Available online at www.sciencedirect.com
SCIENCE DIRECT
Meat Science 72 (2006) 560–566
MEAT SCIENCE
www.elsevier.com/locate/meatsci

Estimation of the genetic admixture composition of Iberian dry-cured ham samples using DNA multilocus genotypes

David García [a], Amparo Martínez [b], Susana Dunner [a], José L. Vega-Pla [c], Cristina Fernández [a], Juan V. Delgado [b], Javier Cañón [a,*]

[a] *Laboratorio de Genética, Facultad de Veterinaria, Universidad Complutense de Madrid, 28040 Madrid, Spain*
[b] *Departamento de Genética, Universidad de Córdoba, Campus de Rabanales, 14071 Córdoba, Spain*
[c] *Laboratorio de Genética Molecular, Fondo de Explotación de los Servicios de Cría Caballar y Remonta, 14071 Córdoba, Spain*

Received 21 May 2005; received in revised form 14 September 2005; accepted 14 September 2005

Abstract

Under current Spanish regulations, the pigs that provide the raw material for the preparation of the country's most appreciated meat-derived product, dry-cured Iberian ham, must be of a specific genetic composition. Only the Duroc breed is accepted for crossing with Iberian pigs, and a maximum of 50% of the Duroc genome is permitted in the animals used to make this ham. This paper describes a set of statistical procedures for detecting the 'breed composition' of Iberian ham via the use of multilocus genotypes obtained by the amplification of 25 microsatellite markers. The proposed procedure detected up to 20% of ham samples with a genetic composition incompatible with present legislation – either because the Duroc genome was present in a percentage greater than that permitted, or because of the significant presence (>25%) of white coat pig genomes. The probability of finding an illegal cured ham was greater in restaurants than in retail grocery stores, and in medium-low category restaurants or stores than in higher category establishments.

Keywords: Iberian ham traceability; Genetic admixture; Multilocus genotypes; Iberian pig

1. Introduction

Dry-cured Iberian ham, characterised by its intense flavour, is one Spain's most appreciated meat-derived products (López-Bote, 1998). Three main factors are associated with the high market price of this product: the breed origin of the raw material (Iberian pigs), the extensive finishing system based on acorns and the use of pasture (known as *montanera*), and a prolonged, traditional processing method involving a 12–24 month ripening period. These factors strongly limit supply, which is usually exceeded by demand.

The term 'Iberian pig' is a racial grouping for the native pigs of the Iberian Peninsula (Dieguez, 1992) which survive in the Mediterranean forest ecosystem traditionally known as *La Dehesa* (pastureland interspersed with Mediterranean oaks) (López-Bote, 1998). The breed is characterised by high levels of subcutaneous and intramuscular fat which, if the animal finishes its fattening period in an extensive system, is rich in monounsaturated (MUFA) and polyunsaturated (PUFA) fatty acids (Flores, Biron, Izquierdo, & Nieto, 1988; Mayoral et al., 1999). Industrial crosses between Duroc and Iberian pigs have become commonplace since this helps improve the growth rate, the food conversion rate and the lean content of the carcass. The effect of genotype (Iberian or Iberian × Duroc [50:50]) has little effect on the sensory and fatty acid profiles of the final ham (Carraspiso, Bonilla, & García, 2003) but has an important effect on triacylglycerol composition (Petrón, Muriel, Timón, Martín, & Antequera, 2004).

Protection by governmental regulation against breed substitution is important for reasons of conservation and trade (Branciari, Nijman, Plas, Di Antonio, & Lenstra, 2000).

* Corresponding author. Fax: +34 913 943 772.
E-mail address: jcanon@vet.ucm.es (J. Cañón).

0309-1740/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.meatsci.2005.09.005

Present Spanish legislation (Boletin Oficial del Estado, 2001, 2003) allows up to 50% Duroc origin in animals used to produce Iberian hams. However, the products made from these different genetic sources are clearly labelled as either *Iberian ham* or *Pure Iberian ham*: the latter requires the Iberian pig be the only breed involved. No other breed of pig may be crossed with Iberian pigs under present legislation.

The high price differences (5–10-fold) between dry-cured Iberian ham and ham derived from other pig breeds can encourage the indiscriminate use of the Iberian ham label, resulting in a negative impact on the regulated product. It would therefore be very advantageous to be able determine the existence of consumer fraud and the degree to which this occurs.

