Archivos de Zootecnia
ISSN: 0004-0592
pa1gocag@lucano.uco.es
Universidad de Córdoba
España

Groenen, M.A.M.; Joosten, R.; Boscher, M.-Y.; Amigues, Y.; Harlizius, B.; van der Poel, J.J.; Crooijmans, R.
The use of microsatellite genotyping for population studies in the pig with individual and pooled dna samples
Archivos de Zootecnia, vol. 52, núm. 198, 2003, pp. 145-155
Universidad de Córdoba
Córdoba, España

THE USE OF MICROSATELLITE GENOTYPING FOR POPULATION STUDIES IN THE PIG WITH INDIVIDUAL AND POOLED DNA SAMPLES

EL USO DEL GENOTIPADO DE MICROSATÉLITES PARA ESTUDIOS DE POBLACIONES EN EL CERDO CON MUESTRAS DE ADN INDIVIDUAL Y MEZCLADO

Groenen, M.A.M.[1], R. Joosten[1], M-Y Boscher[2], Y. Amigues[2], A. Rattink[1], B. Harlizius[1], J.J. van der Poel[1] and R. Crooijmans[1]

[1]Animal Breeding and Genetics group. Wageningen University. Marijkeweg 40. 6709 PG Wageningen. The Netherlands.
[2]Laboratoire d'analyses génétiques pour les espèces animales (LABOGENA). Domaine de Vilvert. Jouy en Josas. 78352 cedex. France.

ADDITIONAL KEYWORDS

Molecular genetics. Genotyping. Characterisation.

PALABRAS CLAVE ADICIONALES

Genética molecular. Genotipado. Caracterización.

SUMMARY

In the present paper the results obtained from the genotyping of DNA microsatellite markers in 60 populations of European Pig are shown. The genotypings have been performed on individual and pooled DNA samples, standing out the technical efficiency of both methods in the characterisation of the European pig biodiversity.

RESUMEN

En el presente trabajo se muestran los resultados obtenidos del genotipado de marcadores microsatélites del ADN en 60 poblaciones de cerdos europeos. Estos genotipados han sido desarrollados sobre muestras de ADN individual y mezclado, destacándose la eficiencia técnica de ambos métodos para la caracterización de la biodiversidad de los cerdos europeos.

INTRODUCTION

Quantitative assessment of genetic diversity within and between populations is an important tool for decision making in genetic conservation plans. The most widely used method to quantify this genetic diversity is by genotyping a selection of unrelated individuals from the populations under investigation. In principle any marker for which there has been described genetic variation can be used for such studies. In earlier studies markers that were commonly used were mainly based on blood groups and easily identifiable enzymes in the blood (Van Zeveren *et al*., 1990; Rohrer *et al*., 1997). However, for technical reasons, these markers were not very well suited for large-scale population studies with

large numbers of markers and individuals. The development of DNA based markers in the last two decades has revolutionised the possibilities to monitor genetic diversity of populations by making it feasible to screen large numbers in a relatively short time. One type of marker that has been intensely used for population studies in the last 10 years, are the so-called microsatellite or single sequence repeat markers. These markers are very abundant, show a high degree of polymorphism and can be analysed by means of PCR, enabling a high degree of automation of the genotyping analysis. In pigs microsatellites have already been used in a number of studies to address the biodiversity in commercial as well as rare breeds (Van Zeveren *et al*., 1995; Martínez *et al*., 2000; Laval *et al*., 2000).

