

Genetic relationships in Spanish dog breeds. II. The analysis of biochemical polymorphism

J Jordana, J Piedrafita, A Sanchez

Universitat Autònoma de Barcelona, Unitat de Genètica i Millora Animal, Departament de Patologia i de Producció Animals, Facultat de Veterinària, 08193-Bellaterra, Barcelona, Spain

(Received 27 July 1990; accepted 24 February 1992)

Summary – The phylogenetic relationships between 10 Spanish dog breeds were studied using the gene frequency values obtained from the electrophoretic analysis of 21 structural genic loci that code for blood-soluble proteins and enzymes. In addition, we studied the genetic differentiation within breeds. In some cases the genetic distances between subpopulations of the same breed were greater than the genetic distances between different breeds. The average between-breed distance has a value of 0.0197 (\pm 0.0128), with extreme values of $D = 0.000$ between Gos d'Atura and Podenco Ibérico, and of $D = 0.051$ for the Mastín Español – Ca de Bestiar pair. The groupings of Spanish dog breeds obtained in our study from morphological and biochemical data were apparently quite similar. The correlation between enzymatic and morphological distances was, however, low ($r = 0.07$) and non-significant. The estimates of the divergence times among the 4 ancestral trunks suggest that the ancestral trunks separated independently in a relatively short interval of time, between 30 000 and 55 000 years ago.

Spanish dog breeds / biochemical polymorphisms / electrophoresis / genetic distance / genetic relationships

Résumé – Relations génétiques entre des races canines espagnoles. II. Analyse du polymorphisme biochimique. À partir des valeurs des fréquences géniques, obtenues par l'analyse électrophorétique de 21 locus qui codent pour des enzymes et des protéines solubles du sang, on a étudié les relations phylogénétiques existant entre dix races canines espagnoles. On a déterminé aussi le niveau de différenciation intraraciale, et constaté que, dans certains cas, les distances génétiques entre sous-populations d'une même race sont supérieures à celles existant entre races différentes. La distance moyenne entre races prend une valeur de 0,0197 (\pm 0,0128), avec des valeurs extrêmes de $D = 0,000$ entre «Gos d'Atura» et «Podenco Ibérico», et de $D = 0,051$ pour le couple «Mastín Español» – «Ca de Bestiar». Les groupements obtenus dans notre étude, à partir de données morphologiques et biochimiques, sont apparemment assez similaires. La corrélation entre distances enzymatiques et morphologiques est cependant très faible ($r = 0,07$) et non

significative. L'estimation des origines de la divergence entre les quatre troncs ancestraux, suggère que ces troncs se sont séparés dans un intervalle de temps relativement court, il y a 30 000 à 5 000 ans.

raças caninas espanholas / polimorfismo bioquímico / eletroforese / distância genética

INTRODUCTION

The genetic relationships in Spanish dog breeds have been studied in a previous paper with data from morphological characters (Jordana *et al*, 1992). Nevertheless, these characters have been, over time, under a great pressure of selection, either natural or artificial, this selection having had a great influence in the process of breed differentiation.

Assuming that genetic variability – detected through biochemical polymorphism – is maintained in populations by the equilibrium between mutation and genetic drift (Kimura, 1983), and that this polymorphism has not been deliberately selected by man, the analysis of that variability would give a more precise estimation of the relationships among populations.

Past electrophoretic and immunological studies of blood proteins and enzymes, to understand the genetic relationships among breeds of dog, include: Leone and Anthony, 1966; Tanabe *et al*, 1974, 1977, 1978; Sugiura *et al*, 1977; Juneja *et al*, 1981; and Kobayashi *et al*, 1987.

This paper is a study of the genetic relationships among Spanish dog breeds by the analysis, using electrophoretic techniques, of “neutral” structural genes that code for soluble proteins and enzymes of the blood. An analysis of within-breed genetic differentiation is also done starting from a total of 24 subpopulations because significant differences might exist among subpopulations of the same breed, owing to the specific characteristics of some subpopulations (size of flocks, reproductive isolation, etc). This will be useful to interpret and discuss the observed genetic relationships among breeds with more precision.

The resulting enzymatic phylogeny is compared with that which is observed from the analysis of morphological characters (Jordana *et al*, 1992), to check whether a possible evolutionary parallelism between both types or characters exists.

MATERIAL AND METHODS

A total of 484 blood samples has been taken in the 10 Spanish dog breeds, with the following distribution: Gos d'Atura (93), Mastín del Pirineo (55), Mastín Español (45), Perdiguero de Burgos (42), Galgo Español (31), Sabueso Español (53), Ca de Bestiar (46), Podenco Ibicenco (71), Podenco Canario (15) and Podenco Ibérico (33).

Blood samples were collected with sodium EDTA (1 mg per ml of blood) as an anticoagulant. The samples were separated into the 3 main blood components; plasma, red blood cells and white blood cells, and stored at -20°C .

The values of the allelic frequencies of the genes studied have been used to measure the genetic variation and to study the divergence among populations. Twenty-one loci were analyzed, according to the methodology that has been described in detail by Jordana (1989), by using electrophoretic techniques: horizontal electrophoresis in starch gel, polyacrylamide and agarose-polyacrylamide (bidimensional) gels. The total number of loci analyzed included 5 red blood cell systems: superoxide dismutase (Sod), glucose phosphate isomerase (Gpi), 6-phosphogluconate dehydrogenase (6-Pgd), phosphoglucomutase-1 (Pgm₁), and glucose 6-phosphate dehydrogenase (G6pd); 4 leucocyte systems: mannose phosphate isomerase (Mpi), malate dehydrogenase soluble form (Mdh_s), malate dehydrogenase mitochondrial form (Mdh_m), and acid phosphatase (Pac), and 12 plasma systems: leucine aminopeptidase (Lap), albumin (Alb), peptidase D (Pep-D), transferrin (Tf), prealbumin (Pr), Gc protein (Gc), α_1 B-glycoprotein (α_1 B), protease inhibitor (Pi-1), protease inhibitor-3 (Pi-3), postalbumin-1 (Pa-1), pretransferrin-1 (Prt-1) and pretransferrin-2 (Prt-2).

The breeds have been subdivided into 24 subpopulations to perform the within-breed analysis of the populations, according to geographical criteria and/or the areas of influence of certain breeders (table I). The 2 subpopulations of the Podenco Canario breed had to be built purely at random to perform the analysis, because there were no data about the origins of the individuals.

A factor analysis of principal components was done using the BMDP-4M program (Frane *et al*, 1985), to study the relationships among populations with data from the allelic frequencies of the polymorphic loci. These were taken as variables to typify the different populations.