In recent years, methods have been proposed for assigning anonymous samples to (or for excluding them from) reference populations by making use of hypervariable molecular markers (Baudouin, Piry, & Cornuet, 2004; Cornuet, Piry, Luikart, Estoup, & Solignac, 1999). Genotyping has also been proposed, based on single genes involved in the coat colour pattern such as MC1R and KIT (Kijas et al., 1998; Kijas, Moller, Plastow, & Andersson, 2001). In fact, the British Wild Boar Association developed a test based on variants of the MC1R gene to differentiate between wild boars and commercial pigs. Recently, Pritchard, Stephens, and Donnelly (2000) developed a Bayesian-based method that provides the posterior probability of each individual originating in each of a set of ancestral populations. The nature of ham samples, most of which come from crossbreeds rather than pure breeds, means that classical assignment or exclusion tests using pure breeds as reference populations are ineffective.

The purpose of the present work was to use multilocus genotypes to estimate the genetic composition of dry-cured Iberian ham offered to consumers in retail stores and restaurants.

2. Materials and methods

2.1. Samples

Two hundred and fifty dry-cured ham samples sold under the Iberian ham label were purchased by volunteers posing as anonymous customers. This avoided any special treatment from sellers who might have provided samples from specially chosen hams had they known the purpose of the experiment. The ham samples came from five different sampling sources: two categories of restaurant (high, $n = 30$; medium-low, $n = 65$), two categories of retail store (high, $n = 30$; medium-low, $n = 65$) in Madrid, and shops and bars in medium-sized villages around Madrid ($n = 60$). Restaurants and bars were chosen following a retrospective-type sampling procedure based on an ad hoc combination of criteria such as the score received in gastronomic guides and the socio-economic level of the surrounding neighbourhood. Once a restaurant/bar was selected, it was visited and the sample to be analysed collected.

All samples were frozen at −20 °C. DNA extraction was performed using the Genomic DNA Purification Kit (Gentra Systems, Minnesota, USA) following the manufacturer's instructions. Briefly, 10 mg of ham tissue were ground in a cell lysis solution, treated with proteinase K and RNase and the protein precipitated. DNA was precipitated in isopropanol and diluted to an approximate final concentration of 50 ng/μl. Samples were screened for 25 pig microsatellites selected from the 27 markers recommended by the FAO for pig biodiversity studies (FAO, 1998). Genotyping was performed using an ABI 377 XL automatic sequencer (Applied Biosystems, Foster City, CA, USA). Details of the primers used and other technical references can be found at http://www.projects.roslin.ac.uk/pigbiodiv/markers.html.

2.2. Reference breeds

A number of different breeds were taken into account as reference breeds when assessing the genetic composition of the ham samples. White coat breeds (Large White [$n = 83$] and Landrace [$n = 104$]) were chosen since they provide 90% of the genes in 90% of marketed dry-cured (but not Iberian) ham. Since Iberian-like products are popular and can provide extra profit for producers (Archibald, 1997), the Mangalica breed was also taken into account ($n = 25$). This breed, developed in Hungary and considered the best fat pig in Europe (Bodó et al., 2002), has been proposed for crossing with Durocs for the production of cured loin and hams. The Duroc breed was included ($n = 169$) since it is the breed of choice for crossbreeding with Iberian pigs ($n = 226$).

Reference samples for all these breeds were provided by the Dept. of Genetics, Cordoba University, within the framework of the Andalusian Research Program AGR218, "*Improvement And Conservation Of Genetic Resources In Domestic Animals*".

2.3. Statistical analysis

Exploratory analyses were first performed on the reference samples to determine the degree of genetic differentiation among the different breeds and the accuracy of traditional assignment methods.

Two different procedures were used to analyse the ham sample data. The first was an unsupervised method (Pritchard et al., 2000) which was used to determine the population structure. The term "unsupervised" alludes to the fact that no defined populations are used as an input in the model. However, the number of clusters (which may be understood in some cases as ancestral populations) into which the entire group of individuals is divided is introduced into the algorithm. The posterior probability of belonging to each of these clusters is then calculated for each individual, as well as the proportion of the genome of each individual belonging to each cluster. The number of clusters is understood as a random variable, so its like-

lihood can actually be calculated. Several values can be tried until a maximum likelihood figure is obtained, which may or may not coincide with the true number of reference populations.

