Comparative F Statistics Analysis of the Genetic Structure of Ten Spanish Dog Breeds

J. Jordana, J. Piedrafita, A. Sanchez, and P. Puig

The genetic structure and relationships among 10 Spanish dog breeds have been studied by using F statistics. Data came from 21 structural genic loci that codify for blood-soluble proteins and enzymes detected by electrophoresis. Of the 21 loci, 11 were found to be polymorphic. The study was done at three levels of hierarchical differentiation: ancestral trunks, breeds, and subpopulations. The deficit of heterozygotes was estimated at the subpopulation, breed, and ancestral trunk levels, with values of 4.0%, 6.5%, and 11.2%, respectively. In the whole population, the deficit of heterozygotes was about 17%. The proportion of genetic variability attributable to differences between subpopulations, breeds, and ancestral trunks was estimated to be 14.2%, 9.9%, and 6.9%, respectively. The dendrogram, obtained by using values of genic differentiation (F_{ST}) as a measure of the genetic distance among populations, is topologically identical to the one obtained using Nei's index of distance, which indicates a high correlation (r = .99) between both distances. These racial groupings, however, differ from the grouping obtained from historical, archeological, and morphological data.

The genetic divergence among populations or breeds is usually quantified by two different measures: statistics of genetic distance and fixation indices, or F statistics (Wright 1965), modified by Nei (1977), and by Wright (1978). Fixation indices seem to be more appropriate than genetic distances when studying differentiation among subpopulations of the same or different breeds.

The fixation indices in Wright's terminology are F_{IS}, F_{IT}, and F_{ST} (Wright 1965). F_{IS} and F_{IT} stand for the correlations between two uniting gametes drawn at random from a subpopulation and from the total population, respectively, whereas F_{ST} is the correlation between two gametes randomly drawn from each subpopulation. These parameters are related through the following expression:

$$1 - F_{IT} = (1 - F_{IS})\cdot(1 - F_{ST}). \quad (1)$$

This equation assumes neutral diallelic loci. Nei (1977) and Wright (1978) extended the treatment to multiple alleles. In particular, Nei (1977) concluded that all the fixation indices can be defined using the observed and expected heterozygosities at the equilibrium of the populations under study. F_{IS} and F_{IT} statistics measure the excess or the deficit of the average heterozygotes in each subpopulation and in the population as a whole, respectively. In each case these statistics can have negative or positive values. In turn, the F_{ST} statistic measures the degree of genetic differentiation among populations. Its value is always positive and lies between zero and 1, so it may be used as a measure of genetic distance among populations (Gregorius and Roberds 1986; Long 1986; Long et al. 1987; Weir 1990).

It is possible to extend the analysis to additional hierarchical layers of population nesting. If the subpopulations are grouped within clusters, then the extension of equation (1) is

$$1 - F_{IT} = (1 - F_{IS})\cdot(1 - F_{SC})\cdot(1 - F_{CT}),$$

where F_{IS} and F_{IT} are defined as before, F_{CT} is the correlation of gametes within a cluster, relative to that of gametes drawn at random from the total population, and F_{SC} is the correlation of gametes within subpopulations, relative to that of gametes drawn at random from the cluster (Smouse and Long 1988).

The canine species has a great number of officially recognized breeds (about 400). But in spite of the considerable morphological variability that exists among them, several questions might be raised. What proportion of total genic variation could be attributed to racial differences? Are two

From the Unitat de Gènetica i Millora Animal, Departament de Patologia i de Producció Animals, Facultat de Veterinària (Jordana, Piedrafita, Sanchez) and the Departament de Matemàtiques (Puig), Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. We would like to thank Gallina Blanca Purina for its initial contribution to the financing of this study, and also several veterinarians and dog breeders who kindly collaborated in obtaining blood samples. We would also like to thank Joan Torrent and Chuck Simmons for assistance with the preparation of this manuscript. The suggestions of two referees are greatly appreciated.

Journal of Heredity 1992;83:367–374; 0022-1503/92/$4.00

breeds more genetically different than two subpopulations of the same breed? What proportion of the observed total genic variation would be explained by the process of breed formation? Are the genetic relationships derived from estimates of F values consistent with those evaluated by some other methods? Obtaining answers to these questions is the aim of this article.

To achieve the objectives stated above we analyze and quantify the inter-, and intraracial genetic differentiation among 10 Spanish dog breeds by using F statistics (Nei 1977; Wright 1978). Additionally, the genetic relationships obtained by using those F statistics as a measure of genetic distance will be compared with the genetic relationships obtained by applying Nei's (1978) index of genetic distance.

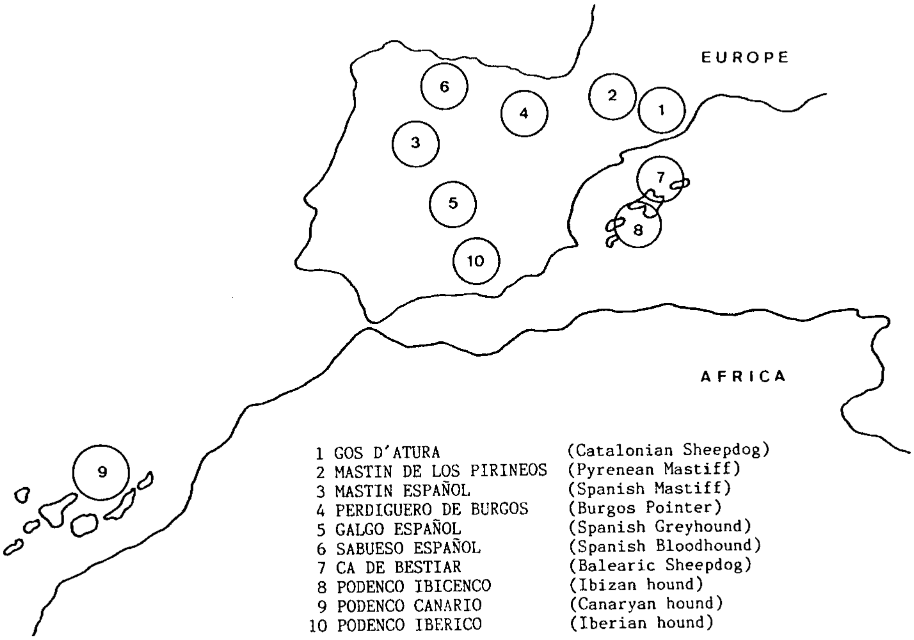

Figure 1. Geographical location of 10 Spanish dog breeds.

