Animal Genetics, 2000, **31**, 117–122

Characterisation of Iberian pig genotypes using AFLP markers

C Óvilo, M T Cervera, C Castellanos, J M Martínez-Zapater

Summary

The use of the AFLP (amplified fragment length polymorphism) technique for the characterisation of highly inbred Iberian pig breed genotypes and the detection of strain-specific polymorphisms is demonstrated. Twelve different primer combinations were used on individual DNA samples from animals belonging to two black hairless Iberian pig strains, Guadyerbas and Coronado. These amplification reactions allowed the detection of more than 1700 amplification products of which 26 were identified as strain-specific markers, present in all individuals of one strain and absent in the other. Comparison of male and female amplification products within one strain also allowed the identification of 8 male-specific amplified bands. AFLP showed a great power of marker detection due to a high multiplex ratio and high reproducibility. Comparison of similarity and co-ancestry coefficient matrices also showed the usefulness of AFLP markers to estimate genetic relationships between individuals pigs.

Keywords: AFLP, Iberian pig

Introduction

An increasing interest in Iberian pig production is arising as a consequence of their high quality standards regarding meat and fat. Traditional strains of Iberian pigs are difficult to distinguish at the genetic level and the origins of commercial populations are generally unknown. Differences in morphological and physiological characters are observed, including variation in production and reproductive traits (Pérez-Enciso & Gianola 1992). The black hairless types of Iberian pigs, which are characterised by very high levels of body fat and are usually managed without genealogical control, are endangered as a consequence of being held in small herds with high inbreeding levels that reduces their reproductive performance.

Correspondence: Cristina Óvilo (e-mail: ovilo@inia.es).

Accepted 5 December 1999

The problem posed by the genetic identification of the pigs and conservation of genetic variability within highly inbred herds could be addressed with the development of strain specific markers and with the application of rapid genotyping analysis to the individuals of the herd. This would allow the selection of the progenitors for inter and intra-strain crosses from those that carry a larger proportion of the existing variability. Many PCR-based fingerprinting methods have been developed. These methods make use of synthetic oligonucleotides to amplify specific or arbitrary DNA sequences of the genome thereby producing a characteristic pattern of amplified fragments (Caetano-Anollés 1996; Cushwa & Medrano 1996; Welsh & McClelland 1991; Williams *et al.* 1990). Of these, AFLP which is based on the selective amplification of restriction fragments ligated to adapters of known sequence (Vos *et al.* 1995), has the highest multiplex ratio, allowing the screening of a very high number of amplified restriction fragments with a few amplification reactions.

The objective of this study was to evaluate the potential of AFLP markers for identification of highly inbred Iberian pig genotypes and their possible application in the conservation and management of populations with unknown genealogy.

Materials and methods

A total number of 14 black hairless Iberian pigs were analysed, including five males and five females from the Guadyerbas strain and four males from the Coronado strain. The Guadyerbas strain represents the black hairless type from the Guadiana valley and is conserved in an experimental herd as a genetically isolated population since 1945 (Rodrigañez *et al.* 1997). This population has a high inbreeding level ($F = 0{\cdot}36$) and this makes it advisable to cross the Guadyerbas pigs with other black hairless populations in order to obtain a new synthetic line with better reproductive traits and lower conservation costs. The Coronado strain is a black hairless Iberian pig population of very small size conserved in one commercial herd without pedigree records.