Once this was done for the reference samples, these and the experimental ham samples were then examined together. The proportion of the genome of each individual in each cluster was then obtained. Ideally, samples from the same reference breed should all have genomes belonging almost entirely to one cluster. The ham samples, however, were expected to show genomes spread across different clusters, according to the cross that gave rise to them. For example, an individual whose father is pure breed A and whose mother is the result of a cross between breed A and B would have 75% of its genome belonging to the cluster to which breed A is allocated, and 25% belonging to the cluster to which breed B is allocated.

The above procedure is based on calculating the posterior distribution of the proportions of the individuals' genomes that belong to the different clusters, the allelic frequencies within these clusters, and the origins of each allelic copy in the individuals. $P(Z,P,Q|X)$ is then estimated, where X represents the genotypes, P the allelic frequencies, Q is given by $q_k^{(i)}$, which is the proportion of individual i's genome that originates in cluster k, and Z is given by $z_l^{(i,a)}$, which is the cluster of origin of individual i' s copy a of allele l.

Parameter estimations of the posterior distributions were performed after executing a Gibbs sampling algorithm. This provides approximate samples for the joint posterior distribution.

This approach offers the advantage of there being no need for information on the allelic frequencies of the reference populations (estimations that are sometimes biased by low sample size). Further, this method can be used with complex genetic models. This was particularly useful in the present situation since many samples were expected to come from Iberian × Duroc crossbreeds.

Structure software (Pritchard et al., 2000) was used for the unsupervised methodology. Besides obtaining estimates for the proportions of each individual's genome originating in the different clusters, the availability of likelihood values for different numbers of clusters allowed a maximum likelihood estimation for the number of ancestral populations forming the reference breeds.

To optimise the parameters of the model, multiple simulations were performed and a burn-in of 50,000 iterations and a run length of 200,000 finally used to carry out the genetic cluster analysis. Four parental populations were inferred assuming an admixture model.

The above procedure formed the first step of this analysis, the results of which conditioned the way the remainder of the methodology was used. The samples involved in this study were thought likely to reflect mixtures of different breeds in different proportions, a context difficult to deal with in a "supervised" framework. The term 'supervised' indicates that a specific set of reference breeds or populations characterised by reference samples needs to be used, and that the ham samples would be assigned to one of the reference breeds with a certain probability. Since many ham samples came from crossed animals, frequencies from dummy mixed breeds were calculated from the original breed frequencies, and these dummy breeds were included in the assignment process. Special emphasis was made in the crossings between Duroc and Iberian breeds, and crossings between them with 1:1, 1:2 and 2:1 ratios were included. The unsupervised method was therefore used to determine the proportion of each breed in the genome of each sample, and supervised methods were subsequently used to confirm each result using the mixed populations as references.

The main supervised approach followed here was developed by Baudouin et al. (2004). The probability of any given individual belonging to one of the breeds is calculated via the Bayes theorem as the conditional probability of finding the individual's genotype given the genetic structure of that breed, divided by the sum of all the conditional probabilities for all breeds, i.e., $P(b_i|g,p_i) = P(g|p_i)/\sum_{j=1}^{N} P(g|p_j)$, where b_i represents the event "belonging to breed i", g is the individual's genotype, p_i is the genetic information of breed i, N is the number of breeds, and a uniform prior is assumed for the b_i. The genetic structure of the breeds is given by the allelic frequencies. However since these frequencies are not population values but sampled values, a Bayesian view is adopted, and posterior estimates of the allelic frequencies are obtained from the sampled values.

Geneclass2 software (Piry et al., 2004) was used for the calculations. An added feature of this software is that the information from the allelic frequencies can be used to simulate genotypes from each of the reference populations, and to calculate the probability that each simulated genotype belongs to its population of origin. Thus, for each reference population, an empirical probability distribution for the probabilities of its genotypes is obtained. Therefore, for each real sample, the probability of its genotype belonging to a certain population can be compared with the distribution of the probabilities of that population's genotypes. A measure of the "rarity" of that sample within that population can thus be obtained in the form of an exclusion probability. In practice, this is calculated as the proportion of simulated genotypes that are more rare than that of the sample.