Nei's unbiased distance (a modified version of D for small sample sizes; Nei, 1978) and the Cavalli-Sforza and Edwards' (1967) chord distance have been calculated. These 2 distances were chosen for the respective construction of phenograms and cladograms, owing to their properties. Nei *et al* (1983), using a "known" simulated phylogeny by computer and assuming a constant rate of molecular evolution, have found that: a), the trees generated using UPGMA and Wagner's methods with the Cavalli-Sforza and Edwards' (1967) chord distance produce the most accurate topology of the branches; and b), Nei's (1972, 1978) standard distances gave the best estimation of the branch lengths, when the tree was built up through the UPGMA algorithm. Besides that, unlike other distances these distances show a close linear relationship with the number of amino acidic substitutions, which makes them useful to obtain rough estimates of divergence times (Hedges, 1986; Nei, 1987).

A jackknife method (Muller and Ayala, 1982) was also used to calculate Nei's distances among populations, since it gives a more accurate estimation when the range of distances is below 0.1.

The reliability of the constructed phenograms has been evaluated by computing the standard errors (SE) at every point of bifurcation of the tree branches. The evaluation of the SE is important because every point of ramification suggests an important event of speciation or division of the population (Nei *et al*, 1985). In the same way, in the phenogram obtained with the values of Nei's distances by using the jackknife method, it is possible to make comparisons among clusters, checking whether the difference between the average distance among clusters and the average intracluster distance is significantly greater than zero. The reliability

Table I. Distribution by breeds and subpopulations of the analyzed material. The numbers of samples analyzed are given in parentheses.

<i>Breeds</i>		<i>Subpopulations</i>		
GA. Gos d'Atura	(93)	GA1	Tipico	(59)
		GA2	Cerdà	(34)
MP. Mastín del Pireneo	(55)	MP1	Zaragoza	(36)
		MP2	Benabarre	(11)
		MP3	Rest	(08)
ME. Mastín Español	(45)	ME1	Prat-Rest	(15)
		ME2	Barna-North	(16)
		ME3	Madrid	(14)
PB. Perdiguero de Burgos	(42)	PB1	Barcelona	(19)
		PB2	Rest	(23)
GE. Galgo Español	(31)	GE1	Catalunya	(21)
		GE2	Rest	(10)
SE. Sabueso Español	(53)	SE1	Catalunya	(25)
		SE2	Asturias	(28)
CB. Ca de Bestiar	(46)	CB1	Mallorca	(20)
		CB2	Palma	(13)
		CB3	Rest	(13)
PE. Podenco Ibicenco	(71)	PE1	Catalunya	(41)
		PE2	Baleares	(30)
PC. Podenco Canario	(15)	PC1	PC1.I	(07)
		PC2	PC2.P	(08)
PI. Podenco Ibérico	(33)	PI1	Urgell	(10)
		PI2	Garrigues	(17)
		P13	Rest	(06)

of the bifurcation points is indirectly checked and, with it, the reliability of the topology of the tree.

The values of the genetic distances among populations, the phenograms and cladograms, as well as the goodness-of-fit statistics of those dendrograms have been computed by using the BIOSYS-1 program (Swofford and Selander, 1981).

RESULTS

Gene frequencies

A total of 38 electromorphs have been identified whose distribution varied from 1 to 5. Taking as a criterion of polymorphism that of 95%, 10 systems (Gpi, 6-Pgd, Pgm-1, Mdh-s, Mdh-m, G6pd, Pac, Pr, Gc and Pi-3) were found to be monomorphic

for all populations. The allele frequencies for each polymorphic locus and breed are shown in table II.

The plasma proteins (Alb, Tf, Pi-1, α_1 -B, Prt-1, Prt-2, Gc, Pr, Pi-3 and Pa-1), which constitute 48% of the 21 analyzed loci, show a greater level of polymorphism than the enzymatic systems analyzed (Sod, Gpi, Lap, Mpi, 6-Pgd, Pgm-1, Mdh-s, Mdh-m, Pep-D and Pac) with the first group explaining 83.33% of the total polymorphism in the studied populations.

Only 2 populations showed disagreement with the expected Hardy-Weinberg proportions for some loci. These populations were: Sabueso Español for Tf ($P < 0.01$) and Prt-1 ($P < 0.05$) systems, and Podenco Ibicenco for α_1 -B ($P < 0.05$). The deficit of heterozygotes (D) was -0.382 , -0.374 , and -0.0269 , respectively.

Principal components analysis

In order to infer the possible relationships among populations, either at a breed level or at a subpopulation level, a principal components analysis with 3 factors has been done. The allelic frequencies of 11 polymorphic systems are used, giving a total of 17 independent variables.

Table III shows, over the total existing variation and over the total explained variation, the different percent values in decreasing order, of the systems that give more information about breed differentiation. 28.08% of the total explained variance corresponds to the transferrin (Tf) system, followed by the Lap, Pi-1, Alb, Sod, Prt-1, α_1 -B, Prt-2, Pa-1, Pep-D and Mpi systems.

At the breed level (fig 1), the first 3 factors explain 65.60% of the total variance. Three groups are closely related: Podenco Canario (PC) and Perdiguero de Burgos (PB) populations; Gos d'Atura (GA), Galgo Español (GE) and, less closely related, Podenco Ibérico (PI); and finally Mastín del Pirineo (MP) and Sabueso Español (SE). Mastín Español (ME) remains as an isolated population, although it is closer to the group formed by Mastín del Pirineo and Sabueso Español than to any other group. Although the Ca de Bestiar (CB) population differs from the others, it has a certain relationship with the group formed by Podenco Canario and Perdiguero de Burgos. Podenco Ibicenco (PE) appears clearly differentiated from the rest of the breeds.

When the analysis at the subpopulation level is done (fig 2), the explained total variance on the first 3 axes decreases to 49.83%. The diagram is, approximately, comparable to the one obtained at the breed level. A close relationship among the subpopulations of the Ca de Bestiar, Mastín Español, Gos d'Atura, Perdiguero de Burgos, Podenco Canario and Podenco Ibérico breeds is observed. The remaining breeds have a smaller relationship among their subpopulations, which suggests the existence of a certain degree of within-breed genetic differentiation.

Genetic distances and dendrograms

From the values of the gene frequencies of the analyzed loci and by means of the application of several indexes of genetic distance, dendrograms of the Spanish dog breeds have been obtained by 2 different methodologies: cluster analysis and Wagner's method. For the cluster analysis, the UPGMA algorithm (Sneath and

Table II. Values of the gene frequencies obtained for each of the 11 polymorphic loci in the 10 Spanish dog breeds. (*N*) stands for the number of animals.