C Óvilo
M T Cervera
C Castellanos
J M Martínez-Zapater
Departamento de Mejora Genética y Biotecnología, Subdirección General de Investigación y Tecnología, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Ctra de la Coruña km 7, 28040 Madrid, Spain
M T Cervera
J M Martínez-Zapater
Departamento de Genética Molecular de Plantas, Centro Nacional de Biotecnología, CSIC, Campus de la Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain

Genomic DNA was extracted from blood samples using the technique described by Wu *et al.* (1995). AFLP analysis was performed as described by Vos *et al.* (1995) with some specific modifications (Cervera *et al.* 1996). For each genomic DNA sample, 500 ng were double digested with 12 U *Eco*RI and 8 U *Mse*I, RL buffer (10 mM Tris–HAc, 10 mM MgAc, 50 mM Kac, 5 mM DTT) in a final volume of 35 μl. The reaction was incubated for three hours at 37 °C. Next, the ligation mixture was added, consisting of 5 μl of a solution containing 5 pmol *Eco*RI-adapter, 50 pmol *Mse*I-adapter, 8 mM ATP and 1·2 U T4 DNA ligase in RL buffer and incubated for three hours at 37 °C and at 4 °C overnight. After ligation of adapters, DNA fragments were diluted to a final volume of 200 μl with water.

Amplification of DNA fragments was performed in two steps. The first step or preamplification was a PCR reaction with two primers which are complementary to the adapters *Eco*RI and *Mse*I with one additional selective 3′ nucleotide. The primer combinations used in this preamplification were: *Eco*RI + A/*Mse*I + C, *Eco*RI + A/*Mse*I + T, *Eco*RI + A/*Mse*I + G, *Eco*RI + C/*Mse*I + C, *Eco*RI + C/*Mse*I + T. PCR reactions were performed in 20 μl containing 3 μl of diluted DNA fragments, 0·4 U *Taq* DNA polymerase (Boehringer Mannheim), 30 ng *Eco*RI primer, 30 ng *Mse*I primer, 4 mM dNTPs and 1X PCR buffer (Boehringer Mannheim). The cycle profile was as follows: 30 s at 94 °C, 60 s at 60 °C and 60 s at 72 °C, for 28 cycles. PCR products were examined in 0·8% agarose gels and diluted 10-fold with water.

The second PCR step or selective amplification was performed using primers containing the same sequences as those used in the preamplification but with three selective 3′ nucleotides. Selective nucleotide combinations tested with *Eco*RI/*Mse*I primers were: ACC/TGT, AAG/TTC, ACT/GCT, CAC/CTA, ACC/GTA, AAG/CAT, ACT/TTC, CAC/TTC, AAG/CTG, AAG/TGT, CAC/CAT, CAC/TGT. Selective amplification reactions were carried out in 20 μl containing 5 μl of the diluted preamplification products, 0·4 U *Taq* DNA polymerase (Boehringer Mannheim), 5 ng *Eco*RI primer radioactively labelled with [γ^{33}P] ATP, 30 ng *Mse*I primer, 2 mM dNTPs and 1X PCR buffer. The PCR profile started with one cycle of 30 s at 94 °C, 30 s at 65 °C and 60 s at 72 °C, in the following 12 cycles the annealing temperature decreased 0·7 °C per cycle and in the last 23 cycles the profile was: 60 s at 94 °C, 30 s at 56 °C and 60 s at 72 °C. The preamplification and selective amplification reactions were performed in a PE-9600 thermocycler (Perkin-Elmer Corp, USA). The amplification products (3 μl) were analysed on 4·5% acrylamide/bisacrylamide 19 : 1, 7·5 M urea and 1X TBE gels using a standard DNA sequencing unit. Electrophoresis were performed at constant power (90 W) for two hours and 30 min and AFLP patterns were visualised by autoradiography.

Polymorphic fragments were scored as 1 for presence and 0 for absence and genetic similarities were calculated using two pairwise similarity indexes: Dice coefficient (Sneath & Sokal 1973):

$$GS(ij) = 2a / (2a + b + c)$$

where a is the number of polymorphic DNA fragments common to both individuals, b is the number of fragments present in i and absent in j, and c is the number of fragments present in j and absent in i, and Rogers and Tanimoto (RT) coefficient (Rogers & Tanimoto 1960):

$$GS(ij) = m / (n + u)$$

where m is the number of matches (present or absent DNA fragments common to both individuals), u is the total number of fragments only possessed by individuals i and j, respectively, and n is the total sample size ($u + m$). The resultant matrices were subjected to cluster analysis by the unweighted pair-group method analysis (UPGMA) and a dendrogram was constructed from the clustering results with the TREE program. Good-fit of cluster analysis (between the similarity matrix and the dendrogram obtained) was measured by calculating the cophenetic correlation between the similarity matrix and the cophenetic matrix for each analysis.