Finally, a more classical supervised approach was followed. The individual hypotheses of belonging to each individual breed and each simulated crossing were tested pairwise, and simulations performed to obtain powers for all the possible contrasts. One million genotypes were simulated under each hypothesis and likelihood ratio distributions estimated in order to calculate power figures. For example, to obtain the power for the Iberian pig breed (H_0) vs. the Duroc pig breed (H_1) at $\alpha = 0.01$, one million Iberian pig genotypes were simulated and likelihood ratios calculated for each. An empirical rejection threshold was then set for this α . Finally, one million Duroc genotypes

were simulated, likelihood ratios calculated, and the power estimated as the proportion of likelihood ratios exceeding the threshold.

3. Results

Table 1 shows the different within-diversity estimators for each population. These can be thought of as factors involved in the accuracy of discrimination among the reference populations. Table 2 shows the pairwise genetic distance matrix used to deduce information regarding the genetic discrimination between populations.

In general, the genetic distances among populations (average F_{ST} value of 0.18), the number of loci used, and the results of Cornuet et al. (1999), suggest that assignment accuracy in the present study could be high. The Duroc and Iberian populations were the closest and Mangalica the most distant from the rest of the breeds considered.

Alternatively, the relative position among these populations may be visualized using multivariate techniques (e.g., correspondence analysis, Lebart, Morineau, & Warwick, 1984) and be represented in a two-dimensional plane, such that the axes correspond to the inertias or fractions of information (Fig. 1). As expected, samples belonging to the same population tended to cluster together. Essentially, the first axis separates the white coat breeds from the Iberian pigs, while the second axis separates Mangalica from Duroc pigs and Landrance from Large White pigs.

Genetic cluster analysis, performed using *Structure* software, was undertaken assuming an admixture model (Pritchard et al., 2000). Table 3 shows the mean posterior estimates of the proportions of the genomes from the sampled populations that belong to the four clusters inferred. A high degree of clustering was observed since, for each of the sampled populations, 95% or more of their genomes was inferred to originate in a single cluster.

Table 1
Diversity averaged over the 25 microsatellite markers

	Expected heterozygosity	Observed heterozygosity	Number of alleles	Effective number of alleles
Mangalica	0.439	0.362	3.3	1.8
Large White	0.625	0.596	6.6	2.7
Landrace	0.635	0.591	7.2	2.7
Duroc	0.599	0.551	6.5	2.5
Iberian	0.601	0.522	8.4	2.5

Table 2
Pairwise genetic distances among reference populations expressed in terms of F_{ST}

	Mangalica	Large White	Landrace	Duroc
Large White	0.27[a]			
Landrace	0.23	0.14		
Duroc	0.27	0.20	0.14	
Iberian	0.21	0.22	0.15	0.13

[a] F_{ST} values can be interpreted as the total genetic variability explained by the genetic differences between two populations.

Table 3
Posterior distributions of the proportions (q_k) of the sampled populations' genomes belonging to the four inferred clusters

Sampled populations	Ancestral populations			
	D	I	LW	M
Mangalica	0.005	0.031	0.004	0.960
Large White	0.004	0.010	0.983	0.004
Landrace	0.005	0.011	0.979	0.005
Duroc	0.979	0.011	0.007	0.003
Iberian	0.003	0.987	0.008	0.002

D, I, LW and M refer to Duroc, Iberian, Landrace-Large White, and Magalica.

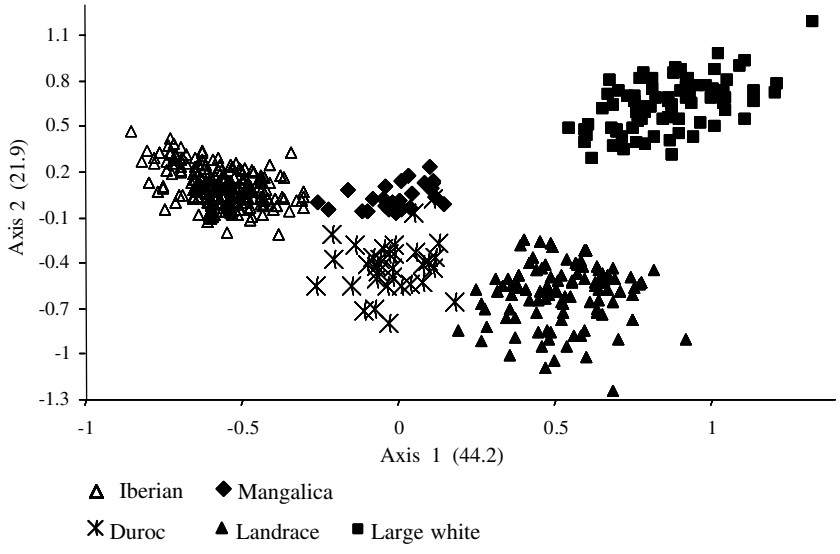

Fig. 1. Correspondence analysis of allele frequencies for the 25 microsatellite markers typed in 5 reference pig breeds. Axis contributions to the total inertia are shown in brackets.