<i>Locus</i>	<i>Allele</i>	<i>GA</i>	<i>MP</i>	<i>ME</i>	<i>PB</i>	<i>GE</i>	<i>SE</i>	<i>CB</i>	<i>PE</i>	<i>PC</i>	<i>PI</i>
Sod	(N)	93	55	45	42	31	53	46	71	15	33
	A	0.973	0.918	0.878	1.000	0.952	0.868	0.967	0.993	1.000	0.955
	B	0.027	0.082	0.122	0.000	0.048	0.132	0.033	0.007	0.000	0.045
Lap	(N)	93	55	45	42	31	53	46	71	15	33
	A	0.968	0.991	1.000	0.988	0.952	1.000	1.000	0.930	0.967	1.000
	B	0.032	0.009	0.000	0.012	0.048	0.000	0.000	0.021	0.000	0.000
	C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.049	0.033	0.000
Mpi	(N)	85	49	27	42	24	35	41	63	10	27
	A	0.906	0.724	0.667	0.821	0.687	0.857	0.817	0.937	0.700	0.981
	B	0.094	0.276	0.333	0.179	0.313	0.143	0.183	0.063	0.300	0.019
Alb	(N)	93	55	45	42	31	53	46	70	15	32
	S	0.575	0.518	0.667	0.881	0.452	0.377	0.848	0.529	0.867	0.578
	F	0.425	0.482	0.333	0.119	0.548	0.623	0.152	0.471	0.133	0.422
Pep-D	(N)	93	55	45	42	31	45	46	71	15	33
	A	0.876	0.964	0.900	0.905	0.903	0.900	0.870	0.958	0.933	0.909
	B	0.124	0.036	0.100	0.095	0.097	0.100	0.130	0.042	0.067	0.091
Tf	(N)	95	55	45	42	31	52	46	70	15	33
	A	0.005	0.000	0.000	0.000	0.000	0.010	0.000	0.029	0.000	0.000
	B	0.458	0.582	0.300	0.107	0.500	0.586	0.272	0.257	0.067	0.425
	C	0.532	0.418	0.489	0.881	0.500	0.404	0.728	0.714	0.933	0.545
	D	0.000	0.000	0.211	0.012	0.000	0.000	0.000	0.000	0.000	0.000
α 1-B	(N)	93	55	45	42	31	53	46	71	15	33
	S	0.602	0.273	0.589	0.429	0.661	0.443	0.348	0.796	0.633	0.561
	F	0.398	0.727	0.411	0.571	0.339	0.557	0.652	0.204	0.367	0.439
Pi-1	(N)	93	55	45	42	31	53	46	71	15	33
	S	0.231	0.155	0.600	0.190	0.290	0.236	0.097	0.585	0.133	0.273
	I	0.161	0.073	0.011	0.334	0.097	0.028	0.033	0.049	0.300	0.136
	F	0.608	0.772	0.389	0.476	0.613	0.736	0.870	0.366	0.567	0.591
Prt-1	(N)	91	55	45	42	31	43	46	69	15	32
	S	0.055	0.018	0.044	0.000	0.000	0.058	0.054	0.022	0.100	0.063
	F	0.709	0.718	0.778	0.665	0.468	0.744	0.402	0.630	0.667	0.656
	D	0.236	0.264	0.178	0.345	0.532	0.198	0.544	0.348	0.233	0.281
Prt-2	(N)	89	54	40	42	29	40	45	67	15	31
	S	0.028	0.028	0.013	0.012	0.052	0.025	0.000	0.007	0.000	0.032
	F	0.972	0.972	0.987	0.988	0.948	0.975	1.000	0.993	1.000	0.968
Pa-1	(N)	66	34	28	30	30	34	37	40	15	23
	S	0.636	0.691	0.750	0.750	0.667	0.691	0.135	0.675	0.800	0.696
	F	0.364	0.309	0.250	0.250	0.333	0.309	0.865	0.325	0.200	0.304

Table III. Percentage values (%) of the 11 polymorphic systems, over the total existing variation (1), and over the total explained variation (2), at the breed level.

<i>Locus</i>	(1)	(2)
Tf	18.423	28.084
Lap	9.151	13.949
Pi-1	7.978	12.161
Alb	5.527	8.425
Sod	5.354	8.162
Prt-1	5.094	7.766
α 1-B	4.517	6.886
Prt-2	4.099	6.248
Pa-1	2.930	4.466
Pep-D	1.956	2.982
Mpi	0.566	0.863

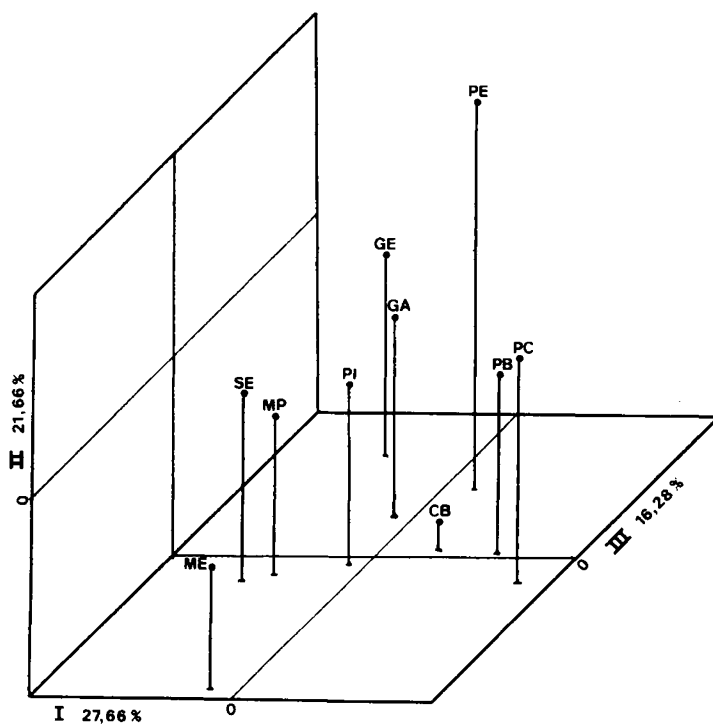


Fig 1. Factor analysis of principal components with 3 factors Spanish dog breeds. The gene frequencies of 11 polymorphic systems were used as variables.