Matrix comparisons were made to determine the agreement between the AFLP similarity matrix, calculated using the RT coefficient, and the co-ancestry coefficient matrix using the Mantel test (Mantel 1967). These calculations were performed using the computer program NTSYS-PC version 1·80 (Rohlf 1993).

Results and discussion

The use of molecular techniques may contribute to our knowledge of the genetic structure of Iberian pig populations. The development of strain specific markers is a first step which would enable objective identification of Iberian pig strains, assist in conservation of the endangered strains and allow the traceability of strains and breeds. Identification of strain-specific markers can in some cases be a difficult task, especially when the genotypes to be

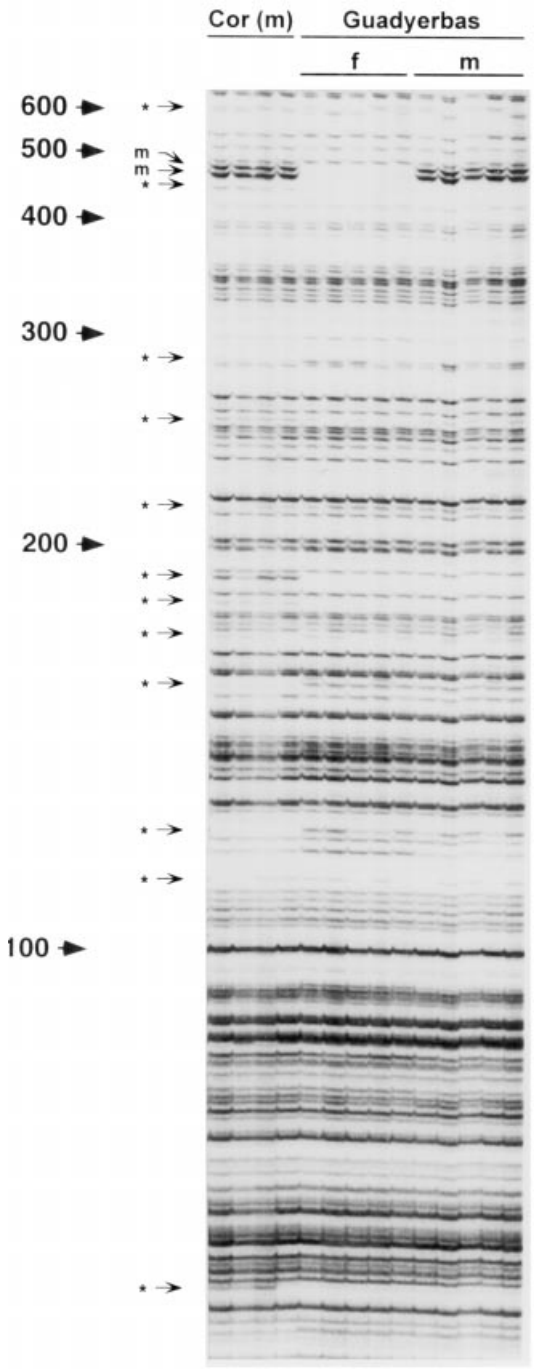

Fig. 1. AFLP analysis of 14 Iberian pigs belonging to two different strains: four Coronado males [Cor(m)] and 10 Guadyerbas [five females (f) and five males (m)]. The DNA fingerprint was generated with primer combination *Eco*RI + AAG/*Mse*I + TTC. The large arrows indicate the size marker positions. Arrows indicated with an m correspond to male specific markers, and those preceded by a * indicate polymorphic bands.

distinguished are very closely related. This was the situation in our work, in which the two strains of Iberian pigs used (Guadyerbas and Coronado) belong to the same type of black hairless pigs.