Allowing for more than four clusters provided a higher posterior likelihood for the number of clusters (i.e., $\log P(K|X)$). However, this did not yield stable results, and multiple runs of the *Structure* program produced varying cluster configurations.

When cluster analysis was performed taking into account individuals from the reference Duroc and Iberian breeds and from the simulated mixed populations, and assuming two clusters, the proportions of the genome of each of the two sampled populations (Duroc and Iberian) and of the three simulated populations (Iberian × Duroc 25%, 50% and 75%) almost perfectly matched the two clusters (Table 4). The clustering algorithm therefore estimated the ancestry of the simulated populations with reasonable accuracy.

Having reference pig individuals from known populations, and wishing to classify ham samples of unknown origin (presumably mixtures of some of the reference populations), the use of prior information was thought likely to improve the accuracy of any inference made (Beaumont et al., 2001). Therefore, to estimate the proportion of each individual ham sample's genome that originated in each ancestral population, prior population information was used (option USEPOPINFO = 1) so that the program could take into account the reference population from which each individual was sampled in the clustering algorithm and the admixture model. Table 5 shows the centrality and dispersion parameter estimates of the mean of the posterior distribution of each individual's admixture coefficient.

The results show that the contribution of Duroc and Iberian pigs to the genetic composition of the ham samples was, on average, of similar magnitude. In 48 of the individual ham samples (19%) (data not shown), the percentage of Iberian origin was <20%. In another 28 (11%), the percentage Duroc origin was >70%, and in 19 (8%), a significant percentage ($\geqslant$25% in the corresponding cluster) of white coat genomes was found. The last two types of sample cannot be considered Iberian ham under current regulations.

The genetic admixture composition of all ham samples not meeting Iberian ham parameters was confirmed using the assignment methods in the *Geneclass2* package (Piry et al., 2004). Dummy mixed breeds with allelic frequencies calculated from the original reference breeds' frequencies were included as reference populations for these analysis. Only five samples (~10%) were assigned to a different category than that estimated by the *Structure* program, but in all cases were adjacent categories (e.g., 25% Duroc–75% Iberian and 50% Duroc–50% Iberian).

Table 6 shows the discrimination powers for two type I error levels, 0.01 and 0.001, obtained when contrasting different combinations of hypotheses, including crossings of Duroc and Iberian populations as described above. These results clearly show that ham samples not meeting current regulations (i.e., those with >50% Duroc origin) can be detected with high probability (>90%) except when the contrast is H0 50% Duroc vs. H1 75% Duroc. This is due to the great similarity between these populations. When applied to the real ham samples these tests also confirmed the assignments made.

Finally, to establish a link between ham quality and store/restaurant category, the mean of the posterior distribution of each individual's Iberian origin coefficient (q_k) was taken as a dependent variable, and the premises where the sample was obtained as the independent variable. Thus, the differences in the q_k between these premises can be estimated and their significance tested. The ham samples with the highest q_k values came from the highest category retail stores, while those with the poorest coefficient of Iberian origin came from stores in the villages surrounding Madrid (Table 7). This result was confirmed when the distribution of illegal ham samples, when split into those with an excess

Table 4
Posterior distributions of the proportions (q_k) of the sampled (Duroc and Iberian) and simulated populations' genomes belonging to two ancestral clusters

	Cluster	
	1	2
Duroc	0.039	**0.961**
75% Duroc	0.221	0.779
50% Duroc	0.448	0.552
25% Duroc	0.726	0.273
Iberian	0.931	0.069

Table 5
Average proportions ($\overline{q_k}$) and standard errors (in brackets) of individual ham sample genomes belonging to the four inferred clusters

Ancestral populations			
D (Duroc)	I (Iberian)	LW (White coat colour)	M (Mangalica)
0.44 (0.20)	0.40 (0.19)	0.10 (0.09)	0.06 (0.05)