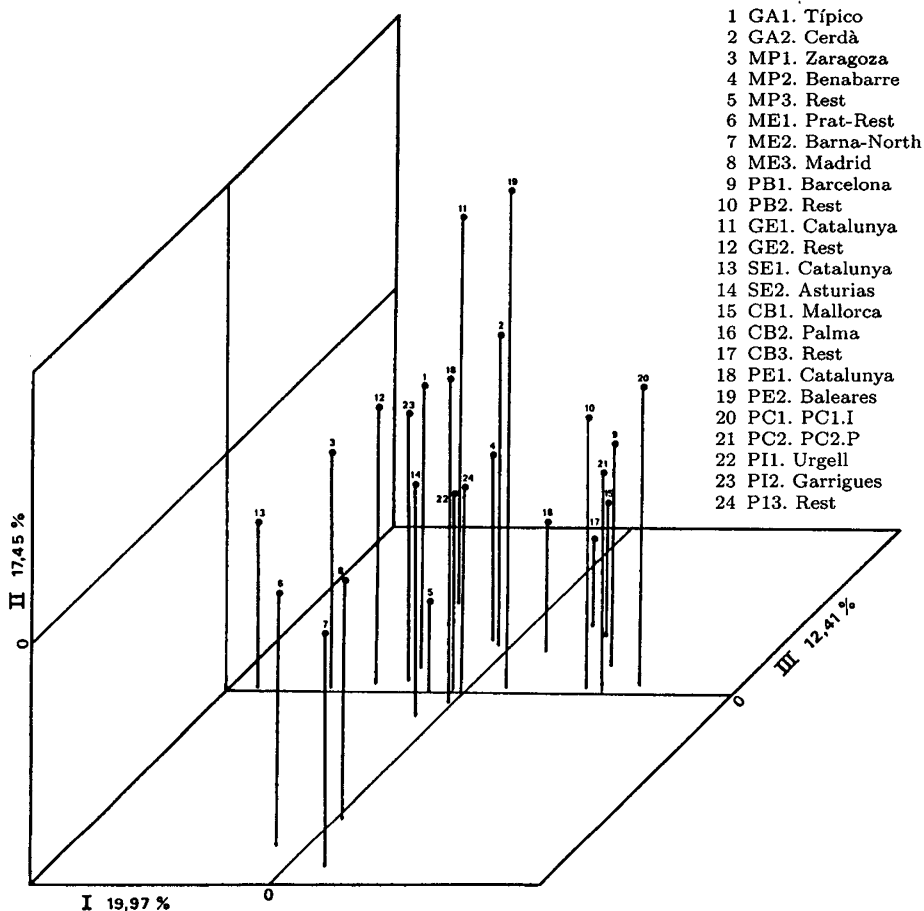


Fig 2. Factor analysis of principal components with 3 factors, at the subpopulation level. The gene frequencies of 11 polymorphic systems were used as variables.

Sokal, 1973) was applied to the distance matrices obtained by using Nei's (1978) index and Cavalli-Sforza and Edwards' (1967) chord distance, respectively.

Nei's (1978) genetic distance among breeds and identity values are shown in table IV. Distance values range between $D = 0.000$ for the Gos d'Atura-Podenco Ibérico pair, and $D = 0.051$ for the Mastín Español-Ca de Bestiar pair. The average value of between-breed distance is $0.0197 (\pm 0.0128)$. The Ca de Bestiar shows, in general, distance values with regard to the other breeds that are much higher than the average of the between-breed comparisons. The phenograms obtained by cluster analysis are shown in figures 3 and 4.

The formation of 2 large clusters is observed: Perdiguero de Burgos and Podenco Canario, and the one formed by the rest of the breeds, except Ca de Bestiar, which separates from the hypothetical common trunk very early. Within the second

Table IV. Nei's distance (1978) (below diagonal) and identity values (above diagonal) for Spanish dog breeds. Within parentheses, associated standard errors of the distance.

Breeds	GA	MP	ME	PB	GE	SE	CB	PE	PC	PI
GA	*****	0.990	0.989	0.985	0.994	0.995	0.969	0.990	0.985	1.000
MP	0.010 (0.022)	*****	0.981	0.975	0.988	0.997	0.966	0.967	0.971	0.991
ME	0.011 (0.023)	0.020 (0.030)	*****	0.981	0.986	0.983	0.950	0.989	0.983	0.989
PB	0.015 (0.027)	0.025 (0.035)	0.020 (0.030)	*****	0.975	0.968	0.969	0.978	0.999	0.987
GE	0.006 (0.017)	0.012 (0.024)	0.014 (0.026)	0.025 (0.035)	*****	0.991	0.965	0.988	0.977	0.994
SE	0.005 (0.015)	0.003 (0.012)	0.017 (0.029)	0.032 (0.040)	0.009 (0.021)	*****	0.958	0.978	0.967	0.995
CB	0.031 (0.039)	0.034 (0.041)	0.051 (0.050)	0.031 (0.039)	0.035 (0.042)	0.043 (0.046)	*****	0.951	0.962	0.968
PE	0.010 (0.022)	0.033 (0.040)	0.011 (0.023)	0.023 (0.033)	0.012 (0.024)	0.023 (0.033)	0.050 (0.049)	*****	0.980	0.992
PC	0.015 (0.027)	0.030 (0.038)	0.017 (0.029)	0.001 (0.007)	0.023 (0.033)	0.034 (0.040)	0.038 (0.043)	0.020 (0.031)	*****	0.985
PI	0.000 (0.000)	0.009 (0.021)	0.011 (0.023)	0.013 (0.025)	0.006 (0.017)	0.005 (0.015)	0.032 (0.040)	0.008 (0.020)	0.015 (0.027)	*****

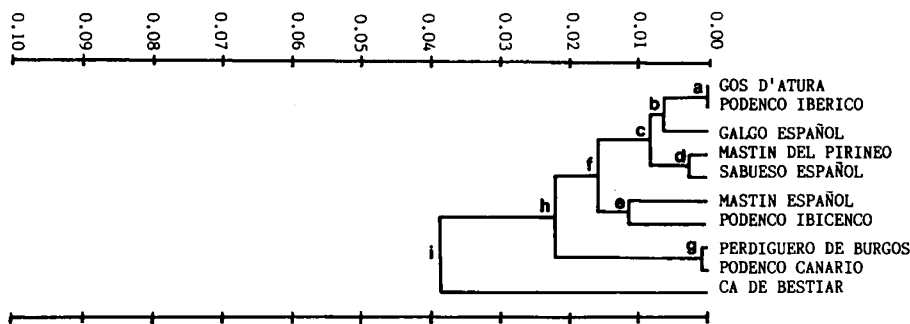


Fig 3. Phenogram obtained by the UPGMA method from Nei's distance (1978). Goodness-of-fit statistics of 0.88 for the cophenetic correlation, 0.20 for Farri's *F* (1972), 23.06 for Prager and Wilson's (1976) *F*, and 34.61% standard deviation.

group, Mastín Español and Podenco Ibicenco would be more related, perfectly differentiated from the other members of the cluster, and forming in their turn a new one. Within the last cluster, 2 new groups would form; on the one hand Mastín del Pirineo and Sabueso Español, and on the other hand, Gos d'Atura and Podenco Ibérico with Galgo Español.

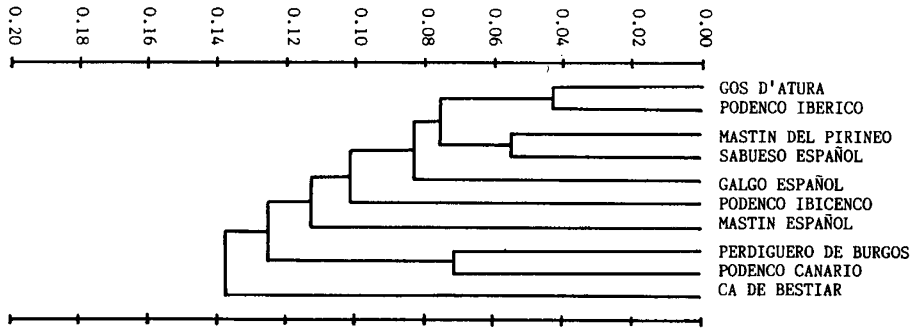


Fig 4. Phenogram obtained by the UPGMA method from Cavalli-Sforza and Edwards' (1967) chord distance. Goodness-of-fit statistics of 0.91 for the cophenetic correlation, 0.39 for Farris' F (1972), 7.81 for Prager and Wilson's (1976) F , and 9.94% of standard deviation.

