

Bayesian analysis of genetic differentiation between populations of the Duroc and Iberian pig breeds

Alves E., Fernández A., Barragán C., Rodríguez C., Silió L.

in

Audiot A. (ed.), Casabianca F. (ed.), Monin G. (ed.).
5. International Symposium on the Mediterranean Pig

Zaragoza : CIHEAM

Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 76

2007

pages 23-28

Article available on line / Article disponible en ligne à l'adresse :

<http://om.ciheam.org/article.php?IDPDF=800553>

To cite this article / Pour citer cet article

Alves E., Fernández A., Barragán C., Rodríguez C., Silió L. **Bayesian analysis of genetic differentiation between populations of the Duroc and Iberian pig breeds.** In : Audiot A. (ed.), Casabianca F. (ed.), Monin G. (ed.). 5. *International Symposium on the Mediterranean Pig* . Zaragoza : CIHEAM, 2007. p. 23-28 (Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 76)



<http://www.ciheam.org/>
<http://om.ciheam.org/>

Bayesian analysis of genetic differentiation between populations of the Duroc and Iberian pig breeds

E. Alves, A. Fernández, C. Barragán, C. Rodríguez and L. Silió

Departamento de Mejora Genética Animal, INIA, Carretera de la Coruña Km 7, 28040 Madrid, Spain

SUMMARY – Bayesian methods have been proposed for estimating hidden population substructures in closely related populations. A sampling of 64 Duroc and 170 Iberian pigs, assigned to two strains (Torbiscal and Guadyerbas), two red varieties (Retinto, Entrepelado) and one black hairless variety (Lampiño) was genotyped for 36 microsatellites. An optimal partition of only five clusters was estimated, when the previous assignment to six pig groups was taken into account. The analysis without pre-assigned groups showed a more complex partition: (i) most of the Duroc pigs were grouped into the same cluster; (ii) the cluster of greater size included entirely the Entrepelado variety and most of the Retinto and Lampiño pigs; and (iii) the remaining seven clusters grouped pigs from five small isolated herds or from Torbiscal and Guadyerbas closed strains. Genotypes of MC1R coat color gene also revealed a large intercross between black and red Iberian varieties. The future definition of conservation units in the Iberian breed should consider these results.

Keywords: Iberian, Duroc, varieties, genetic differentiation, clustering.

RESUME – "Analyse bayésienne de la différenciation génétique entre populations des races porcines Duroc et Ibérique". Des méthodes bayésiennes ont été proposées pour évaluer la sous-structure cachée de populations très proches. Un échantillon de 64 porcs Duroc et 170 porcs Ibériques, appartenant à deux lignées (Torbiscal et Guadyerbas), deux souches rouges (Retinto et Entrepelado) et une noire sans poil (Lampiño) a été génotypé pour 36 microsatellites. Une partition optimale en seulement cinq clusters a été estimée, quand l'assignation préalable en six groupes de porcs a été prise en compte. L'analyse sans assignation préalable a montré une partition optimale plus complexe : (i) la plupart des porcs Duroc étaient groupés dans le même cluster ; (ii) le groupe le plus grand contenait tous les porcs de souche Entrepelado et la plupart des animaux Retinto et Lampiño ; et (iii) sept autres clusters englobaient des porcs de cinq petits troupeaux isolés ou des lignées fermées Torbiscal et Guadyerbas. Les génotypes du gène de couleur du pelage MC1R montrent aussi un important croisement entre les souches Ibériques noires et rouges. La future définition des unités de conservation de la race Ibérique devrait tenir compte de ces résultats.

Mots-clés : Ibérique, Duroc, lignées, différenciation génétique.

Introduction

The morphology of the Iberian pig makes it resistant to sunstroke and high summer temperatures, and enables it for pasturing acorns, grass, small roots and bulbs, available during autumn and winter in the sparse woodlands of the SW of the Iberian Peninsula. Although all the Iberian pigs share the aforementioned features, several local varieties were differentiated showing phenotypic differences (black hairless, red and blond pigs), and with only rare genetic interchange among them. Some of these varieties were exported to some American countries during the colonization, being the direct ancestors of Creole pig breeds and also contributed to the origin of the Duroc-Jersey breed in the United States.