In order to identify strain-specific polymorphisms, we tested 12 different AFLP primer combinations on 14 individual pig samples belonging to the two strains. These primer combinations yielded a total of 1733 scorable bands, ranging from 70 to 700 nucleotides in length, representing an average of 140 scorable amplified bands per primer combination (see Fig. 1 for an example of the autoradiograms obtained). Out of these 1733, only 106 (6%) were polymorphic among the samples analysed (Table 1). Within those polymorphic bands, 26 were present in all individuals of one strain but absent in the other, 12 were specific for the Coronado strain and 14 others for the Guadyerbas strain. Comparison of the amplified bands present in males and females of the Guadyerbas strain revealed eight markers present in all male samples but absent in females (Table 1).

These results show the effectiveness of the high multiplex ratio of AFLP for rapid identification of candidate strain-specific markers with a few amplification reactions. The presence of the specific amplified bands in all the individuals of one strain and their complete absence from the individuals of the other strain strongly indicates that these are specific amplified products and not simple amplification artefacts, and that they could be fixed in the strain. This later comment however, requires confirmation through the analysis of more individual samples belonging to the same strain and comparison with a larger number of genetically diverse strains. Unfortunately, genotyping a larger number of individuals of this type of inbred population, for which the application of this technology is highly useful, is not always possible. This is the case for the black hairless Iberian pig strains, due to the limited size of the available herds.

Less than 0·3% of the amplified bands were male specific. This percentage is in the same order of magnitude as the percentage represented by the Y chromosome in the pig karyotype (1·2%). The slight under-representation could either be due to errors in the karyotype determination or to the presence of common genomic regions between the X and Y chromosomes. The detection of male specific markers based on AFLP technology could be a useful tool to determine the sex of animal embryos mainly in cattle, but could also be used in the certification of high quality meat products exclusively derived from males in other species such as ducks and geese.

Both strain and male specific markers have been previously obtained in Iberian pig breed using RAPDs, microsatellites or the RDA tech-

Table 1. AFLP results: number of amplified DNA fragments, polymorphisms and specific markers obtained with the 12 primer combinations tested

Primer combination	Number of bands	Total number of polymorphisms	Coronado markers	Guadyerbas markers	Male markers
E + ACC/M + TGT	140	8	0	2	0
E + AAG/M + TTC	173	13	2	3	2
E + ACT/M + GCT	121	9	0	0	3
E + CAC/M + CTA	112	7	0	0	0
E + ACC/M + GTA	129	7	2	2	0
E + AAG/M + CAT	150	4	1	1	0
E + ACT/M + TTC	155	15	2	2	0
E + CAC/M + TTC	141	8	1	0	2
E + AAG/M + CTG	162	6	1	1	0
E + AAG/M + TGT	165	14	1	1	1
E + CAC/M + CAT	139	5	0	1	0
E + CAC/M + TGT	146	10	1	1	0
Total	1733	106	12	14	8

nique (Castellanos *et al.* 1996; Óvilo *et al.* 1998; Pérez-Pérez & Barragán 1998). The use of AFLP represents a faster and cheaper alternative to scan the genome in the search for those markers when it comes to work on closely related individuals. A similar strategy can be applied to any species without any previous information on its genome sequence, which represents a useful tool in the study of wild species.