Table 6
Discrimination power for different type I errors

H1:	Duroc		Dur75 × Iber25		Dur50 × Iber50	
H0:	0.01	0.001	0.01	0.001	0.01	0.001
Iberian	1	1	1	0.9999	0.9980	0.9921
Dur25 × Iber75	1	1	0.9748	*0.8820*	*0.4293*	*0.1785*
Dur50 × Iber50	0.9999	0.9978	*0.4072*	*0.1525*		

H_0 refers to the population considered in the null hypothesis while H_1 is the population in the alternative hypothesis. Values lower than 0.90 are italicised. Dur 25% × Iber 75% indicates a simulated population in which the proportions of genes coming from Duroc and Iberian pigs are 25% and 75%, respectively.

Table 7
Average posterior distribution of the Iberian origin coefficient (q_k) across retail outlet categories

Sampling unit	q_k
High level retail store	0.553[a]
High level restaurant	0.426[b]
Medium-low level retail store	0.401[bc]
Medium-low level restaurant	0.360[bc]
Villages surrounding Madrid	0.344[c]

Different letters indicate significant differences at 0.05.

Table 8
Distribution of ham samples with an excess of Duroc genome or with a significant presence of white coat genomes

Sampling unit	Category	Percentage of ham samples in which Duroc genome represent more than 70%	Percentage of ham samples in which a significant proportion ($\geqslant$25%) of white coloured breed is present
Bar-Restaurant	High	11	0
	Medium-low	13	10
Retail store	High	9	0
	Medium-low	15	9
Villages surrounding Madrid	All	14	14

of Duroc genes or from a white breed cross, was portrayed with respect to the establishments from which the samples were collected (Table 8).

4. Discussion

Unsupervised methods such as the Bayesian based-model proposed by Pritchard et al. (2000) allow the inclusion of population mixes. However, as indicated by Baudouin et al. (2004), this requires a certain degree of experience in their use; their routine employment is somewhat difficult. On the other hand, the more classical assignment methods have the problem that any anonymous sample will always be assigned to one of the reference populations, so if crosses are not included as reference populations, samples that are actually a mix of different ''pure'' breeds will be incorrectly assigned. To overcome this problem a type of *exclusion* test has been proposed (Cornuet et al., 1999) that rejects anonymous samples as belonging to a population when their likelihood is below a certain threshold. However, a problem remains when the anonymous samples are from crossed animals; the solution proposed in this paper is to simulate synthetic populations and to use these to confirm the results obtained with the unsupervised method.

The high regard among consumers for Iberian ham, and the limited possibilities of increasing supply, result in a high market price. Although new regulations have been proposed by the Spanish government (Boletin Oficial del Estado, 2001, 2003) regarding the precise genetic composition of animals to be used for making this type of ham, a certain degree of indiscriminate use of the term 'Iberian ham' is to be expected.

The *c*20% of ham samples that apparently did not meet official demands for this product (because of too great a Duroc contribution or the significant presence of genes from non-acceptable breeds such as Landrace or Large White) are the consequence of problems of very different origin. While an excess of Duroc genes is a problem of the production system (e.g., crossbreeding strategies using F1 *Duroc* × *Iberian* females and Duroc males, or the existence of 'Iberian' populations showing a certain genetic mixture), the presence of genes from white coat breeds is plain fraud. Higher frequencies of such fraud might be expected in restaurants, where consumers usually do not have the opportunity to see the ham from which their portion will be cut, and where the potential profit from knowingly selling these illegal products is very high (Table 7). The lowest quality products (<35% of sampled genes from Iberian pigs) were mostly found in the villages around Madrid.

Two important aspects must be tackled in future surveys: (a) linking ham samples to labelling information on the product, and (b) determining whether 50% of the Iberian genome is female or male in origin.

5. Conclusions

In conclusion, determining the genetic composition of mixed samples requires complex statistical analysis and abundant molecular information. No single statistical tool can provide a convincing estimate of the genetic composition of a sample, and therefore several must be used to confirm results. The present analysis confirmed the existence of a significant level of illegal dry-cured Iberian ham. Consumers appear to have more chance of buying better quality Iberian ham if they purchase it from a retail store than if they order it in a restaurant. Probably this is a consequence of the difficulties that consumers have in supervising the product when it is consumed in a restaurant; in contrast, shoppers can check the appearance of the product and its label.