According to Nei *et al* (1985), when the identity values (I) are higher than 0.9 for most pairs of populations and the average of heterozygosity (H) is high (higher than 0.1), as it is in this case, an overestimation of the values of the variances of the distances is produced. For this reason the distance between breeds has been calculated by a jackknife method (Mueller and Ayala, 1982) in an attempt to correct this bias. The average value obtained by this last method for interracial distance is 0.025 9 (\pm 0.016 8). The topology of the tree is identical to the topology obtained before by using standard distance values.

As it has been said before, a way to evaluate the stability of the phenograms obtained from Nei's index is to compute the standard errors (SE) at every bifurcation point of the tree branches (Nei *et al*, 1985). Our results show that the (SE) of all bifurcation points are considerably greater than the length of the branch. This implies that any relationship among OTUs (operative taxonomic units) would be possible within the tree. The same conclusion is reached by using jackknife values in the intra- and intercluster comparisons.

Nevertheless, this is not the only criterion to check the stability of a classification, because a classification can be considered as stable if its topology is not altered when new characters and/or new OTUs are included, or when different algorithms of taxonomic resemblance are used (Sokal *et al*, 1984). In this way, figure 5 shows the relationships among subpopulations. The topology of this tree is nearly the same as the topology obtained at the breed level, with the exception of 3 subpopulations: MP2, PE2 and SE2. Nei's (1978) average intersubpopulational distance is 0.0206 (\pm 0.0149), the average distance between subpopulations that belong to the same breed being 0.0068 (\pm 0.0087). The average within-breed distance (table V) takes the values of 0.023 for Mastín del Pirineo, 0.019 for Podenco Ibicenco and 0.015 for Sabueso Español. In the rest of the breeds these values range between 0.000 and 0.005, showing that the genetic differentiation among subpopulations of the same breed is nearly null.

When Wagner's method (Farris, 1972) is applied to the chord distance values of Cavalli-Sforza and Edwards (1967), the cladogram of figure 6 is obtained. The central criterion of this method is that of "parsimony", having the "maximum

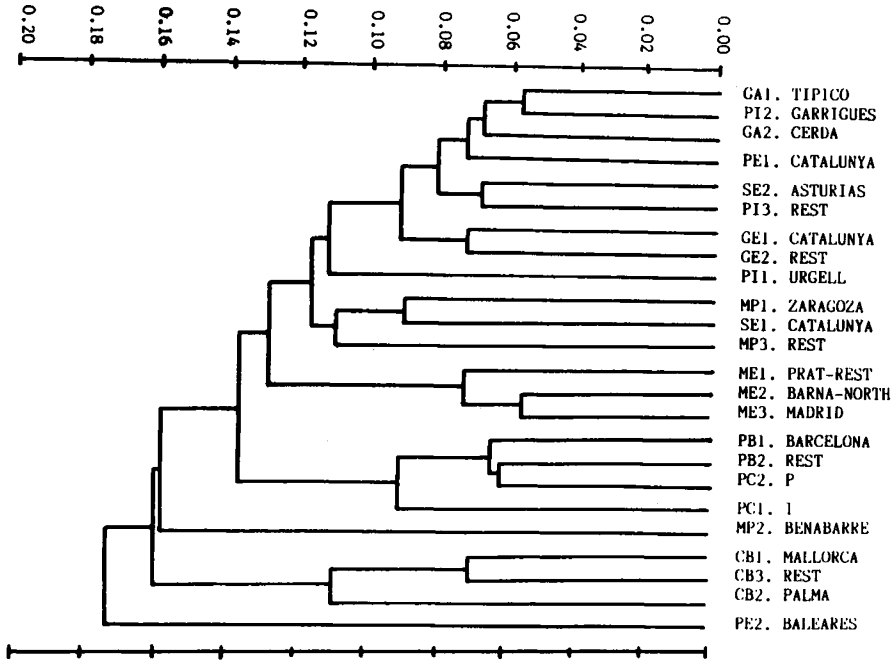


Fig 5. Phenogram obtained by the UPGMA method from Cavalli-Sforza and Edward's (1967) chord distance, for the 24 subpopulations into which the Spanish dog breeds have been subdivided.

parsimony" when all the OTUs with the minimum possible distance are related. The cladogram is topologically similar to the previous phenograms, which would corroborate the stability of the classification proposed.

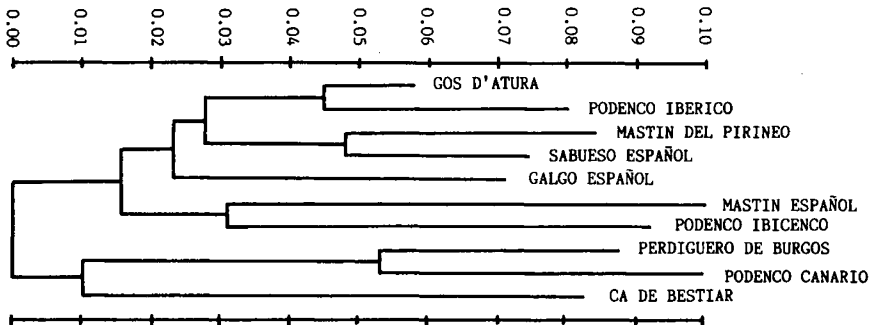
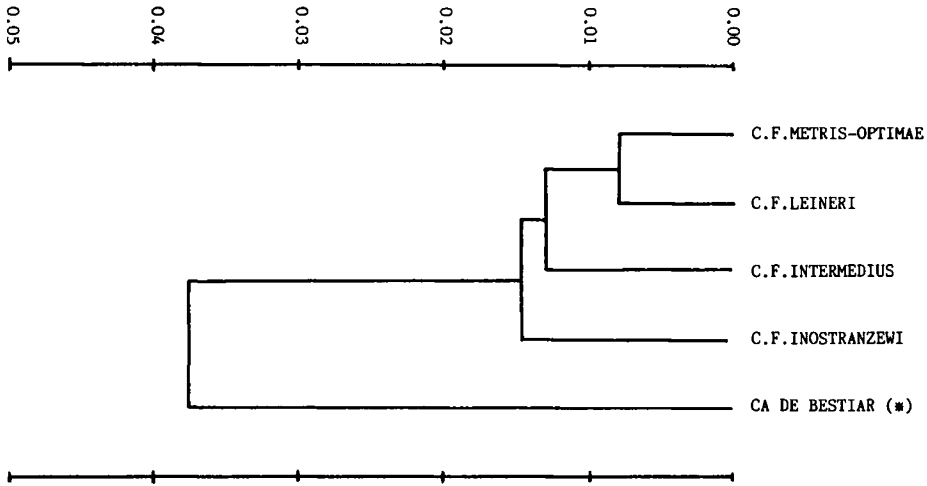


Fig 6. Cladogram obtained by the distance-Wagner method (Farris, 1972), by using the multiple addition criterion (Swofford, 1981), and rooted at the midpoint of the longest path. The Cavalli-Sforza and Edwards' (1967) chord distance was used as the measure of distance.