In order to genetically identify Guadyerbas and Coronado individuals and to establish the genetic relationships between them, a similarity matrix was calculated using Dice coefficient based on the presence or absence of polymorphic bands (excluding the eight male-specific markers). A dendrogram illustrating the data of genetic similarity obtained with the 12 primer combinations used in this study is displayed in Fig. 2. All the individuals could be easily distinguished and two main and well-differentiated branches were observed. One of them included all the Guadyerbas individuals while the other contained all the Coronado samples, indicating a clear distinction between both groups. The cophenetic correlation between the similarity matrix and the cophenetic matrix was very high (0·97, $P = 0{\cdot}002$) indicating a very good fit of the cluster analysis.

The possibility of extinction is a major issue for Iberian pig strains due to the reduced size of

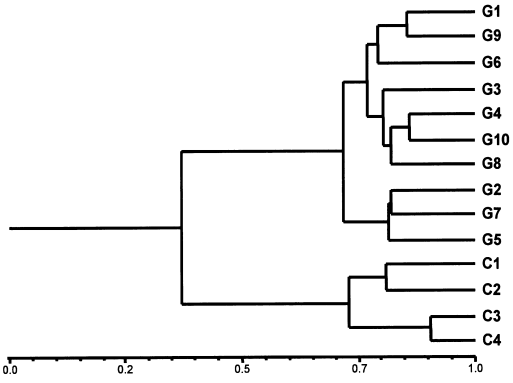

Fig. 2. Dendrograms representing the genetic similarity among individuals of both Iberian pig strains. Samples named with G are from Guadyerbas animals and those named with C are from Coronado animals. The dendrogram was constructed using the UPGMA clustering method on the Dice estimates of genetic similarities based on AFLP data

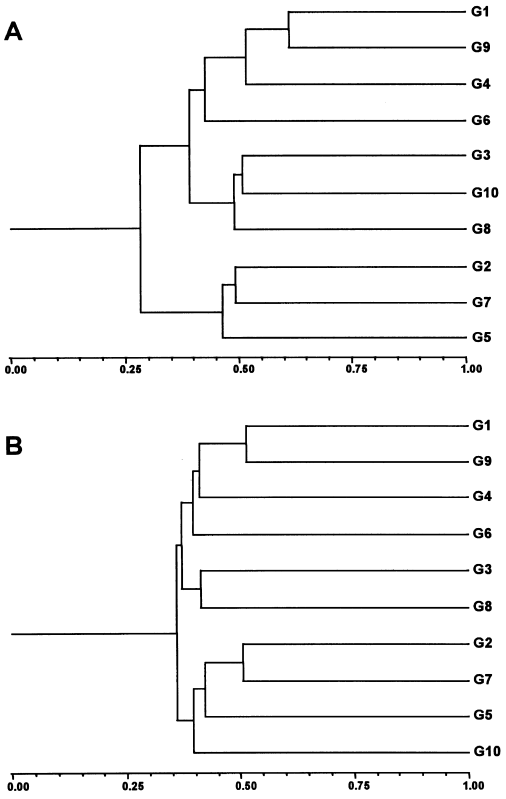

Fig. 3. Genetic relationships among 10 Guadyerbas individuals based on UPGMA cluster analysis using either RT similarity coefficient to analyse AFLP data (A) or co-ancestry coefficients (B).

Table 2. Comparison of AFLP similarity values, calculated using RT coefficient (over the diagonal), and co-ancestry coefficients (below the diagonal) between all the pairs of Guadyerbas individuals analysed

	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
G1	**1.000**	0.316	0.380	0.538	0.289	0.370	0.235	0.429	0.613	0.449
G2	0.362	**1.000**	0.380	0.370	0.441	0.333	0.493	0.316	0.250	0.333
G3	0.383	0.355	**1.000**	0.400	0.315	0.324	0.195	0.485	0.380	0.508
G4	0.392	0.363	0.360	**1.000**	0.273	0.316	0.220	0.408	0.493	0.515
G5	0.362	0.419	0.355	0.363	**1.000**	0.241	0.485	0.273	0.195	0.361
G6	0.384	0.353	0.357	0.386	0.353	**1.000**	0.316	0.370	0.587	0.316
G7	0.362	0.504	0.352	0.346	0.419	0.355	**1.000**	0.266	0.176	0.250
G8	0.426	0.355	0.411	0.360	0.355	0.357	0.352	**1.000**	0.351	0.493
G9	0.511	0.368	0.354	0.424	0.368	0.412	0.372	0.354	**1.000**	0.648
G10	0.361	0.419	0.359	0.381	0.419	0.351	0.343	0.359	0.364	**1.000**