Materials and Methods

Breeds

Blood samples from 484 dogs were taken from 10 Spanish dog breeds: Gos d'Atura (93), Mastín de los Pirineos (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). All breeds except the Podenco Ibérico are officially recognized by the Féderation Cynologique Internationale (FCI). Individuals were assigned to their breeds on the basis of their morphologic similarity with the standard of the breed (Gómez-Toldrà 1985), and, according to the available information, sampled individuals were unrelated. The geographical distribution of these breeds is shown in Figure 1. There exist several hypotheses about their origin, which we summarize in the following way (Jordana et al. 1990).

Gos d'Atura (Catalonian Sheepdog). Descendant of the old shepherd dogs (probably the Bergamasco breed), which the Roman legions brought to Spain in their campaigns. The origin of the Bergamasco dog is the Polish shepherd dog, which might have descended from the old eastern shepherds (Andreu 1984; Delalix 1986; Gómez-Toldrà 1985).

Mastín Español and Mastín de los Pirineos (Spanish Mastiff and Pyrenean Mastiff). These breeds are included in the "ortognated moloses" which seem to descend from the legendary Mastiff of Tibet (in Central Asia). These dogs are supposed to have reached Spain by two routes: the Central-European route and via the Mediterranean (Esquiró 1982).

Perdiguero de Burgos (Burgos Pointer). Probably originated from matings between the Sabueso Español and the short-coated "Pachones" from Navarra (Boivin et al. 1983; Delalix 1986; Gómez-Toldrà 1985; Sanz Timón 1982). These "Pachones" from Navarra, also called "Perros de Punta Ibéricos," are the ancestors of the current English Pointer (Boivin et al. 1983; Sotillo and Serrano 1985).

Sabueso Español (Spanish Bloodhound). Several authors (Boivin et al. 1983; Gómez-Toldrà 1985; Gondrexon and Browne 1982; Villemont 1970) have attributed to the bloodhounds a Celtic origin. Most of the European bloodhound breeds seem to descend from the Saint Hubert, a modern Belgian breed, the direct descendant of the *Segusius* of the Celts and the Gauls.

Ca de Bestiar (Balearic Sheepdog). Several authors (Delalix 1986; Guasp 1982; Sotillo and Serrano 1985) agree that the origin of this breed seems to be the result of crossings between Podencos Ibicencos, Perdigueros (Ca Mé), and Mastiffs.

Galgo Español (Spanish Greyhound). For some authors (Sotillo and Serrano 1985; Villemont 1970) the English Greyhound and the Galgo Español are descendants of the Arabian Sloughi, brought to Europe via Spain during the Moslem invasion. Another hypothesis (Boivin et al. 1983) supports the idea that the Galgo was brought to Western Europe by the ancient Celts when they settled in Gaul. Nevertheless, the same author points out a second contribution of blood from the Sloughi.

Podenco Ibicenco (Ibizan Hound). It is generally accepted that the Podenco Ibicenco breed descends from the "Dog of the Pharaohs" (Boivin et al. 1983; Gómez-Toldrà 1985; Gondrexon and Browne 1982; Mora 1982; Villemont 1970) and that it was brought to Ibiza by the Phoenicians (Delalix 1986; Maza 1982; Pugnetti 1981), even though other hypotheses state that its arrival took place much later, with the Moslems, at the same time as the Galgo (Boivin et al. 1983; Villemont 1970).

Podenco Canario (Canary Hound). Certain hypotheses (Delalix 1986) suppose that this hunter came from Egypt and that it was taken to the Canary Islands, probably by the Phoenicians, Greeks, Carthaginians, or even the Egyptians, but it is possible that Majorcan monks, forced to emigrate to these islands by the Vatican, introduced these dogs (Anonymous 1982).

Podenco Ibérico (Iberian Hound). Its origin might be the same as that of the Podenco Ibicenco. This breed later fused with the Podenco Rondeño from Andalusia, which originated the breed we know at present (García et al. 1982).

Samples

We collected blood samples with EDTA 2Na (1 mg per ml of blood) as an anticoagulant. The samples were separated into the three main blood components—plasma, red blood cells, and white blood cells—and stored at −20°C.

We analyzed 21 loci, according to the

Table 1. Assumed hierarchical structure of Spanish dog populations[a]

Ancestral trunks		Breeds		Subpopulations	
1. C. F. Metris-Optimae	(93)	1.1. Gos d'Atura	(93)	GA1.Tipico	(59)
				GA2.Cerdà	(34)
2. C. F. Inostranzewi	(100)	2.1. Mastín de los Pirineos	(55)	MP1.Zaragoza	(36)
				MP2.Benabarre	(11)
				MP3.Rest	(08)
		2.2. Mastín Español	(45)	ME1.Prat-Rest	(15)
				ME2.Barna-North	(16)
				ME3.Madrid	(14)
3. C. F. Intermedius	(95)	3.1. Perdiguero de Burgos	(42)	PB1.Barcelona	(19)
				PB2.Rest	(23)
		3.2. Sabueso Español	(53)	SE1.Catalunya	(25)
				SE2.Asturias	(28)
4. C. F. Leineri	(150)	4.1. Galgo Español	(31)	GE1.Catalunya	(21)
				GE2.Rest	(10)
		4.2. Podenco Ibicenco	(71)	PE1.Catalunya	(41)
				PE2.Baleares	(30)
		4.3. Podenco Canario	(15)	PC1.	(15)
		4.4. Podenco Ibérico	(33)	PI1.Urgell	(10)
				PI2.Garrigues	(17)
				PI3.Rest	(06)
5. Ca de Bestiar[b]	(46)	5.1. Ca de Bestiar	(46)	CB1.Mallorca	(20)
				CB2.Palma	(13)
				CB3.Rest	(13)

[a] Sample size for each population is in parentheses.

[b] The Ca de Bestiar breed, due to its particular formation, has not been assigned to any specific ancestral trunk.

methods described in detail by Jordana (1989), by using electrophoretic techniques: horizontal electrophoresis in starch gel and polyacrylamide and agarose-polyacrylamide (bidimensional) gels. The loci analyzed coded for the following proteins and enzymes: superoxide dismutase (SOD), glucose phosphate isomerase (GPI), leucin aminopeptidase (LAP), 6-phosphogluconate dehydrogenase (6-PGD), mannose phosphate isomerase (MPI), phosphoglucomutase-1 (PGM_1), malate dehydrogenase soluble form (MDH_S), malate dehydrogenase mitochondrial form (MDH_M), glucose 6-phosphate dehydrogenase (G6PD), albumin (ALB), peptidase D (PEP-D), acid phosphatase (ACP), transferrin (TF), prealbumin (PR), Gc protein (GC), α_1B-glycoprotein (α_1B), α_1-protease inhibitor (PI-1), protease inhibitor-3 (PI-3), postalbumin-1 (PA-1), pretransferrin-1 (PRT-1), and pretransferrin-2 (PRT-2).