The Iberian breed constitutes the largest of the surviving Mediterranean type pig populations. Its described heterogeneity was acknowledged in the breed standard type, which was lately proposed during the past century. The breed census was drastically reduced since 1960 due to the outbreak of the African swine fever, the depreciation of animal fats and the introduction of more efficient foreign breeds. The old breed structure, with differentiated varieties locally distributed, was replaced by a pyramidal structure based on crossbreeding with Duroc, and a strong dependence on a small number of breeding nuclei supplying purebred Iberian animals to all production tiers. Along the last years, this population bottleneck has been reversed, and the production of Iberian type pigs has largely increased to satisfy the new demand of top quality meat and dry-cured products. However, as a consequence of the past critical period, some ancestral varieties have disappeared and other ones could be endangered or blended.

Phylogenetic techniques based on genetic distances estimated from polymorphic microsatellite markers have been the method of choice to assess the genetic diversity of livestock breeds. Martínez *et al.*, (2000) analyzed the genetic structure of Iberian breed using this classical approach. Other methods that construct genetic clusters from a set of individual multilocus have been proposed as a more flexible alternative to those based on genetic distances. These clustering methods allow to separate a set of individuals in several populations if their genetic origin is unknown beforehand or to study the correspondence between inferred genetic clusters and known predefined population categorizations (Pritchard *et al.*, 2000). Both genetic distances and clustering methods have been used by Fabuel *et al.*, (2004) to analyze genetic diversity and conservation priorities in Iberian pigs. The results of these previous studies and other ones based on the analysis of mitochondrial DNA sequences (Alves *et al.*, 2003) supported evidence of intercross between some of the traditional varieties.

Recently, Bayesian methods have been proposed for estimating hidden population substructure, which treat both the allele frequencies of the molecular markers and the number of genetically diverged populations as random variables (Corander *et al.*, 2003). The objectives of this study were to apply this new statistical tool: (i) for assessing the genetic diversity of the varieties and preserved strains of Iberian pigs; and (ii) for defining more accurately the possible conservation units.

Material and methods

Animals

Two out of the groups of Iberian pigs considered here, Guadyerbas and Torbiscal, belong to an early conservation programme (Rodrigáñez *et al.*, 2000; Toro *et al.*, 2000). The complete genealogy of all the animals is available since 1945, with 18.9 (Guadyerbas) and 21.0 (Torbiscal) generations from the population founders until the animals genotyped here. The remaining Iberian pig groups represent the main three extant varieties: black hairless (Negro Lampiño) and red (Retinto) pigs and a black hairy variety (Entrepelado), whose piglets show a chestnut colour at birth. Blood samples were collected from 170 individuals inscribed in the breed herdbook, their distribution being as follows: 31 Torbiscal, 32 Guadyerbas, 50 Retinto (seven herds), 30 Lampiño (three herds) and 27 Entrepelado (five herds). Due to their historical and current relations with the Iberian pigs, a further 64 Duroc pigs from seven herds were also sampled and analyzed.

Microsatellites

All the animals were genotyped for 36 microsatellite markers, two on each autosome (Table 1). They were chosen for their reproducibility, position on the chromosome, polymorphism and absence of null alleles. Amplified microsatellite markers were analyzed with Genescan software on capillary electrophoresis equipment with fluorescent detection (ABI PRISM 310 genetic analyzer). To increase the accuracy of allele size determination, four control animals were genotyped in all the gels.