existing populations. The establishment of conservation programs requires pedigree information to select the founder animals, however, this information is lacking in most cases. Toro *et al.* (1998) proposed the use of molecular markers to identify founder animals and minimise the losses of genetic variability in the first generations. In order to test the efficiency of the use of molecular information in the prediction of co-ancestry we compared the matrix of genetic similarity obtained for the Guadyerbas individuals, using AFLP markers, with their co-ancestry matrix. Due to the high inbreeding level of the Guadyerbas only 50 polymorphic AFLP markers with frequencies ranging from 0·2 to 0·8 were used (Lynch & Milligan 1994). Due to the close relationship between Guadyerbas individuals both presence and absence of DNA fragments were considered as indication of similarity using the RT coefficient as similarity index. Table 2 shows the values of the AFLP similarity, calculated using the RT coefficient, and the co-ancestry coefficients between all the pairs of Guadyerbas individuals. Figure 3 shows dendrograms derived from the AFLP-based similarity data (Fig. 3A) and the co-ancestry coefficients (Fig. 3B). Both dendrograms showed similar relationships between individuals, allowing the distinction of three main groups: individuals G1, G9, G4 and G6 were placed on one group, individuals G3 and G8 on another group and individuals G2, G7 and G5 in a third group. Only individual G10 did not hold an identical position in both cluster analyses. This could suggest an incorrect parentage assignment for this sample. The statistical significance of the correlation between both matrices was verified with the Mantel permutation test, (Mantel 1967). Although the results of the test were in the lowest level of correlation that can be considered statistically significant (0·61, $P = 0·002$) we can explain these results based on the differences of the methodology employed. The results suggest that in the absence of pedigree information, genetic markers could be useful to estimate genetic relationships between individuals and thus to design marker assisted conservation programs, as a way to select the progenitors within a breed which allow to preserve the existing variation and avoid gene losses.

Although AFLPs are technically more demanding than microsatellites, this study indicates that AFLP is a fast and reliable method to scan the genome in search of specific polymorphisms. In a preliminary work, focused on the application of microsatellite analysis to identify highly inbred Iberian pig populations, we found a lower power for the detection of strain-specific alleles compared with AFLP technique (results not published). With the analysis of 19 microsatellite markers we were not able to find any strain-specific allele fixed in the population, although we found some alleles present only in one strain but with low frequencies, which are therefore less useful to discriminate strains or breeds.

Moreover, intra-strain polymorphisms detected with AFLP technique provide high amount of information about the genetic relationships between individuals, even in the case of highly inbred populations like pig populations used in this study. The use of microsatellites has been recommended by FAO for biodiversity studies when different strains or nonrelated populations have to be genotyped. AFLP markers overcome microsatellites when the genetic analysis of closely related individuals is required, and both AFLP and microsatellites could complement each other for the study of other type of outbred populations.

Although AFLP has been considered an expensive and complex technique, it becomes highly cost-efficient when time and expenses are estimated on the basis of the number of markers scored (Costa *et al.* 1999). Moreover,

the strain specific AFLP markers identified can be transformed into SCAR markers (Paran & Michelmore 1993) which are easily screened by PCR assays using specific pairs of primers.