To perform the intraracial analysis of the populations, we divided the breeds into 23 subpopulations, according to geographical criteria and/or areas of influence of certain breeders. The a priori subdivision of the different breeds in subpopulations, as well as their grouping in their hypothetical ancestral trunks, is shown in Table 1. Assignation of the breeds to their ancestral trunks was based on historical, archaeological, morphological, and enzymatic data (Boivin et al. 1983; Jordana 1989; Tanabe et al. 1974; Villemont 1970). The Ca de Bestiar breed was not included in any particular ancestral trunk because it seems to be the result of recent crossings among several breeds belonging to different trunks (Delalix 1986; Guasp 1982).

The inter- and intraracial genetic differentiation among Spanish dog breeds was analyzed by using F statistics (Wright 1965) as modified by Nei (1977) and Wright (1978). The bias in the genotypic proportions due to small sample sizes was corrected following Levene (1949). The statistical significance of gene frequency differences among populations was tested for each locus by the chi-square test, $\chi^2 = 2NF_{ST}(k-1)$, with $(k-1)\cdot(s-1)$ degrees of freedom, where N is the total sample size, k is the number of alleles for the locus, and s is the number of subpopulations (Chesser 1983). The deviation of F_{IT} from zero was tested by the chi-square test (Hedrick 1985) $\chi^2 = NF^2$.

Computations were performed with the BIOSYS-1 package (Swofford and Selander 1981, 1989). The method of cluster analysis using the UPGMA algorithm (Sneath and Sokal 1973) was applied to the F_{ST} estimated values, used as genetic distances.

Because of the small sample sizes, we performed a permutational test to contrast statistically the interracial genetic differentiation measured by the F_{ST} statistic (see Table 5). Such a permutational test was based on the study of the F_{ST} statistical distribution in considering all the possible combinations of the $n + m$ individuals (n and m being the sample size of each of two breeds), grouped in two samples of n and m elements taken at random. Owing to the high number of possible combinations, we took 1,000 random combinations in a procedure similar to that of Lewis and Orav (1989), which consequently gave 1,000 F_{ST} values. The distribution obtained by simulation allowed us to test the null hypothesis (HO: $F_{ST} = 0$) against the alternative hypothesis (H1: $F_{ST} > 0$).

Results

A total of 38 electromorphs were identified whose distribution varied from one to five. Taking as a criterion of polymorphism that of 95%, 10 systems (GPI, 6-PGD, PGM-1, MDH-S, MDH-M, G6PD, ACP, PR, GC, and PI-3) were found to be monomorphic for all breeds. Only two breeds showed disagreement with the expected Hardy-Weinberg proportions in some loci. These populations were Sabueso Español for TF ($P < .01$) and PRT-1 ($P < .05$) systems, and Podenco Ibicenco for α_1-B ($P < .05$). The deficit of heterozygotes (D) was −0.382, −0.374, and −0.269, respectively. At the subpopulational level (the allele frequencies for each polymorphic locus and subpopulation are shown in Table 2), only five breeds—Gos d'Atura (GA), Mastín de los Pirineos (MP), Mastín Español (ME), Podenco Canario (PC), and Ca de Bestiar (CB) showed genetic equilibrium for all loci in all subpopulations. The rest presented some subpopulations with significant disagreement in some loci, particularly, PB1.Barcelona for MPI ($P < .01$); GE1.Catalunya for ALB ($P < .05$), SE2.Asturias for PRT-1 ($P < .05$), PE1.Catalunya for PI-1 ($P < .05$), PE2.Baleares for ALB ($P < .05$), PI1.Urgell for PA-1 ($P < .05$), and PI2.Garrigues for ALB ($P < .01$).

F Statistics Analysis

The intra- and interpopulational analyses through F statistics were done at five different levels, as follows.

Subpopulations versus the total. There exists a global deficit of heterozygotes of 4% as an average ($F_{IS} = 0.040$), in all loci as a whole, for each subpopulation. The deficit of heterozygotes in the population as a whole—Spanish dog breeds—equals 17.6% ($F_{IT} = 0.176$). The average of genetic differentiation among subpopulations is 14.2% ($F_{ST} = 0.142$), a highly significant value when all the loci are considered. For each particular locus the same level of significance is held, with the exception of the *Prt-2* and *Pep-D* loci, where the differences are not significant.

Breeds versus the total. There exists a deficit of heterozygotes of 6.5% as an average for each breed, whereas the deficit of heterozygotes for the whole population, 15.7%, is similar to the deficit in the analysis of subpopulations. The average of genetic differentiation among breeds equals 9.9% ($P < .001$). This value is smaller, as expected, than the value that represents the differentiation among subpopulations. The degree of differentiation among breeds is highly significant for all loci as a whole and for each particular locus, except for *Prt-2* and *Pep-D,* as in the former case.

Ancestral trunks versus the total. The average deficit of heterozygotes for each ancestral trunk is 11.2%, whereas the estimated value of the deficit of heterozygotes for the entire population remains about the same (17.4%); the degree of differentiation among ancestral trunks is 6.9% on average ($P < .001$).

Table 3 summarizes the three previous sections. F_{IS} values increase as the hierarchical level increases, suggesting that the different levels are a pool of subdivided populations. The "breed" level would be formed by differentiated subpopulations and the "ancestral trunk" level would be, in turn, composed of different breeds. Finally, F_{ST} decreases, as expected, as the hierarchical level increases—that is, in approaching the hypothetical canid ancestors the genetic divergence among them must be much smaller than, for example, the divergence among subpopulations of different breeds. These subpopulations may have differentiated mainly through mutation and genetic drift.

Subpopulations within breeds. Table 4 summarizes the genetic structure of the breeds from an analysis of their subpopulations. There is a significant genetic differentiation among subpopulations of the same breed for Gos d'Atura (GA) ($F_{ST} = 0.013$), Mastín de los Pirineos (MP) ($F_{ST} = 0.108$), Sabueso Español (SE) ($F_{ST} = 0.062$), Podenco Ibicenco (PE) ($F_{ST} = 0.054$), and Ca de Bestiar (CB) ($F_{ST} = 0.045$). The average of differentiation among subpopulations within the same breed (Table 3) is 3.7% ($F_{SC} = 0.037$).

