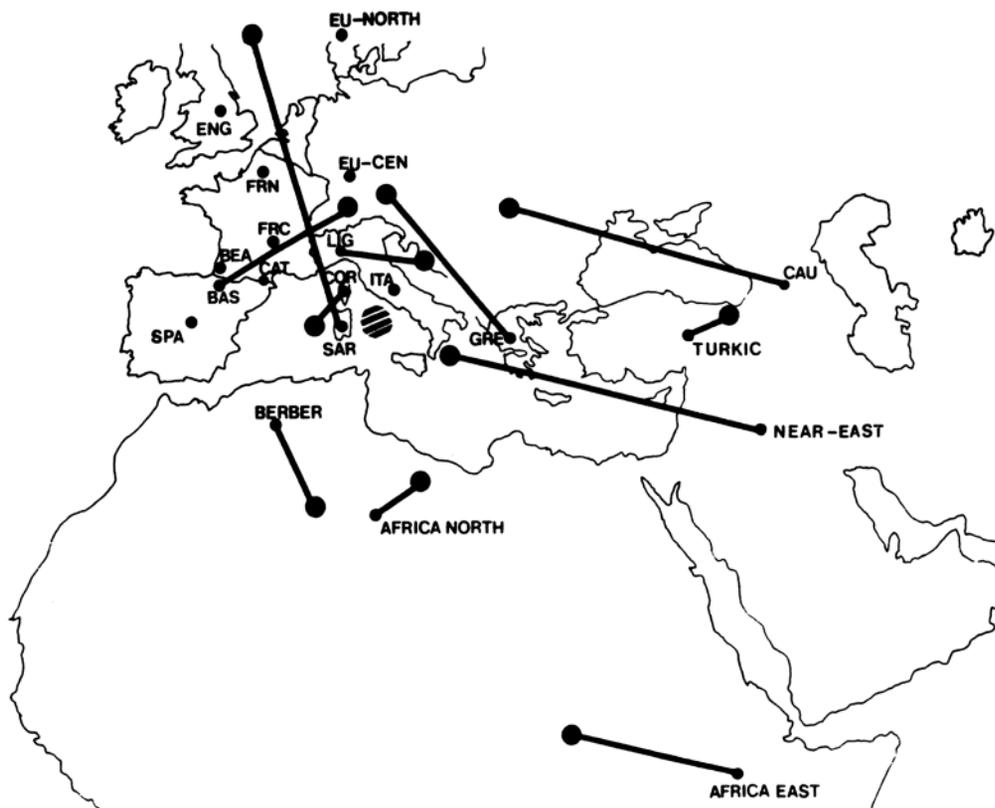
Fig. 2. Genetic and geographical distances as they result by applying the technique by Schöemann and Carrol (1970).

no major hybridization, no convergence due to similarity of environment, etc.) then it can be reconstructed in terms of phylogenetic trees (CAVALLI-SFORZA and EDWARDS 1967, CAVALLI-SFORZA AND PIAZZA 1975). From this point of view, the case of Europe is hopeless: the history of European populations is a sequence of invasions, colonizations, movements and settlements, which are exactly the factors usually preventing a process of independent evolution. In fact it has been shown that most of the European genetic variation has been determined by the spread of farmers from Near East in the Neolithic times (MENOZZI, PIAZZA, CAVALLI-SFORZA 1978): obviously a process of this kind cannot be modeled by a phylogenetic tree. In spite of all that, the euristic value of a tree representation cannot be disregarded: it provides at least an easy way of looking at clusters.

Figure 3 shows a tree-like display of our data. The «bootstrap» approach proposed by FELSENSTEIN (1985) has been used in order to validate the tree topology. The best supported tree is shown. Multifurcations indicate a low confidence limit when further hierarchical levels in the phylogeny are tried.

This representation, like the previous ones, exhibits the already considered genetic similarity between Basques and the French from Béarn. Most European populations seem to follow parallel lines of evolution without intermediate interactions. Notice also the position of the East Mediterranean populations: they seem to group with West Mediterranean rather than with African populations.