MATERIAL AND METHODS

TYPING PROTOCOL FOR INDIVIDUAL GENOTYPING OF MICROSATELLITES

The PCR reactions were performed in a total volume of 12 µl containing 80 ng of genomic DNA, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris.HCl pH=8.3, 1 mM Tetramethylammoniumchloride (TMAC), 0.1 percent triton X-100, 0.01 percent gelatin, 200 µM dNTP, 0.25 Unit Goldstar polymerase (Eurogentec S.A., Belgium), 2.3 pmoles of each primer and covered with 10 µl of mineral oil (Sigma). The PCR protocol was as follows: 5 min denaturation at 95°C; 36 cycles of [30", 95°C/30", 50-60°C/30", 72°C] finally followed by 2 minutes at 72°C. PCR reactions for the different microsatellite markers were pooled and a mixture of up to 1µl pooled PCR products was added to 3.2 µl loading buffer (which contained the GENESCAN-350 TAMRA internal standard and formamide; final concentration of 80 percent). The sample was denatured at 95°C and loaded on a 6 percent denaturing polyacrylamide gel (sequagel 6; National Diagnostics) on an ABI 373A sequencing machine (12 cm well-to-read; loading 4 µl). PCR products of different markers from DNA of the same animal were pooled in such a way that each marker signal on the ABI automated sequencers has a peak height of about 1000. The fragment sizes were calculated relative to the GENESCAN-350 TAMRA with the GENESCAN fragment analysis software (Perkin Elmer, Applied Biosystems Division). Genotyping was performed using the Genotyper 2.0 software.

TYPING PROTOCOL FOR GENOTYPING OF MICROSATELLITES ON POOLED DNA SAMPLES

Genotyping of pooled DNA was performed essentially as described for individual typings with the following modifications:

(1)Microsatellite markers were pooled in such a way to avoid overlap of alleles, even if markers differed for the fluorescent dye. This means that in general 3 or 4 markers are analysed on the same gel.

(2)Longer gels (36 cm well-to-read distance) in combination with a different gel matrix (long ranger 5.75 percent acrylamide, 7 M Urea) were used to increase the resolution of the individual peaks.

(3)In stead of allele frequencies,

peak frequencies were calculated based on the area under the peaks (Using the GENOTYPER 2.0 software). Previous results on the analysis of pooled chicken DNA revealed that it was more reliable to use peak frequencies, rather then to correct for stutter bands.

RESULTS AN DISCUSSION

MICROSATELLITE MARKER PANEL SELECTED FOR GENOTYPING

Previously we have described the selection of 27 microsatellites to be used as a standard panel for population studies in the pig (Laval *et al*., 2000). These markers were chosen based on their quality, size, polymorphism and location on the porcine genome (Archibald *et al*., 1995; Rohrer *et al*., 1994, 1996) as proposed by the FAO (Barker *et al*., 1998). Quality was based mainly on the absence of any known null alleles, the sharpness of the peaks on automatic sequencers and robustness of the amplification reaction. Markers were chosen to maximise the genome coverage provided by these 27 markers. All pig chromosomes, except chromosome 18 were represented by this marker set. Furthermore, large chromosomes often were represented by two different markers. When two markers were on the same chromosome they were chosen at a minimum distance of at least 30 cM. The third criterion used for the selection of the markers was the size of the amplified product and the fluorescent label of the amplification product. This eventually enabled us to design 3 different sets with 9 markers each that could be used for multiloading and simultaneous analysis on ABI automatic sequencers. To avoid overlap between adjacent markers labelled with the same fluorescent dye, markers were combined in such a way that the size difference between the smallest allele of the larger marker was at least 30 bp longer then the largest allele of the smaller marker. The 27 markers and their distribution over these three sets are shown in **table I** (set I-III).

Within the EU PigBioDiv project, we decided to increase the coverage of the markers across the pig genome. Eventually 23 markers were added to the original selection of 27 markers, resulting in a total of 50 markers divided over 6 sets that could be analysed simultaneously on ABI automatic sequencers (see **table I**). Although the same criteria were applied for selecting the markers as those that were used for selecting set I to III, it was no longer possible to use the criterion for a minimum distance of 30 cM between the markers. Nevertheless, the majority of the markers were located at least 20 cM apart.