Direct comparison among breeds. To calculate the degree of genetic differentiation among particular breeds, an analysis was done through F statistics taking breeds in couples (Table 5). The values of genic differentiation, F_{ST}, among Spanish dog breeds range from 0.3% for the Gos d'Atura–Podenco Ibérico (GA-PI) pair to 13.7% for the Ca de Bestiar–Podenco Ibérico (CB-PI) pair. Only two pairs of breeds, Gos d'Atura–Podenco Ibérico (GA-PI) and Perdiguero de Burgos–Podenco Canario (PB-PC), did not show statistically significant genetic differentiation between themselves. Differentiation in the rest of the pairs was significant at $P < .001$ with the exception of the Mastín de los Pirineos–Sabueso Español (MP-SE) pair, which was significant at $P < .05$.

The levels of significance of the F_{ST} statistic, when permutational techniques were applied, were similar to those found by the chi-square test for all the pairs in comparison. The exceptions were the Gos d'Atura–Galgo Español (GA-GE) and Galgo Español–Podenco Ibérico (GE-PI) pairs, which showed significant differences at $P < .01$), and Sabueso Español–Podenco Ibérico (SE-PI), which showed differences at $P < .05$.

Table 2. Gene frequencies estimated for each of 11 polymorphic loci in the 23 Spanish dog subpopulations[a]

	Subpopulations[b]										
Locus	GA1	GA2	MP1	MP2	MP3	ME1	ME2	ME3	PB1	PB2	GE1
Sod	(59)	(34)	(36)	(11)	(08)	(15)	(16)	(14)	(19)	(23)	(21)
A	0.992	0.941	0.958	0.955	0.688	0.800	0.938	0.893	1.000	1.000	0.929
B	0.008	0.059	0.042	0.045	0.313	0.200	0.063	0.107	0.000	0.000	0.071
Lap	(59)	(34)	(36)	(11)	(08)	(15)	(16)	(14)	(19)	(23)	(21)
A	0.975	0.956	0.986	1.000	1.000	1.000	1.000	1.000	1.000	0.978	0.929
B	0.025	0.044	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.022	0.071
C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mpi	(52)	(33)	(31)	(10)	(08)	(15)	(07)	(05)	(19)	(23)	(16)
A	0.952	0.833	0.758	0.550	0.813	0.600	0.714	0.800	0.789	0.848	0.656
B	0.048	0.167	0.242	0.450	0.188	0.400	0.286	0.200	0.211	0.152	0.344
Alb	(59)	(34)	(36)	(11)	(08)	(15)	(16)	(14)	(19)	(23)	(21)
S	0.585	0.559	0.639	0.136	0.500	0.767	0.656	0.571	0.895	0.870	0.405
F	0.415	0.441	0.361	0.864	0.500	0.233	0.344	0.429	0.105	0.130	0.595
Pep-D	(59)	(34)	(36)	(11)	(08)	(15)	(16)	(14)	(19)	(23)	(21)
A	0.839	0.941	0.958	1.000	0.938	0.933	0.844	0.929	0.921	0.891	0.881
B	0.161	0.059	0.042	0.000	0.063	0.067	0.156	0.071	0.079	0.109	0.119
Tf	(59)	(34)	(36)	(11)	(08)	(15)	(16)	(14)	(19)	(23)	(21)
A	0.000	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
B	0.508	0.368	0.694	0.318	0.438	0.367	0.219	0.321	0.079	0.130	0.452
C	0.483	0.618	0.306	0.682	0.563	0.433	0.531	0.500	0.921	0.848	0.548
D	0.000	0.000	0.000	0.000	0.000	0.200	0.250	0.179	0.000	0.022	0.000
E	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
α_1B	(59)	(34)	(36)	(11)	(08)	(15)	(16)	(14)	(19)	(23)	(21)
S	0.653	0.515	0.278	0.364	0.125	0.567	0.563	0.643	0.395	0.457	0.643
F	0.347	0.485	0.722	0.636	0.875	0.433	0.438	0.357	0.605	0.543	0.357
Pi-1	(59)	(34)	(36)	(11)	(08)	(15)	(16)	(14)	(19)	(23)	(21)
S	0.288	0.132	0.194	0.045	0.125	0.467	0.781	0.536	0.105	0.261	0.286
I	0.110	0.250	0.069	0.000	0.188	0.033	0.000	0.000	0.289	0.370	0.119
F	0.602	0.618	0.736	0.955	0.688	0.500	0.219	0.464	0.605	0.370	0.595
Prt-1	(59)	(32)	(36)	(11)	(08)	(15)	(16)	(14)	(19)	(23)	(21)
S	0.051	0.063	0.000	0.000	0.125	0.033	0.031	0.071	0.000	0.000	0.000
F	0.712	0.703	0.861	0.318	0.625	0.867	0.781	0.679	0.526	0.761	0.405
D	0.237	0.234	0.139	0.682	0.250	0.100	0.188	0.250	0.474	0.239	0.595
Prt-2	(58)	(31)	(36)	(11)	(07)	(15)	(13)	(12)	(19)	(23)	(19)
S	0.034	0.016	0.042	0.000	0.000	0.033	0.000	0.000	0.026	0.000	0.053
F	0.966	0.984	0.958	1.000	1.000	0.967	1.000	1.000	0.974	1.000	0.947
Pa-1	(43)	(23)	(17)	(11)	(06)	(11)	(08)	(09)	(13)	(17)	(20)
S	0.616	0.674	0.794	0.591	0.583	0.773	0.688	0.778	0.769	0.735	0.625
F	0.384	0.326	0.206	0.409	0.417	0.227	0.312	0.222	0.231	0.265	0.375

[a] Sample size for each population is in parentheses.

[b] See Table 1 for codes.

Table 3. Average estimates of the F statistic at different levels of hierarchy. F_{SC} represents the among-subdivisions variations within a cluster, and F_{CT} represents the among-cluster component of the variation

	Estimated F statistics					Subdivided F_{ST}	
Level	F_{IS}	F_{IT}	F_{ST}	Chi square[a]	df	F_{SC}	F_{CT}
Subpopulations	0.040	0.176	0.142	1,923.2***	374	0.037	0.099
Breeds	0.065	0.157	0.099	1,325.8***	153	0.041	0.069
Ancestral trunks	0.112	0.174	0.069	697.6***	68	—	—

*** $P < .001$.

0,3122

[a] Significance of F_{ST} is indicated by the chi-square value.