Acknowledgements

We are grateful to the Genetics Laboratory of the Animal Production Dept., Complutense University of Madrid, for collecting the ham samples. We thank Adrian Burton for linguistic assistance. This work was funded by the Regional Government of Madrid (*Consejería de Sanidad y Consumo*).

References

Archibald, A. L., 1997. (Coordinator). (1997). Trait gene mapping in a cross between two divergent pig breeds – Duroc and Mangalica pigs. Scientific Progress Report. Contract no.: IC15CT96-0915.

Baudouin, L., Piry, S., & Cornuet, J. M. (2004). Analytical Bayesian approach for assigning individuals to populations. *Journal of Heredity, 95*, 217–224.

Beaumont, M., Gottelli, D., Barratt, E. M., Kitchener, A. C., Daniels, M. J., Pritchard, J. K., et al. (2001). Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology, 1*, 319–336.

Bodó, I., Komlósi, I., Mihók, S., Szabó, P., Jávor, A., & Béri, B., (2002). Genetic resources and their management in Hungary. In *Seventh world congress on genetics applied to livestock production*, 19–23 August 2002, Montpellier, France.

Boletin Oficial del Estado, (2001). Real Decreto 1083/2001, de 5 de cotubre, por el que se aprueba la norma de calidad para el jamón ibérico, paleta ibérica y caña de lomo ibérico elaborados en España. B.O.E., 247, pp. 37830–37833.

Boletin Oficial del Estado, (2003). Orden APA/213/2003, de 10 de febrero por la que se establecen normas de desarrollo del Real Decreto 1083/2001, de 5 de octubre, por el que se aprueba la norma de calidad para el jamón ibérico, paleta ibérica y caña de lomo ibérico elaborados en España. B.O.E., 36, pp. 5305–5310.

Branciari, R., Nijman, I. J., Plas, M. E., Di Antonio, E., & Lenstra, J. A. (2000). Species origin of milk in Italian mozzarella and Greek feta cheese. *Journal of Food Protection, 63*, 408–411.

Carraspiso, A. I., Bonilla, F., & García, C. (2003). Effect of crossbreeding and rearing system on sensory characteristics of Iberian ham. *Meat Science, 65*, 623–629.

Cornuet, J. M., Piry, S., Luikart, G., Estoup, A., & Solignac, M. (1999). New methods employing multilocus genotypes to select or exclude populations as origin of individuals. *Genetics, 153*, 1989–2000.

Dieguez, E., (1992). Historia, evolución y situación actual del cerdo ibérico pg 9–35. In El Cerdo Ibérico, la naturaleza, la dehesa. Ministerio de Agricultura, Pesca y Alimentación, Madrid.

FAO, (1998). Domestic Animal Diversity Information System Stage 2.0. CD-ROM and <http://www.fao.org/dad-is> Rome, FAO.

Flores, J., Biron, C., Izquierdo, L., & Nieto, P. (1988). Characterization of green hams from Iberian pigs by fast analysis of subcutaneous fat. *Meat Science, 23*, 253–262.

Kijas, J. M. H., Moller, M., Plastow, G., & Andersson, L. (2001). A frameshift mutation in MC1R and a high frequency of somatic reversion cause black spotting in pigs. *Genetics, 158*, 779–785.

Kijas, J. M. H., Walles, R., Tornsten, A., Chardom, P., Moller, M., & Andersson, L. (1998). Melanocortin receptor (MC1R) mutations and coat colour in pigs. *Genetics, 150*, 1177–1185.

Lebart, L., Morineau, A., & Warwick, K. (1984). *Multivariate descriptive statistical analysis*. New York: Wiley.

López-Bote, C. J. (1998). Sustained utilization of the Iberian pig breed. *Meat Science, 49*(Suppl.), S17–S27.

Mayoral, A. I., Dorado, M., Guillén, M. T., Robina, A., Vivo, J. M., Vázquez, C., et al. (1999). Development of meat and carcass quality characteristics in Iberian pigs reared outdoors. *Meat Science, 52*, 315–325.

Petrón, M. J., Muriel, E., Timón, M. L., Martín, L., & Antequera, T. (2004). Fatty acids and triacylglycerols profiles from different types of Iberian dry-cured hams. *Meat Science, 68*, 71–77.

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L., & Estoup, A. (2004). GeneClass2: a software for genetic assignment and first generation migrants detection. *Journal of Heredity, 95*, 536–539.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*, 945–959.