Table V. Average values of genetic distance (Nei, 1978) among subpopulations of the same breed (on diagonal) and among subpopulations of different breeds (below diagonal). Within parentheses, extreme distance values among subpopulations.

<i>Breeds</i>	<i>Sub</i>	1	2	3	4	5	6	7	8	9	10	
1	GA	2	0.003 (0.003-0.003)									
2	MP	3	0.019 (0.005-0.040)	0.023 (0.007-0.041)								
3	ME	3	0.011 (0.003-0.018)	0.031 (0.012-0.061)	0.001 (0.000-0.002)							
4	PB	2	0.015 (0.009-0.024)	0.034 (0.021-0.080)	0.020 (0.013-0.030)	0.002 (0.002-0.002)						
5	SE	2	0.011 (0.001-0.020)	0.017 (0.006-0.035)	0.022 (0.001-0.044)	0.039 (0.017-0.063)	0.019 (0.019-0.019)					
6	GE	2	0.005 (0.000-0.010)	0.017 (0.010-0.025)	0.012 (0.000-0.022)	0.025 (0.022-0.029)	0.013 (0.000-0.028)	0.000 (0.000-0.000)				
7	PE	2	0.017 (0.002-0.034)	0.048 (0.022-0.075)	0.015 (0.000-0.036)	0.028 (0.016-0.040)	0.035 (0.003-0.083)	0.016 (0.002-0.028)	0.015 (0.015-0.015)			
8	PC	2	0.014 (0.003-0.023)	0.038 (0.018-0.057)	0.016 (0.011-0.025)	0.002 (0.000-0.005)	0.040 (0.013-0.074)	0.021 (0.018-0.025)	0.000 (0.013-0.034)	0.000 (0.000-0.000)		
9	PI	3	0.001 (0.000-0.005)	0.017 (0.000-0.037)	0.009 (0.000-0.017)	0.013 (0.004-0.023)	0.010 (0.000-0.024)	0.004 (0.000-0.011)	0.012 (0.000-0.026)	0.014 (0.002-0.028)	0.001 (0.000-0.004)	
10	CB	3	0.034 (0.025-0.045)	0.040 (0.016-0.055)	0.052 (0.044-0.059)	0.035 (0.023-0.049)	0.051 (0.027-0.076)	0.035 (0.034-0.037)	0.059 (0.36-0.081)	0.042 (0.029-0.057)	0.032 (0.018-0.056)	0.005 (0.001-0.009)

When the different breeds are grouped within their hypothetical ancestral trunks (Jordana *et al*, 1992) by means of a hierarchical analysis taking the breeds as OTUs (Swofford and Selander, 1981), a matrix of distances among ancestral trunks is computed, obtaining an average value of intertrunk distance of 0.022 8 (\pm 0.013 3). The resultant phenogram (fig 7) shows a well-defined cluster that includes *Cf metris-optimae* and *Cf leinieri*; *Cf inostranzewi* and *Cf intermedius* join afterwards, forming in their turn a new cluster, leaving Ca de Bestiar clearly separated from it (this breed, due to its particular formation (Guasp, 1982), has not been assigned to any specific ancestral trunk).



(*) Ca de Bestiar breed, due to its particular formation, it has not been assigned to any specific trunk

Fig 7. Phenogram generated by the UPGMA method using Nei's (1978) distance values, obtained by a hierarchical analysis.

DISCUSSION

Genetic differentiation among populations

In this study, the average distance values among subpopulations (0.0206), among breeds (0.019 7), and among ancestral trunks (0.022 8) do not substantially differ from one another. These values are in the range of distances indicated by Nei (1987) for local breeds. It could suggest that there is not enough genetic differentiation among the so-called ancestral trunks to give them the taxonomic rank of subspecies.

From the comparison between tables IV and V, in some cases there is more genetic differentiation between subpopulations of the same breed than between different breeds. This would be the case of the Mastín del Pirineo, Sabueso Español

and Podenco Ibicenco breeds. Similar situations have been described in other domestic species (Vallejo *et al*, 1979; Ordás and San Primitivo, 1986). Nei and Roychoudhury (1982) also point out that the genetic variation among the 3 major human races is sometimes smaller than the variation among subpopulations of the same race.

Theoretically, the divergence between 2 populations can be the result of one or more causes: mutation, geographical and reproductive isolation, natural and/or artificial selection and genetic drift, so it is difficult to determine precisely the possible factors causing the observed within-breed differentiation in Mastín del Pirineo, Sabueso Español and Podenco Ibicenco. Nevertheless, genetic drift could be the factor that has contributed the most to the observed within-breed differentiation, owing to the low effective population size in the subpopulations studied. Besides that, in most domestic species the drift process is accelerated, because both sexes are not equally represented, which is especially common in dogs.

Congruence between enzymatic and morphological phylogenies

So far, the different breeds have been classified into their respective hypothetical ancestral trunks by using mainly dental and cranial morphology, and historical and behavioral comparative criteria (Studer, 1901; Antonius, 1922; Villemont *et al*, 1970; Rousselet-Blanc, 1983).

In a first attempt (Jordana *et al*, 1992) the Spanish dog breeds can be classified into several groups: Mastín del Pirineo with Mastín Español, both from the *Cf inostranzewi*; Podenco Ibicenco, Podenco Canario, Podenco Ibérico and Galgo Español would form another group, with *Cf leineri* as common ancestor. Historical data point out that Perdiguero de Burgos was formed in the intersection points of Sabuesos and Pachones, so it would be grouped with Sabueso Español in the common trunk of *Cf intermedius*, while Gos d'Atura would be the only representative of *Cf metris-optimae*. Finally, Ca de Bestiar, due to a quite uncertain origin – even though most authors (Guasp, 1982; Sotillo and Serrano, 1985; Delalix, 1986) impute its origin to crossings among Podencos, Mastiffs and Perdigueros – has not been assigned to any particular ancestral trunk. The phylogenies resulting from the qualitative and quantitative analysis of morphological data confirm this classification (Jordana *et al*, 1992).

All the enzymatic phylogenies evaluated are similar which supports the stability of the classification obtained using electrophoretic data. Nevertheless, these phylogenies show some differences from the phylogenies obtained using morphological data (Jordana *et al*, 1992).

By excluding the subpopulation PE2 Baleares from the analysis, a great similarity is observed between enzymatic and morphological phylogenies. This subpopulation was shown to differentiate clearly from all the other subpopulations (see fig 5). With this exclusion (fig 8), the relationships between the Greyhounds (Podencos and Galgo) and Gos d'Atura, and between the Mastines and Sabuesos are more obvious, in a way similar to the morphological analysis.