Table 2. Extended

Subpopulations^{b}											
GE2	SE1	SE2	CB1	CB2	CB3	PE1	PE2	PC1	PI1	PI2	PI3
(10)	(25)	(28)	(20)	(13)	(13)	(41)	(30)	(15)	(10)	(17)	(06)
1.000	0.920	0.821	0.975	0.962	0.962	0.988	1.000	1.000	1.000	0.971	0.833
0.000	0.080	0.179	0.025	0.038	0.038	0.012	0.000	0.000	0.000	0.029	0.167
(10)	(25)	(28)	(20)	(13)	(13)	(41)	(30)	(15)	(10)	(17)	(06)
1.000	1.000	1.000	1.000	1.000	1.000	0.976	0.867	0.967	1.000	1.000	1.000
0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.033	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.100	0.033	0.000	0.000	0.000
(08)	(22)	(13)	(15)	(13)	(13)	(35)	(28)	(10)	(07)	(16)	(04)
0.750	0.818	0.923	0.867	0.615	0.962	0.886	1.000	0.700	1.000	0.969	1.000
0.250	0.182	0.077	0.133	0.385	0.038	0.114	0.000	0.300	0.000	0.031	0.000
(10)	(25)	(28)	(20)	(13)	(13)	(41)	(29)	(15)	(10)	(16)	(06)
0.550	0.240	0.500	0.925	0.885	0.692	0.463	0.621	0.867	0.600	0.563	0.583
0.450	0.760	0.500	0.075	0.115	0.308	0.537	0.379	0.133	0.400	0.438	0.417
(10)	(25)	(20)	(20)	(13)	(13)	(41)	(30)	(15)	(10)	(17)	(06)
0.950	0.840	0.975	0.875	0.923	0.808	0.927	1.000	0.933	0.950	0.882	0.917
0.050	0.160	0.025	0.125	0.077	0.192	0.073	0.000	0.067	0.050	0.118	0.083
(10)	(25)	(27)	(20)	(13)	(13)	(41)	(29)	(15)	(10)	(17)	(06)
0.000	0.000	0.019	0.000	0.000	0.000	0.049	0.000	0.000	0.000	0.000	0.000
0.600	0.800	0.389	0.225	0.385	0.231	0.317	0.172	0.067	0.450	0.412	0.417
0.400	0.200	0.593	0.775	0.615	0.769	0.634	0.828	0.933	0.450	0.588	0.583
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.000
(10)	(25)	(28)	(20)	(13)	(13)	(41)	(30)	(15)	(10)	(17)	(06)
0.700	0.320	0.554	0.400	0.346	0.269	0.707	0.917	0.633	0.550	0.588	0.500
0.300	0.680	0.446	0.600	0.654	0.731	0.293	0.083	0.367	0.450	0.412	0.500
(10)	(25)	(28)	(20)	(13)	(13)	(41)	(30)	(15)	(10)	(17)	(06)
0.300	0.180	0.286	0.025	0.231	0.077	0.427	0.800	0.133	0.450	0.206	0.167
0.050	0.020	0.036	0.025	0.000	0.077	0.085	0.000	0.300	0.050	0.147	0.250
0.650	0.800	0.679	0.950	0.769	0.846	0.488	0.200	0.567	0.500	0.647	0.583
(10)	(22)	(21)	(20)	(13)	(13)	(39)	(30)	(15)	(10)	(16)	(06)
0.000	0.114	0.000	0.025	0.000	0.154	0.038	0.000	0.100	0.200	0.000	0.000
0.600	0.841	0.643	0.400	0.385	0.423	0.731	0.500	0.667	0.500	0.750	0.667
0.400	0.045	0.357	0.575	0.615	0.423	0.231	0.500	0.233	0.300	0.250	0.333
(10)	(21)	(19)	(20)	(12)	(13)	(37)	(30)	(15)	(10)	(16)	(06)
0.050	0.048	0.000	0.000	0.000	0.000	0.000	0.017	0.000	0.000	0.063	0.000
0.950	0.952	1.000	1.000	1.000	1.000	1.000	0.983	1.000	1.000	0.938	1.000
(10)	(17)	(17)	(19)	(07)	(11)	(22)	(18)	(15)	(08)	(10)	(05)
0.650	0.706	0.676	0.184	0.000	0.136	0.614	0.750	0.800	0.500	0.800	0.700
0.350	0.294	0.324	0.816	1.000	0.864	0.386	0.250	0.200	0.500	0.200	0.300

Genetic Relationships

Through cluster analysis methods and by applying the UPGMA algorithm (Sneath and Sokal 1973) to the matrix of distances-F_{ST} estimates (Table 6), the dendrogram of Figure 2 was obtained. The formation of two large clusters is seen—one formed by Perdiguero de Burgos (PB) and Podenco Canario (PC) and the other the rest of the breeds. The Ca de Bestiar (CB) breed is an exception because it seems to diverge very early from the hypothetical common trunk. Within the first group, the Mastín Español (ME) and Podenco Ibicenco (PE) would be more closely related, differentiating themselves perfectly from the other members of the cluster and forming, in turn, a new one. The last cluster—two new groups—would be differentiated; on one hand Mastín de los Pirineos (MP) and Sabueso Español (SE), and on the other hand Gos d'Atura (GA), Podenco Ibérico (PI), and Galgo Español (GE), the first two breeds being more closely related to each other than to the Galgo Español (GE).

Table 4. Genetic structure of the Spanish dog breeds from an analysis of their subpopulations

Breed^{a}	F_{IS}	F_{IT}	F_{ST}	Chi square^{b}	df
GA	0.082	0.094	0.013	38.3***	15
MP	−0.011	0.098	0.108	128.2***	26
ME	−0.011	0.015	0.026	30.7	26
PB	0.077	0.095	0.020	16.0	12
GE	0.009	0.024	0.015	11.9	12
SE	0.102	0.158	0.062	77.9***	13
CB	0.060	0.103	0.045	42.7**	22
PE	−0.055	0.002	0.054	100.9***	15
PI	0.053	0.082	0.030	27.0	26

** $P < .01$; *** $P < .001$.

^{a} The Podenco Carnario breed was not included in the analysis because a unique population was available.

^{b} Significance of F_{ST} is indicated by the chi-square value.