Acknowledgements

The authors would like to thank M.C. Rodríguez, L. Silió, and M.A. Toro, for providing information of co-ancestry among Guadyerbas samples and for all their comments and contribution along the development of the work, Carlos A. Malpíca for critical reading of the manuscript, Javier Paz-Ares, Antonio Leyva and Emilio Rodríguez-Cerezo for the laboratory facilities, and Angel Sanz for photographic art work. C. O. was funded by a predoctoral fellowship from the INIA and M.-T.C. by a postdoctoral fellowship from the Comunidad Autónoma de Madrid. Support to research activity at Centro Nacional de Biotecnología is provided through a specific agreement CSIC-INIA.

References

Caetano-Anollés, G. (1996) Scanning of nucleic acids by in vitro amplification: new developments and applications. *Nature Biotechnology* **14,** 1668–74.

Castellanos, C., Barragán, C. & Rodríguez, M.C. (1996) Detection of four porcine Y-specific markers by RAPD. *Animal Genetics* **27,** 433–4.

Cervera, M.T., Gusmao, J., Steenackers, M. *et al.* (1996) Identification of AFLP molecular markers for resistance against *Melampsora larici-populina* in *Populus*. *Theoretical and Applied Genetics* **93,** 733–7.

Costa, P., Pot, D., Dubos, C. *et al.* (1999) A genetic map of Maritime pine based on AFLP, RAPD and protein markers. *Theoretical and Applied Genetics* **100 (1)**, 39–48.

Cushwa, W.T. & Medrano, J.F. (1996) Applications of the Random Amplified Polymorphic DNA (RAPD) assay for genetic analysis of livestock species. *Animal Biotechnology* **7** (1)**,** 11–31.

Lynch, M. & Milligan, B.G. (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology* **3,** 91–9.

Mantel, N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* **27,** 209–20.

Óvilo, C., Barragán, C., Castellanos, C., Rodríguez, M.C., Silió, L. & Toro, M.A. (1998) *Application of molecular markers to Iberian pig genotype characterization*. IVth International Symposium on Mediterranean Pigs, Evora, Portugal.

Paran, Y. & Michelmore, R.W. (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and Applied Genetics* **85,** 985–93.

Pérez-Enciso, M. & Gianola, D. (1992) Estimates of genetic parameters for litter sizes in six strains of Iberian pigs. *Livestock Production Science* **32,** 283–93.

Pérez-Pérez, J. & Barragán, C. (1998) Isolation of four pig male-specific DNA fragments by RDA. *Animal Genetics* **29,** 157–8.

Rodrigañez, J., Silió, L., Toro, M.A. & Rodríguez, C. (1997) *Fifty years of conservation of a black hairless strain of Iberian pigs*. Proceedings of the International Symposium on Mediterranean Animal Germplasm and Future Human Challenges. EAAP Publication No. 85**,** pp. 183–6.

Rogers, D.J. & Tanimoto, T.T. (1960) A computer program for classifying plants. *Science* **132,** 115–8.

Rohlf, F.J. (1993) *NTSYS-PC Numerical Taxonomy and Multivariate Analysis System, Version 1.8.* Exeter Publishers, Setauket, NY.

Sneath, P.H.A. & Sokal, R.R. (1973) *The Principles and Practice of Numerical Classification*. W. H. Freeman, San Francisco, CA.

Toro, M.A., Silió, L., Rodrigáñez, J. & Rodríguez, C. (1998) The use of molecular markers in conservation programmes of live animals. *Genetics, Selection and Evolution* **30,** 585–600.

Vos, P., Hogers, R., Bleeker, M. *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23,** 4407–14.

Welsh, J. & McClelland, M. (1991) Genomic fingerprints produced by PCR with consensus tRNA gene primers. *Nucleic Acids Research* **19,** 861–6.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. & Tingey, S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18** (22)**,** 6531–5.

Wu, Q., Chen, M., Buchwald, M. & Phillips, R.A. (1995) A simple, rapid method for isolation of high quality genomic DNA from animal tissues. *Nucleic Acids Research* **23** (24)**,** 5087–8.