The breeds whose position shows less congruence with the morphological classification are Ca de Bestiar, Podenco Canario and Perdiguero de Burgos, which form a well defined and separated cluster in the cladogram generated using Wagner's

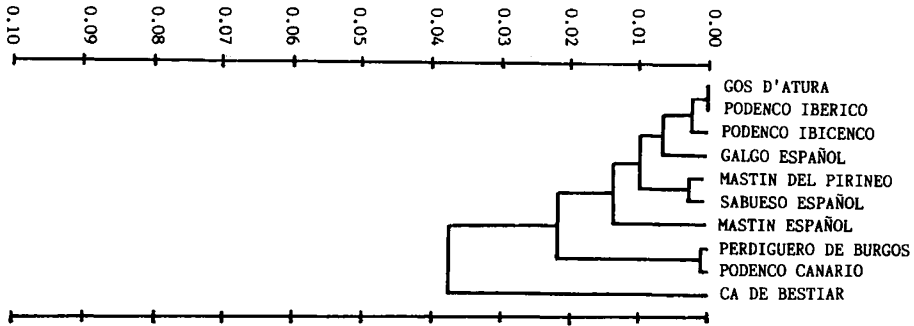


Fig 8. Phenogram obtained by the UPGMA method from Nei's (1978) distance, when PE2 Baleares subpopulation of Podenco Ibicenco breed is excluded from the study.

method (fig 6). It is not very probable that these 3 breeds had a common origin, so the explanation for their location in the phylogenetic tree should be searched for in their respective population structures. Studies done on these 3 breeds, referring to the levels of genetic variability (Jordana *et al*, 1991), have shown that these breeds have suffered important "bottlenecks" throughout their history. As a consequence, the genetic distance estimates relative to the other breeds could be more influenced by genetic drift, due to a small population size, than by the real divergence time among them.

In observing the values of distance found with respect to the other breeds and the topology of the trees, this hypothesis is strengthened. It is known that when a population is under the effects of a bottleneck, genetic distances increase quickly (Nei and Roychoudhury, 1982; Nei, 1987). This increase of genetic distances distorts the topology of the evolutionary trees. Besides that, their own history confirms this hypothesis. In the Ca de Bestiar breed we could even assume a founder effect, because this breed had nearly disappeared in the sixties, starting its recovery in the seventies from only 4 males and 2 females (Guasp, 1982).

Similar discordances in the interpretation of the evolutionary trees in other species, due to bottlenecks, have been described by Nei and Roychoudhury (1982) in human races, by Chesser (1983) in *Cynomys ludovicianus*, or Black-Tailed Prairie Dog, and by Gyllensten *et al* (1983) in the Red European Deer, among others.

Nei and Roychoudhury (1982), in their study of human races, support the hypothesis proposed by King and Wilson (1975), that macromolecular and anatomical characteristics of the organisms evolve at independent rates. The faster evolutionary change of the morphological characters is produced by a few gene substitutions, the genes that control these characters being under stronger natural selection in the process of human racial differentiation than the "average of genes". Nevertheless, they are more sceptical about a possible evolutionary parallelism between both types of characters, because they proved that the genetic distances among populations are not always correlated with the morphological differences.

Wayne and O'Brien (1986) found a non-significant correlation of $r = 0.24 \pm 0.1$ in comparing genetic and morphological distances in 15 inbred mouse strains, and concluded that structural gene and morphometric variation of mandible traits are

uncoupled between mouse strains. Fitch and Atchley (1987) also concluded that there was no correlation between distances based on single loci and mandible shape, in a study of the divergence in inbred strains of mice. Similarly, Crouau-Roy (1990) found no congruence between biochemical and morphological data in a study of 3 species of troglobitic beetles.

Festing and Roderick's (1989) results, however, are in contrast with the results obtained by Wayne and O'Brien (1986). In a study involving 12 inbred strains of mice, Festing and Roderick (1989) found strong and statistically highly significant correlations among all measures of genetic distance, ranging from 0.58 for the comparison of single loci with the logarithm of the Mahalanobis distance based on 24 measurements on 4 bones, to 0.72 for estimates of genetic distance based on single loci and the morphology of the mandible.

Wayne and O'Brien (1987) in a study of the enzymatic divergence in 12 genera of the Canidae family, affirm that, in general, qualitative and quantitative morphologic studies of the Canidae (Clutton-Brock *et al*, 1976; Wayne, 1986) support the groupings represented in the consensus tree they obtain from enzymatic data.

The groupings of Spanish dog breeds obtained in our study from morphological and biochemical data were apparently quite similar, particularly for the populations that have not been long under bottleneck effects. However, the correlation between morphological and enzymatic distances was low ($r = 0.07$), and non-significant, even excluding from the calculations the populations that suffered strong bottlenecks.

From this study, it can be concluded that the large morphological variability among dog breeds, where the process of differentiation has been strongly accelerated by a great pressure of selection on some characters, has no correspondence with differences at the protein and enzymatic levels, where the genetic differences among breeds are very small. This is in accordance with Wayne (1986), who affirms that the domestic dog (*Canis familiaris*) is a group which is morphologically diverse but genetically very homogeneous.

Estimated times of evolutionary divergence

According to the neutral theory (Kimura, 1983), it can be assumed that there is a correlation between evolutionary time and genetic divergence measured by an index of distance such as that of Nei. Also, assuming that through electrophoretic techniques it is possible to detect a third of the amino-acid substitutions in the proteins, the following formula allows us to obtain approximately the time of divergence between 2 populations (Nei, (1987):

$$t = 3.75 \cdot 10^6 \cdot D$$

There are 2 important factors that can distort this estimation. The first is migration between populations, which produces an underestimation of the times of divergence. Migration can be neglected with regard to the dog breeds, because the populations - breeds - were closed shortly after their formation. The second is occurrence of bottlenecks, which have a great influence upon the values of the distance with a subsequent overestimation of divergence times. Taking these

considerations into account, only (t) values among different ancestral trunks have been estimated. We assume that the errors in the calculation of the distances for the populations affected by bottlenecks are diluted when these breeds are included in their ancestral trunk. The Ca de Bestiar breed, however, has not been used in the calculation of the times of divergence because, on the one hand, it has not been assigned to any particular ancestral trunk, and, on the other hand, it has suffered an extreme founder effect.

The divergence between *Cf metris-optimae* and *Cf leineri* would have taken place approximately 30 000 years ago. These 2 trunks would have separated from the common cluster that formed with *Cf intermedius* 49 000 years ago, while *Cf inostranzewi* would have separated 55 000 years ago from the cluster that forms with the other 3 ancestral trunks.

Nei and Roychoudhury (1982) assign a similar time of divergence of the separation from the common trunk of the Caucasoid and Mongoloid human races, approximately 41 000 years ago. Negroids would have separated from the common trunk with Caucasoid and Mongoloid approximately 110 000 years ago.

Nevertheless, it must not be overlooked that these divergence times are only indicative, because the associated errors of the distances are fairly large and the estimates depend upon several assumptions. In our case, the divergence times would be overestimated due to the bias implied in the choice of the loci analyzed, because most of the known enzymatic polymorphic loci have been included deliberately. As a consequence, the true divergence times should be lower in magnitude than those presented in this paper.