Discussion

Differentiation among Populations

The degree of genic differentiation (F_{ST}) among breeds of dogs has been estimated in this study as 0.099. This value is very close to 0.088, found among the three major races of man—Negroid, Mongoloid, and Caucasoid—and among local breeds of some other organisms (Nei and Roychoudhury 1982). According to these figures, the proportion of genic variation attributed to breed differences would only be about 10% of the total genetic variation, the remaining 90% corresponding to differences among individuals. The breed differences at the proteic loci level are essentially more the consequence of changes of the genic frequencies than of fixations of different alleles in the genes analyzed.

In a study about the genetic variation at the blood group loci, Lewontin (1972) concluded that the proportion of genic variation due to differences among human races was small. This author stated that there was no biological basis for the classification of human races. Our results support the disagreement expressed by Nei and Roychoudhury (1982) and Smouse et al. (1982) with that conclusion. While the interracial genic variation is small, it is important to realize that the genetic differentiation is real and generally statistically significant.

The degree of genic differentiation (F_{ST}) among subpopulations of different breeds is 14.2%. However, the average of differentiation among subpopulations within the same breed is 3.7% ($F_{SC} = 0.037$). This value is similar to the values reported for other organisms: black-tailed prairie dogs from different wards within a population (Chesser 1983), house mice from different barns or farms (Selander and Kaufman 1975), and Indians from different villages (Nei 1975). At a superior level, the average of differentiation among breeds within the same ancestral trunk is 4.1% ($F_{SC} = 0.041$).

Assuming that in the hypothetical ancestral population of canids the genetic differentiation among individuals was not important and that from the total current differentiation among individuals it has been possible to assign 14.2% of the dif-

Table 5. F statistics analysis of the Spanish dog breeds compared in pairs

Breeds	F_{IS}	F_{IT}	F_{ST}	Chi square[a]	df	P.T.[b]
GA-MP	0.075	0.101	0.028	89.03***	15	0.000+++
GA-ME	0.060	0.090	0.032	116.17***	16	0.000+++
GA-PB	0.092	0.133	0.045	186.68***	16	0.000+++
GA-GE	0.059	0.077	0.020	62.54***	15	0.003++
GA-SE	0.123	0.138	0.017	62.28***	15	0.000+++
GA-CB	0.095	0.172	0.086	225.38***	15	0.000+++
GA-PE	0.048	0.076	0.029	133.64***	16	0.000+++
GA-PC	0.065	0.112	0.050	187.95***	16	0.000+++
GA-PI	0.093	0.095	0.003	15.29NS	15	0.941NS
MP-ME	0.040	0.090	0.052	124.07***	14	0.000+++
MP-PB	0.070	0.137	0.073	173.60***	14	0.000+++
MP-GE	0.039	0.073	0.035	66.96***	13	0.000+++
MP-SE	0.104	0.115	0.012	25.72*	14	0.041+
MP-CB	0.072	0.162	0.096	145.37***	13	0.000+++
MP-PE	0.026	0.111	0.087	242.20***	15	0.000+++
MP-PC	0.042	0.126	0.088	109.32***	14	0.000+++
MP-PI	0.073	0.101	0.030	57.35***	14	0.000+++
ME-PB	0.054	0.107	0.056	122.77***	14	0.000+++
ME-GE	0.026	0.064	0.039	84.20***	14	0.000+++
ME-SE	0.088	0.130	0.046	109.15***	14	0.000+++
ME-CB	0.056	0.177	0.128	235.98***	13	0.000+++
ME-PE	0.013	0.047	0.034	111.85***	16	0.000+++
ME-PC	0.027	0.082	0.057	93.23***	14	0.000+++
ME-PI	0.058	0.089	0.034	63.84***	14	0.000+++
PB-GE	0.053	0.118	0.069	115.78***	13	0.000+++
PB-SE	0.123	0.200	0.088	193.78***	14	0.000+++
PB-CB	0.091	0.179	0.096	135.96***	14	0.000+++
PB-PE	0.041	0.106	0.068	156.11***	16	0.000+++
PB-PC	0.058	0.072	0.014	16.01NS	14	0.384NS
PB-PI	0.090	0.129	0.043	100.69***	15	0.000+++
GE-SE	0.086	0.111	0.028	57.57***	14	0.000+++
GE-CB	0.055	0.143	0.093	122.31***	13	0.000+++
GE-PE	0.013	0.048	0.036	87.54***	15	0.000+++
GE-PC	0.027	0.094	0.069	65.90***	14	0.000+++
GE-PI	0.056	0.079	0.024	36.52***	14	0.004++
SE-CB	0.127	0.227	0.115	193.96***	13	0.000+++
SE-PE	0.079	0.136	0.062	194.34***	15	0.000+++
SE-PC	0.096	0.184	0.097	156.93***	14	0.000+++
SE-PI	0.122	0.138	0.018	39.60***	14	0.017+
CB-PE	0.042	0.173	0.137	286.04***	15	0.000+++
CB-PC	0.061	0.173	0.120	113.87***	12	0.000+++
CB-PI	0.093	0.177	0.093	123.96***	13	0.000+++
PE-PC	0.012	0.078	0.068	122.24***	15	0.000+++
PE-PI	0.045	0.072	0.027	82.09***	16	0.000+++
PC-PI	0.062	0.112	0.054	66.17***	14	0.000+++

[a] Significance of F_{ST} indicated by the chi-square value: * $P < .05$; *** $P < .001$; NS = not significant.

[b] Significance of F_{ST} by the permutational test (P.T.). The values correspond to the area of the right tail: + $P < .05$; ++ $P < .01$; +++ $P < .001$; NS = not significant.

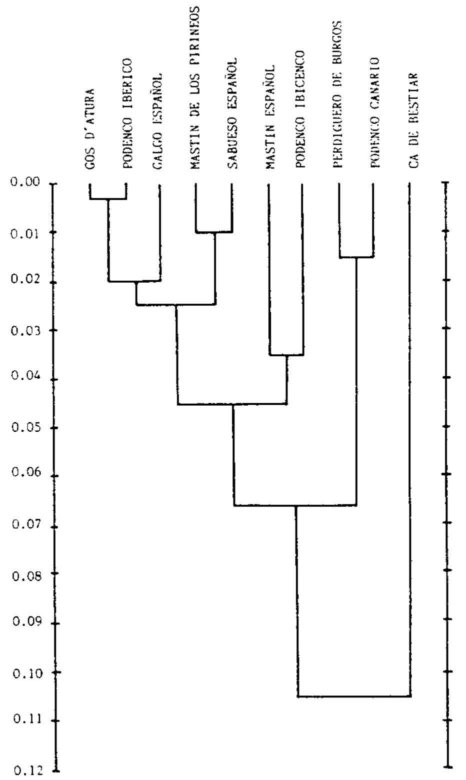

Figure 2. Phenogram obtained by the UPGMA method using F_{ST} distance values.

ferentiation to the lowest hierarchical level (subpopulations), it is possible to divide the percentage of the observed genic differentiation through evolutionary time in the following way. The 48.6% of the variation could be explained by the differentiation of ancestral trunks. The differentiation that took place during the breed formation process would explain 30.3%, while the remaining 21.1% would be explained by the process of subdivision of breeds in subpopulations.