Clustering analysis

Microsatellite data were analyzed for estimating jointly the posterior distributions of the parameters related to the underlying population substructure: (i) the number of populations (n_p) with differing allele frequencies; (ii) the allele frequencies, jointly represented by the symbol q_j for population j ; and (iii) the population structure matrix S , with elements defined as $S_{mr} = 1$, if $q_m = q_r$ and $S_{mr} = 0$, if $q_m \neq q_r$. Both Hardy-Weinberg and linkage equilibrium were assumed. Details of the method can be found in Corander *et al.* (2003), and the calculations were performed using the BAPS2.0 software. Group level and individual level clustering were performed in this study. In the first one, the previous assignation of the pigs to six groups (varieties or strains) was used, and these groups are the sampling units to be clustered to study the correspondence between inferred genetic clusters and known predefined population categorization. A second analysis fitting an individual model was also carried out, without preassigned groups, to identify the optimal allocation of individuals to genetically divergent clusters. After the structure was inferred, we calculated the conditional estimates for the model averaged coefficient of genetic differentiation F_{ST} (Nei, 1977) and the pairwise genetic distances D_m (Nei, 1972) and D_R (Reynolds *et al.*, 1983).

Table 1. Number of alleles (*n*) and observed (*Ho*) and expected (*He*) heterozygosities of microsatellite markers in Iberian and Duroc breeds

Chr.	Markers	Iberian			Duroc			Markers	Iberian			Duroc		
		n	Ho	He	n	Ho	He		n	Ho	He	n	Ho	He
1	S0113	5	0.671	0.728	3	0.375	0.412	S0155	5	0.447	0.483	4	0.672	0.593
2	SW240	8	0.594	0.724	7	0.797	0.774	SW395	6	0.565	0.713	3	0.266	0.300
3	SW72	6	0.512	0.687	6	0.656	0.708	S0002	8	0.606	0.677	5	0.672	0.752
4	S0301	5	0.365	0.524	5	0.641	0.730	S0097	12	0.718	0.826	7	0.531	0.576
5	SW413	4	0.576	0.720	5	0.484	0.765	S0005	14	0.647	0.783	15	0.625	0.803
6	SW1057	8	0.653	0.722	5	0.453	0.723	SW2419	8	0.706	0.758	5	0.609	0.674
7	S0025	8	0.594	0.672	5	0.425	0.608	SW632	10	0.541	0.678	7	0.469	0.520
8	S0178	9	0.635	0.758	7	0.656	0.712	S0225	7	0.571	0.661	4	0.594	0.721
9	SW911	7	0.547	0.699	6	0.656	0.678	SW1349	5	0.482	0.553	5	0.171	0.626
10	S0038	8	0.465	0.765	6	0.484	0.447	S0070	11	0.712	0.819	10	0.609	0.625
11	S0385	5	0.529	0.665	6	0.734	0.770	SW703	5	0.418	0.566	4	0.453	0.579
12	SW874	7	0.582	0.774	9	0.578	0.783	S0106	13	0.694	0.816	6	0.500	0.578
13	S0219	4	0.476	0.650	4	0.359	0.472	S0291	7	0.523	0.719	6	0.719	0.751
14	SW210	5	0.535	0.676	4	0.453	0.555	SW857	5	0.624	0.711	5	0.453	0.668
15	SW1111	6	0.653	0.785	6	0.651	0.656	SW936	9	0.723	0.810	5	0.641	0.723
16	S0026	7	0.341	0.460	5	0.562	0.667	S0061	5	0.659	0.657	5	0.672	0.713
17	SW1920	8	0.612	0.759	7	0.453	0.760	SW24	6	0.582	0.692	7	0.719	0.763
18	S0120	7	0.582	0.723	4	0.406	0.544	SW787	6	0.612	0.680	7	0.797	0.812
								Average	7.2	0.576	0.697	5.8	0.556	0.653

MC1R genotyping

The *MC1R* intragenic haplotypes present in Iberian and Duroc populations were determined by the procedures described by Fernández *et al.* (2004).

Results and discussion

Table 1 shows for each autosome, the number of alleles of each genotyped microsatellite, that ranges from 4 to 14 (average 7.2) summed across the Iberian breed as a whole, and from 3 to 15 in the Duroc population (average 5.4). Table 1 also gives, for each locus, the observed heterozygosity by direct count and the expected heterozygosity under Hardy-Weinberg equilibrium. The observed values were generally lower than the expected values, indicating heterogeneity between populations within each breed.