ACKNOWLEDGMENTS

The suggestions of the referee are greatly appreciated. We thank J Torrent and C Simmons for assistance with the preparation of this manuscript. Finally, the authors thank Gallina Blanca Purina for its initial contribution to the funding of this study.

REFERENCES

- Antonius O (1922) *Grundzüge einer Stammesgeschichte der Haustiere*. Gustav Fischer, Jena
- Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis models and estimation procedures. *Am J Hum Genet* 19, 233-257
- Chesser RK (1983) Genetic variability within and among populations of the Black-tailed prairie dog. *Evolution* 37, 320-331
- Clutton-Brock J, Corbet GB, Hills M (1976) A review of the family Canidae, with a classification by numerical methods. *Bull Br Mus (Nat Hist)*. Zool 29, 119-199
- Crouau-Roy B (1990) Evolutionary systematic of three species of troglolitic beetles: electrophoretic and morphological evidence. *Genet Sel Evol* 22, 189-203
- Delalix AM (1986) *Los Perros Españoles*. De Vecchi, Barcelona
- Farris JS (1972) Estimating phylogenetic trees from distance matrices. *Am Nat* 106, 645-668

- Festing MFW, Roderick TH (1989) Correlation between genetic distances based on single loci and on skeletal morphology in inbred mice. *Genet Res (Camb)* 53, 45-55
- Fitch WM, Atchley WR (1987) Divergence in inbred strains of mice: a comparison of three different types of data. In: *Molecules and Morphology in Evolution: Conflict or Compromise* (Patterson C, ed) Cambridge University Press, Cambridge, 203-216
- Frane J, Jennrich R, Sampson P (1985) P4M factor analysis. In: *BMDP, Statistical Software* (Dixon WJ, ed) University of California Press, CA
- Guasp A (1982) Pastor Mallorquín o Ca de Bestiar. I Symp Nac Razas Caninas Españolas, Córdoba, 19-21 March 1982, 45-51
- Gyllensten U, Ryman N, Reuterwal C, Dratch P (1983) Genetic differentiation in four European subspecies of red deer (*Cervus elaphus* L). *Heredity* 51, 561-580
- Hedges SB (1986) An electrophoretic analysis of holarctic hylid frog evolution. *Syst Zool* 35, 1-21
- Jordana J (1989) Relaciones Genéticas en Cánidos Españoles. Tesis Doctoral, Facultad de Veterinaria, Universidad Autónoma de Barcelona, Barcelona
- Jordana J, Piedrafitra J, Sánchez A (1991) Variabilidad genética en diez razas caninas españolas. *Arch Zootec* 40, 147, 115-128
- Jordana J, Piedrafitra J, Sánchez A (1992) Genetic relationships in Spanish dog breeds. I. The analysis of morphological characters. *Genet Sel Evol* 24 (in press)
- Juneja RK, Christensen K, Andresen E, Gahne B (1981) Frequencies of transferrin types in various breeds of domestic dogs. *Anim Blood Groups Biochem Genet* 12, 79-88
- Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge
- King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science* 188, 107-115
- Kobayashi R, Miyakawa H, Tanabe Y, Hashimoto Y, Ota K, Faruque MO (1987) Blood protein polymorphism in the Bangladesh native dogs. *Genet Stud Breed Differentiation Native Dom Anim Bangladesh* 2, 93-103
- Leone CA, Anthony RL (1966) Serum esterases among registered breeds of dogs as revealed by immunoelectrophoretic comparisons. *Comp Biochem Physiol* 18, 359-368
- Mueller LD, Ayala FJ (1982) Estimation and interpretation of genetic distance in empirical studies. *Genet Res (Camb)* 40, 127-137
- Nei M (1972) Genetic distance between populations. *Am Nat* 106, 283-292
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583-590
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York
- Nei M, Roychoudhury AK (1982) Genetic relationship and evolution of human races. *Evol Biol* 14, 1-59
- Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J Mol Evol* 19, 153-170
- Nei M, Stephens JC, Saitou N (1985) Methods for computing the standard errors of branching points in an evolutionary tree and their application to molecular data from humans and apes. *Mol Biol Evol* 2, 1, 66-85

- Ordás JG, San Primitivo F (1986) Genetic variations in blood proteins within and between Spanish dairy sheep breeds. *Anim Genet* 17, 3, 255-266
- Prager EM, Wilson AC (1976) Congruency of phylogenies derived from different proteins. A molecular analysis of the phylogenetic position of cracid birds. *J Mol Evol* 9, 45-57
- Rousselet-Blanc P (ed) (1983) *Nuestro Amigo el Perro*. Reader's Digest (Iberia), Madrid (translation from French)
- Sneath PHA, Sokal RR (1973) *Numerical Taxonomy*. WH Freeman, San Francisco
- Sokal RR, Fiala KL, Hart G (1984) OTU stability and factors determining taxonomic stability: examples from the Caminalcules and the Leptodomorpha. *Syst Zool* 33, 387-407
- Sotillo JL, Serrano V (1985) *Producción Animal. I. Etnología Zootécnica*. Tebas-Flores, Madrid
- Studer T (1901) Die prähistorischen Hunde in ihrer Beziehung zu den gegenwärtig lebenden Hunderassen. *Abh Schweiz Paleont Ges* 28, 1-137
- Sugiura S, Tanabe Y, Ota K (1977) Genetic polymorphism of eserine resistant esterases in canine plasma. *Anim Blood Groups Biochem Genet* 8, 121-126
- Swofford DL, Selander RB (1981) BIOSYS-1: A Fortran program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *J Hered* 72, 281-283
- Tanabe Y, Sugiura S, Asanoma M, Ota K (1974) Genetic polymorphism of leucine aminopeptidase in canine plasma. *Anim Blood Groups Biochem Genet* 5, 225-230
- Tanabe Y, Omi T, Ota K (1977) Genetic variants of glucose phosphate isomerase (EC 5.3.1.9) in canine erythrocytes. *Anim Blood Groups Biochem Genet* 8, 191-195
- Tanabe Y, Omi T, Ota K (1978) Genetic variants of hemoglobin in canine erythrocytes. *Anim Blood Groups Biochem Genet* 9, 79-83
- Vallejo M, Zarazaga I, Sierra I (1979) Clasificación etnológica de ovinos españoles. I. Elaboración genético - taxonómica en siete ecotipos de la Rasa Aragonesa. IV Jornadas Científicas, de la Sociedad Española de Ovinotecnia, Zaragoza, 7-9 June 1979, Zaragoza, 63-76
- Villemont M (1970) *Le Grand Livre du Chien*. Edito-Service, Geneva
- Wayne RK (1986) Cranial morphology of domestic and wild canids: the influence of development on morphological change. *Evolution* 40, 243-261
- Wayne RK, O'Brien SJ (1986) Empirical demonstration that structural gene and morphometric variation of mandible traits are uncoupled between mouse strains. *J Mammal* 67, 441-449
- Wayne RK, O'Brien SJ (1987) Allozyme divergence within the Canidae. *Syst Zool* 36, 339-355