The genetic structure of the breeds is summarized in Table 4. There are highly significant differences among subpopulations that integrate the Mastín de los Pirineos, Sabueso Español, and Podenco Ibicenco breeds in accordance with what was observed by the study of genetic distances (Jordana 1989). But besides that, significant differences can be seen in two other breeds, Gos d'Atura and Ca de Bestiar. The remaining breeds do not show differentiation among their subpopulations.

An important factor to explain the intraracial differences—that is, differences among subpopulations of the same breed—would be the probable gene introduction or migration from some other very close breed in a particular subpopulation of the breed but not in other subpopulations, having as its goal the improvement of some morphological or functional characteristics of that subpopulation. The possible flow of genes from the Mountain of the Pyrenees (French dog breed) in the subpopulations of the Mastín de los Pirineos could be an example (Jordana et al. 1991a). Assuming that the loci analyzed are selectively neutral, the percentage of genes of each of the migrated loci from one population to the other will be given by chance, depending on the values of the genic frequencies that each locus has in the given population.

There are highly significant differences among the subpopulations of Podenco Ibicenco ($F_{ST} = 0.054$, $P < .001$). A migration of genes from the PE1.Catalunya subpopulation to PE2.Baleares does not seem to

Table 6. F_{ST} estimates used as a measure of genetic distance between Spanish dog breeds

Breeds	GA	MP	ME	PB	GE	SE	CB	PE	PC
MP	0.028								
ME	0.032	0.052							
PB	0.045	0.073	0.056						
GE	0.020	0.035	0.039	0.069					
SE	0.017	0.012	0.046	0.088	0.028				
CB	0.086	0.096	0.128	0.096	0.093	0.115			
PE	0.029	0.087	0.034	0.068	0.036	0.062	0.137		
PC	0.050	0.088	0.057	0.014	0.069	0.097	0.120	0.068	
PI	0.003	0.030	0.034	0.043	0.024	0.018	0.093	0.027	0.054

exist, but there is a migration in the other direction (Jordana 1989). This hypothesis is supported by the fact that hare and rabbit hunters from Catalonian areas have always been potential buyers of Majorcan "podencos." "Podencos" pure breeders would have remained in the islands, whose "blood lines" would not have received the influence of genes from the Peninsula.

The differences among the Gos d'Atura subpopulations, GA1.Típico and GA2.Cerdà, are probably due mainly to a reproductive isolation motivated by economic reasons (Jordana et al. 1991b). Dog breeders are reluctant to intercross long-haired dogs with short-haired dogs. Long-haired Gos d'Atura dogs live mainly in big cities and usually are registered in the official herdbook; they have been integrated in the subpopulation GA1.Típico. In contrast, short-haired dogs, included in the subpopulation GA2.Cerdà, live in small villages or in farms dispersed across the country and are usually nonregistered. Nevertheless, both subpopulations diverged from a common population where the recessive allele (*l*), which seems to govern long hair, segregated with the dominant allele (*L*) that governs the short-haired phenotype.

There are differences among subpopulations of Ca de Bestiar, although there is no significant deficit of heterozygotes in the population as a whole (F_{IT}). These differences might be explained as an accumulative effect of small differences in the genic frequencies of subpopulations, caused by a random process of genetic drift.

Genetic Relationships among Spanish Dog Breeds

The dendrogram in Figure 2 shows the genetic relationships among Spanish dog breeds. This dendrogram is topologically identical to the one obtained using Nei's index of distance (Nei 1978) from values of allelic frequencies of genes that codify for soluble proteins and enzymes in the blood (Jordana 1989). The correlation coefficient between both distances (Nei's genetic distance versus F_{ST} values) was $r = .99$, and highly significant. Nevertheless, the grouping obtained from genetic distances differs from that derived from historical, archaeological, and morphological data (Boivin et al. 1983, Jordana 1989; Villemont 1970).

The breeds showing less congruence with the archaeological and morphological classification are Ca de Bestiar, Podenco Canario, and Perdiguero de Burgos. Data from levels of genetic variability performed in these three breeds (Jordana et al. 1991a), as well as results of the heterozygosities and the number of detected alleles per locus, show that these three breeds have suffered important bottlenecks at some time in their history. According to Nei and Roychoudhury (1982) and Nei (1987), when a population is submitted to a bottleneck effect the genetic distances increase quickly. This increase of genetic distances might subsequently distort the topology of the evolutive trees. In the Ca de Bestiar breed the magnitude of the bottleneck effect was so high that it could be considered a founder effect. According to Guasp (1982) this breed would have nearly disappeared in the 1960s; its recovery began in the 1970s starting from only four males and two females. Similar discrepancies in interpretation of evolutive trees in other species, due to bottleneck effects, have been described by Nei and Roychoudhury (1982) in the human races, by Chesser (1983) in *Cynomius ludovicianus*, or prairie dogs, and by Gyllensten et al. (1983) in the European red deer.

The large divergence in the three mentioned breeds versus all the other breeds is due primarily to the close fixation of the rare allele $Pa\text{-}1^F$ (Table 2) in the Ca de Bestiar population, as well as to the low frequencies in the alleles Alb^F and $Pi\text{-}1^S$, which contribute to the low genetic identity of Ca de Bestiar with the other breeds. In the populations of Perdiguero de Burgos and Podenco Canario the high frequencies of the alleles Alb^S and Tf^C contribute mainly to the low genetic identity with the rest of the breeds.

As we have demonstrated in this article, the F statistics can be used both for the analysis of the genetic structure of the populations and for the reconstruction of phylogenies. It should be emphasized, however, that while the dendrograms show the genetic relationships between breeds, they cannot show the actual evolutionary history of populations if these are not completely isolated (Nei 1987).