The last analysis we like to present has the purpose of testing whether the origin of Basques can
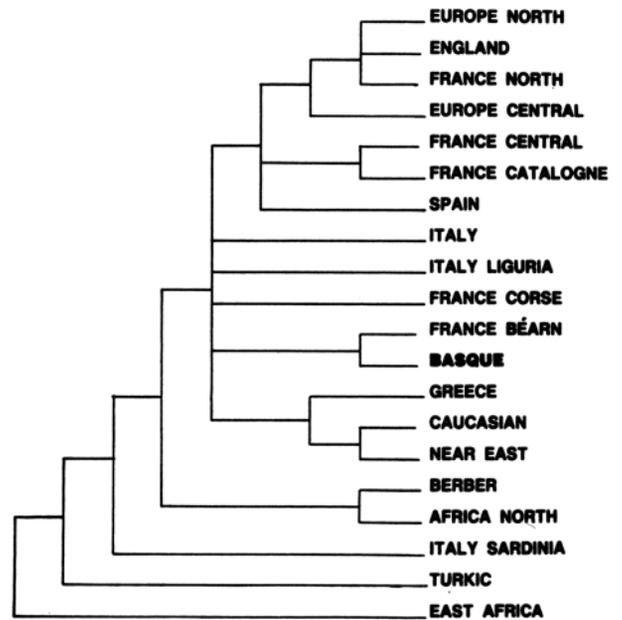
Fig. 3. A tree-like display showing the relationships of Basques with other populations.

be explained in terms of a hypothetical admixture of two of the remaining populations we had considered. The technique proposed by WIJSMAN (1984) has been employed and the results for all genes taken simultaneously are summarized in *Figure 4.* The upper side shows all the pairs of putative ancestor populations which have been tested. Only the statistically significant estimates are shown in the lower side of the figure. Significant and not significant estimates of possible gene flow effects due to admixture are consistent in suggesting the following.

Fig. 4. Estimates of possible admixture in Basques.

RESULT 4

Gene flow of European origin can reasonably be excluded in Basques. However, a percentage of Basque genes seems to derive from Sardinian ancestors. Their origin is unknown but the archeology of Sardinia shows a substantial colonization from North Africa. Therefore it is reasonable to think of a North African component in the Basque gene pool.

7. TENTATIVE CONCLUSIONS

a) Modern BASQUES are more closely related to EUROPEAN rather than AFRO-ASIATIC (see TABLE II) populations.

b) On the basis of a hypothetical PYRENEAN population originally settled in a large area including the homonymous mountains, we can postulate that the genetic differentiation of modern BASQUES from the neighbour BÉARN sample occurred at the latest during the Romanization time, about 2,000 years B.P.

c) On the assumption that genetic distances are proportional to the times when population separated (which is strictly valid only if genetic drift is a major cause of differentiation) we can rewrite TABLE III in the following TABLE IV

Table 4. Genetic distances and times of possible admixture of basques with relevant populations.

POPULATION	DISTANCE	YEARS B.P.
FRANCE BÉARN	100	2,000
SPAIN	146	3,000
FRANCE CATALOGNE	153	3,100
CAUCASIAN	344	6,800
ITALY SARDINIA	446	8,900
AFRICA NORTH	507	10,100
BERBER	560	11,200

where the times of differentiations are calculated from the genetic distances by setting the time of differentiation between Basques and the French from Béarn, 2,000 years B.P.. In order to compare this times with historical events which have been relevant for the Iberian peninsula, we have summarized some of them in TABLE V.

Table 5. Some historical events in the Iberian Peninsula.