Group level clustering

The results show a best partition of only five clusters with differing population frequencies. Four of them correspond to the Duroc breed (C1), the Lampiño variety (C4), and the Guadyerbás (C3) and Torbiscal (C5) strains. The last cluster (C2) combines the Entrepelado and Retinto varieties of Iberian pigs. The posterior mean value of the coefficient of genetic differentiation among these five clusters F_{ST} , was 0.170 (psd = 0.003). The means of the posterior distributions of D_m and D_R distances indicate a maximum distance between Duroc and Guadyerbás, and a minimum distance between the Lampiño variety and the mixed cluster Retinto/Entrepelado (Table 2). The values of genetic distances are strongly dependent on the number of generations since divergence and the effective sizes of populations. The D_m distance also depends on the founder frequencies (Laval *et al.*, 2002). It could explain the high values of this distance between the Duroc and Iberian clusters.

Table 2. Group level clustering. Posterior means and standard deviations of genetic distances among the five inferred clusters. (Nei's distance D_m , above the diagonal and Reynolds's distance D_R , below the diagonal)

Inferred Cluster	C1	C2	C3	C4	C5
C1 Duroc		0.561 (0.018)	0.905 (0.026)	0.524 (0.020)	0.769 (0.025)
C2 Entrepelado/Retinto	0.183 (0.005)		0.482 (0.019)	0.194 (0.013)	0.301 (0.015)
C3 Guadyerbás	0.337 (0.006)	0.243 (0.007)		0.412 (0.022)	0.416 (0.022)
C4 Lampiño	0.175 (0.005)	0.080 (0.005)	0.222 (0.008)		0.324 (0.020)
C5 Torbiscal	0.253 (0.006)	0.137 (0.006)	0.248 (0.009)	0.144 (0.007)	

Individual level clustering

The results of the clustering without preassigned groups show a best partition of nine genetically divergent clusters. The composition of these clusters is described in the Table 3. Four results can be outlined: (i) two of the inferred clusters (IC1 and IC2) separately include animals from the Torbiscal and Guadyerbás closed strains, corresponding to clusters obtained in the previous analysis; (ii) the cluster of greatest size (IC3) includes entirely the Entrepelado variety (27 animals) and most of the Retinto and Lampiño pigs, showing that the traditional Iberian varieties have been largely blended; (iii) most of the Duroc pigs (47 animals) are grouped in the same cluster; and (iv) the remaining five clusters correspond to individuals sampled from small isolated herds: IC5 and IC8 clusters join Duroc pigs imported from United States forty years ago, IC6 and IC10 clusters separately join animals of two ancient herds of the Retinto variety, and IC7 cluster join pigs of other ancient herd of the Lampiño variety.

The posterior mean of the coefficient of genetic differentiation among these nine clusters F_{ST} , was 0.286 (psd = 0.005). According to this more complex partition, the F_{ST} value largely exceeds the one obtained in the previous clustering. To simplify the presentation of results of pairwise genetic distances, only the values between clusters grouping 30 or more individuals (IC1 to IC4) are shown in Table 3.

Table 3. Individual level clustering. Posterior means and standard deviations of genetic distances among the four inferred clusters grouping at least 30 pigs. (Nei's distance D_m , above the diagonal and Reynolds's distance D_R , below the diagonal)

Inferred Clusters [†]	Number and group of individuals clustered	IC1	IC2	IC3	IC4
IC1	32 Guadyerbas		0.416 (0.022)	0.455 (0.018)	0.935 (0.029)
IC2	31 Torbiscal	0.249 (0.022)		0.277 (0.014)	0.824 (0.028)
IC3	33 Retinto, 18 Lampiño 27 Entrepelado	0.232 (0.006)	0.127 (0.006)		0.571 (0.019)
IC4	47 Duroc	0.345 (0.006)	0.266 (0.006)	0.184 (0.005)	

[†] Number and group of pigs clustered into other clusters of lower size: IC5 = 8 Duroc; IC6 = 7 Retinto; IC7 = 12 Lampiño; IC8 = 9 Duroc; IC9 = 10 Retinto.