References

Andreu X, 1984. El Gos d'Atura: Perro de Pastor Catalán. Barcelona: De Vecchi.

Anonymous, 1982. Esquema del estudio sobre el perro Podenco Canario. In: I Symposium Nacional de las Razas Caninas Españolas, Córdoba, March 19–21. Córdoba, Spain: Universidad de Córdoba; 257–262.

Boivin R et al., 1983. Nuestro amigo el perro. Madrid: Selecciones del Reader's Digest (Iberia).

Chesser RK, 1983. Genetic variability within and among populations of the black-tailed prairie dog. Evolution 37:320–331.

Delalix AM, 1986. Los Perros españoles. Barcelona: De Vecchi.

Esquiró L, 1982. El Mastín Español. In: I Symposium Nacional de las Razas Caninas Españolas, Córdoba, March 19–21. Córdoba, Spain: Universidad de Córdoba; 123–141.

García JV, Flores AJ, and Aranda RJ, 1982. Contribución al estudio de algunos caracteres faneróptícos y zoométricos del perro Podenco ibérico andaluz malagueño. In: I Symposium Nacional de las Razas Caninas Españolas, Córdoba, March 19–21. Córdoba, Spain: Universidad de Córdoba; 273–278.

Gómez-Toldrà S, 1985. Los perros. Gran Enciclopedia Canina. Barcelona: Planeta-De Agostini.

Gondrexon A and Browne I, 1982. Guía de los perros del mundo. Barcelona: Omega.

Gregorius HR and Roberds JH, 1986. Measurement of genetical differentiation among subpopulations. Theor Appl Genet 71:826–834.

Guasp A, 1982. Pastor Mallorquín o Ca de Bestiar. En: I Symposium Nacional de las Razas Caninas Españolas, Córdoba, March 19–21. Córdoba, Spain: Universidad de Córdoba; 45–51.

Gyllensten U, Ryman N, Reuterwall C, and Dratch P, 1983. Genetic differentiation in four European subspecies of red deer (*Cervus elaphus* L.). Heredity 51:561–580.

Hedrick PW, 1985. Genetics of populations. Boston: Jones and Bartlett.

Jordana J, 1989. Relaciones genéticas en cánidos españoles (PhD dissertation). Barcelona: Universidad Autónoma de Barcelona.

Jordana J, Piedrafita J, and Sánchez A, 1991a. Variabilidad genética en diez razas caninas españolas. Arch Zootec 147:115–129.

Jordana J, Piedrafita J, and Sánchez A, 1991b. Variabilidad y relaciones genéticas de cinco poblaciones de la raza canina "Gos d'Atura." Invest Agrar Prod y Sanid Anim 6(3):211–223.

Jordana J, Sánchez A, and Piedrafita J, 1990. Los cánidos españoles; Orígenes y clasificación. ONE 89:2–8.

Levene H, 1949. On a matching problem arising in genetics. Ann Math Stat 20:91–94.

Lewis PAW and Orav EJ, 1989. Simulation methodology for statisticians, operations analysts, and engineers, vol 1. Pacific Grove, California: Wadsworth and Brooks/Cole Advanced Books and Software.

Lewontin RC, 1972. The apportionment of human diversity. Evol Biol 6:381–398.

Long JC, 1986. The allelic correlation structure of Gainj and Kalam speaking people: I. The estimation and interpretation of Wright's F-statistics. Genetics 112:629–647.

Long JC, Smouse PE, and Wood JW, 1987. The allelic correlation structure of Gainj and Kalam speaking people: II. The genetic distance between population subdivisions. Genetics 117:273–283.

Maza JL, 1982. Problemática actual del Podenco Ibicenco. In: I Symposium Nacional de las Razas Caninas Españolas, Córdoba, March 19–21. Córdoba, Spain: Universidad de Córdoba; 227–235.

Mora J, 1982. Aproximación al Podenco Ibicenco. In: I Symposium Nacional de las Razas Caninas Españolas, Córdoba, March 19–21. Córdoba, Spain: Universidad de Córdoba; 243–256.

Nei M, 1975. Molecular population genetics and evolution. Amsterdam: North-Holland.

Nei M, 1977. F-statistics and analysis of gene diversity in subdivided populations. Ann Hum Genet 41:225–233.

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. New York: Columbia University Press.

Nei M and Roychoudhury AK, 1982. Genetic relationship and evolution of human races. Evol Biol 14:1-59.

Pugnetti G, 1981. Guía de perros. Barcelona: Grijalbo.

Sanz Timón JM, 1982. El perro perdiguero de Burgalés. In: I Symposium Nacional de las Razas Caninas Españolas, Córdoba, March 19-21. Córdoba, Spain: Universidad de Córdoba; 359-379.

Selander RB and Kaufman DW, 1975. Genetic structure of populations of the brown snail (*Helix apersa*): I. Microgeographic variation. Evolution 29:385-401.

Smouse PE and Long JC, 1988. A comparative F-statistics analysis of the genetic structure of human populations from Lowland South America and Highland New Guinea. In: Proceedings of the 2nd International Conference on Quantitative Genetics (Weir BS, Eisen EJ, Goodman MM, and Namkoong G, eds). Sunderland, Massachusetts: Sinauer Associates; 32-46.

Smouse PE, Spielman RS, and Park MH, 1982. Multiple-locus allocation of individuals to groups as a function of the genetic variation within and differences among human populations. Am Nat 119:445-463.

Sneath PHA and Sokal RR, 1973. Numerical taxonomy. San Francisco: W.H. Freeman.

Sotillo JL and Serrano V, 1985. Producción animal: I. Etnología zootécnica. Madrid: Tebas-Flores.

Swofford DL and 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.

Swofford DL and Selander RB, 1989. BIOSYS-1 user's manual, release 1.7: a computer program for the analysis of allelic variation in population genetics and biochemical systematics. Champaign: Illinois Natural History Survey.

Tanabe Y, Sugiura S, Asanoma M, and Ota K, 1974. Genetic polymorphism of leucine aminopeptidase in canine plasma. Anim Blood Grps Biochem Genet 5: 225-230.

Villemont M, 1970. Le grand livre du chien. Geneva: Edito-Service.

Weir BS, 1990. Genetic data analysis. Sunderland, Massachusetts: Sinauer Associates.

Wright S, 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395-420.

Wright S, 1978. Evolution and the genetics of populations: Vol. 4. Variability within and among natural populations. Chicago: University of Chicago Press.