Reference: Ramos–Oliveira 1971

Event	Years B.P.
Small population size (5,000 ?)	50,000
Capsian North African invasion	20,000
Saharian invasion	8,000
Introduction of farming	6,000
Celtic invasion	2,600
Population size between 2 and 3 millions	
Minor establishments of Far-Easteners, Greeks, Romans	
About 30,000 Arabian invaders	1,276
Population size about 4 million	

By comparing TABLE V with TABLE IV it seems plausible that genetic differentiation from EUROPEAN samples occurred before the Celtic invasion. This confirms the linguistic evidence that BASQUES HAVE NOTHING TO SHARE WITH CELTS.

d) IF BASQUES share a common ancestry with CAUCASIAN speaking and NORTH-AFRICAN populations, the corresponding modern samples (including SARDINIA whose African component is documented) show hypothetical separation times before the introduction of farming (see TABLES IV and V). This genetic evidence supports the idea that the BASQUES are the descendants of a PALEO-LITHIC population. The linguistic evidence of possible Afro-Asiatic substrata could reflect a common genetic origin with CAUCASIAN speaking and NORTH-AFRICAN populations which might have occurred at the latest before the Neolithic introduction of agriculture.

BIBLIOGRAFIA

ALLIÈRES, J.

1977.   *Les Basques,* Press Universitaires de France, Paris.

AMMERMAN, A.J., CAVALLI-SFORZA LL.

1984.   *The Neolithic Transition and the Genetics of Populations in Europe, Princeton University Press,* Princeton.

ARNAIZ-VILLENA, A.; RODRIGUEZ DE CORDOBA, S.; VELA, F.; PASCUAL, J.C.; CERVERO, J.; BOOTELLO, A.

1981.   HLA antigens in a sample of the Spanish Population: common features among Spaniard Basques an Sardinians. *Human Genetics 58,* 344.

BOSCH-GIMPERA, A.

1943.   El problema de los orígenes vascos. *Eusko-Jakintza 3,* 39.

CAVALLI-SFORZA LL.; EDWARDS AW.

1963.   Analysis of Human Evolution. *Proc. XI Int. Congr. Genet. 3,* 923.

1967.   Phylogenetic analysis. Models and estimation procedures. Amer. *J. Hum. Genet. 19,* 233.

CAVALLI-SFORZA LL.; PIAZZA A.

1975.   Analysis of evolution. Evolutionary rates, independence and treeness. *Theoretical Population Biology, 8,* 127.

HUBSCHMID J.

1960   Mediterrane Substrate mit besonderer Berücksichtigung des Baskischen und der west-östlichen Sprachbeziehungen. *A. Francke, Berne.*

FELSENSTEIN, J.

1985.   Confidence limits on phylogenies. An approach using the bootstrap. *Evolution 39,* 783.

GÈZE, L.

1883.   *De quelques rapports entre les languages Berbéres et le Basque,* Toulouse.

LAFON, R.

1933.   Basque et langue Kartvèles, *Revue Internationale des Etudes Basques 24,* 2.

LALOUEL, G.M.

1980.   Distance analysis and multidimensional scaling. In *Current Developments in Anthropology* (ed. Mielke and Crawford) Plenum Press, New York, 1, 135.

LUCHAIRE, A.

1877.   *Les origines linguistiques d'Aquitaine,* Paris.

MENOZZI, P.; PIAZZA, A.; CAVALLI-SFORZA

1978.   Synthetic maps of human gene frequencies in Europeans. *Science 201,* 786.

MOURANT, A.E.

1948.   Basque blood groups. *Nature 162,* 27.

MOURANT, A.E.; KOPEC, A.C.; DOMANIEWSKA-SOBCZAK, K.

1976.   *The distribution of Human Blood Groups and other Polymorphisms. Oxford University Press,* London.

REYNOLDS, J.; WEIR, B.S.; COCKERHAM, C.C.

1983.   Estimation of the coancestry coefficient. Basis for a short-term genetic distance. *Genetics 105,* 767.

RUFFIÉ, J.: BERNARD, J.

1974.   Peuplement du Sud-Ouest Européen. Les relations entre la biologie et la culture. *Cahiers d'Anthropologie et d'Ecologie humaine II,* 3.

SCHÖNEMANN, P.H.; CARROL MR.

1970.   Fitting one matrix to another under choice of a central dilatation and a rigid motion. *Psychometrika 35.* 245.

TOVAR, A.

1950.   *Léxico de las Inscripciones Ibéricas,* Madrid.

TROMBETTI, A.

1926.   *Le origini della lingua Basca,* Bologna.

WIJSMAN, E.M.

1984.   *Techniques for estimating genetic admixture Human Genetics 67,* 441.