MC1R genotypes

The complementary analysis, based on the coat color *MC1R* gene revealed some results unexpected under the hypothesis of strict isolation between the traditional varieties of Iberian pigs (Table 4): (i) the presence in some Lampiño pigs of *MC1R*6* or *MC1R*7* alleles, jointly with the *MC1R*3* allele, characteristic of black populations; and (ii) the presence of *MC1R*3* alleles in some Retinto pigs, jointly with the *MC1R*6* or *MC1R*7* alleles, characteristic of red populations. However, the joint segregation of *MC1R*3*, *MC1R*6* or *MC1R*7* alleles in Torbiscal and Entrepelado pigs could be expected. Torbiscal is a composite strain obtained from ancient black and red varieties (Rodrigáñez *et al.*, 2000). The Entrepelado variety is an emergent type of Iberian pigs that should proceed from the intercross among Iberian pigs of diverse coat color and hair density. The presence of the black *MC1R*3* allele in the red Retinto, Torbiscal and Entrepelado populations refutes the assumed dominance of this allele.

Table 4. Frequencies of *MC1R* alleles in the six Iberian and Duroc pig populations studied

	Guadyerbas (N = 32)	Torbiscal (N = 31)	Lampiño (N = 30)	Retinto (N = 50)	Entrepelado (N = 27)	Duroc (N = 64)
Black Allele <i>MC1R*3</i>	1.00	0.10	0.92	0.04	0.07	0.00
Red Alleles						
<i>MC1R*4</i>	0.00	0.00	0.00	0.00	0.00	1.00
<i>MC1R*6</i>	0.00	0.73	0.08	0.65	0.54	0.00
<i>MC1R*7</i>	0.00	0.17	0.00	0.31	0.39	0.00

References

- Alves, E., Ovilo, C., Rodríguez, M.C. and Silió, L. (2003). Mitochondrial DNA sequence variation and phylogenetic relationship among Iberian pigs and other domestic and wild pig populations. *Anim. Genet.*, 34: 319-324.
- Corander, J., Waldmann, P. and Sillanpää, M.J. (2003). Bayesian analysis of genetic differentiation between populations. *Genetics*, 163: 367-374.
- Fabuel, E., Barragán, C., Silió, L., Rodríguez, M.C. and Toro, M. (2004). Analysis of genetic diversity and conservation priorities in Iberian pigs based on microsatellite markers. *Heredity*, 93: 104-113.
- Fernández, A., Fabuel, E., Alves, E., Rodríguez, M.C., Silió, L., and Ovilo, C. (2004). DNA tests based on coat colour genes for authentication of the raw material of meat products from Iberian pigs. *J. Sci. Food Agric.*, 84: 1855-1860.
- Laval, G., San Cristóbal, M. and Chevalet, C. (2002). Measuring genetic distances between breeds: Use of some distances in various short term evolution models. *Genet. Sel. Evol.*, 34: 481-507.
- Martínez, A.M., Delgado, J.V., Rodero, A. and Vega-Pla, J.L. (2000). Genetic structure of the Iberian pig breed using microsatellites. *Anim. Genet.*, 31: 295-301.

- Nei, M. (1972). Genetic distance between populations. *Am. Naturalist*, 106: 283-292.
- Pritchard, J.K., Stephens, M. and Donnelly, T. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155: 945-959.
- Reynolds, J., Weir, B.S. and Cockerham, C.C. (1983). Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics*, 105: 767-779.
- Rodrigáñez, J., Toro, M.A., Rodríguez, M.C. and Silió, L. (2000). Alleles survival from Portuguese and Spanish strains in a population of Iberian pig. In: Tradition and innovation in Mediterranean pig production, Afonso, J.A., Tirapicos, J.L. (eds). *Options Méditerranéennes*, Series A, pp. 57-61.
- Toro, M.A., Rodrigáñez, J., Silió, L. and Rodríguez, M.C. (2000). Genealogical analysis of a closed herd of black hairless Iberian pigs. *Conserv Biol*, 14: 1843-1851.