MICROSATELLITE GENOTYPING ON INDIVIDUAL SAMPLES

As previously observed in the EU funded PigMaP pilot project on diversity, microsatellite genotyping must be organised so that each marker is typed in only one laboratory, to avoid problems in allele calling. This is particularly true if the analysis is performed on gel based systems such as on the ABI373 and ABI377. As a consequence, it is recommended that when data from different facilities must be combined, each facility genotypes

Table I. *Pig microsatellite markers selected for populations studies. The number of alleles and the allele size range are based upon the PiGMaP (Archibald* et al.*, 1995) and USDA (Rohrer* et al.*, 1996) reference populations. Distance in cM indicates the distance between that particular marker and the preceding marker in the table. Set refers to the markers that are analysed simultaneously on ABI automatic sequencers.* (Marcadores microsatélites de cerdo seleccionados para los estudios de poblaciones. El número de alelos y el rango de tamaño de los alelos se basan en las poblaciones de referencia del PIGMAP (Archibald *et al.*, 1995) y el USDA (Rohrer *et al.*, 1996). Las distancias en cM indican las distancias entre ese marcador particular y el marcador precedente en la tabla. Las series se refieren a los marcadores que son analizados simultáneamente con secuenciadores automáticos ABI).

Marker	Chr arm	Set	#Alleles	Allele size range		Distance[1] (cM)
				Min	Max	
CGA	1p	I	12	250	320	-
S0155	1q	III	6	150	166	45
SW1828	1q	VI	6	95	105	24
SW240	2p	III	8	96	115	-
S0226	2q	II	9	181	205	20
SW72	3p	III	8	100	116	-
SW902	3q	IV	7	195	214	40
S0002	3q	III	7	190	216	45
S0227	4p	II	10	231	256	-
S0301	4p	V	6	254	266	20
S0217	4q	IV	5	145	165	45
S0097	4q	IV	8	208	248	50
S0005	5q	III	10	205	248	-
IGF1	5q	III	7	197	209	30
SW2406	6p	V	7	220	256	-
SW1067	6q	VI	7	144	175	50
SW122	6q	II	10	110	122	15
S0228	6q	II	12	222	249	20
S0025	7p	IV	9	104	120	-
SW632	7q	I	9	159	180	100
S0101	7q	I	6	197	216	30
SW2410	8p	VI	9	104	120	-
S0225	8q	II	8	170	196	82
S0178	8q	II	4	110	124	45

Table I. *Pig microsatellite markers selected for populations studies.* (Marcadores microsatélites de cerdo seleccionados para los estudios de poblaciones). *(continuación).*

Marker	Chr arm	Set	#Alleles	Allele size range		Distance[1] (cM)
				Min	Max	
SW911	9p	I	9	153	177	-
SW174	9q	IV	5	123	133	90
SW830	10p	VI	7	176	205	-
S0070	10q	IV	7	265	299	60
SW951	10q	II	5	125	133	35
SW2008	11p	IV	6	95	105	-
S0386	11q	III	10	156	174	55
S0143	12p	VI	5	148	166	-
S0090	12q	II	4	244	251	75
SWR1941	13	VI	7	202	223	-
S0068	13	I	10	211	260	45
SW769[2]	13	VI	7	106	140	55
S0215	13	I	9	135	169	5
SW857	14	III	6	144	160	-
SW295	14	V	8	109	139	30
S0355	15	I	14	243	277	-
SW1111	15	V	6	165	183	25
SW936	15	I	13	80	117	50
SW742	16	V	9	193	224	-
S0026	16	III	8	92	106	30
SWR1004[2]	17	V	5	147	167	-
SW24	17	I	8	96	121	5
SW1023	18	V	5	84	117	-
SW787	18	IV	8	153	165	25
SW2476	Xq	V	5	88	106	-
S0218	Xq	II	8	164	184	35

[1]Distance on the USDA map.
[2]Marker SW769 (Chr 13) and SWR1004 (Chr 17) are included though they are both located at a distance of only 5cM from the next marker.

the same individuals for different markers. It is possible that the increased standardisation on the capillary based sequencers (ABI 3100, 3700, 3730) will allow for an easier transferability of the data across different laboratories, eliminating the need to take this approach in the future.

Individual genotyping has been done for 60 breeds (considering the French and British samples of *Meishan* as two different breeds) for all 50 markers described in **table I**. Genotyping of sets I, II, III and VI was performed at Labogena on an ABI3700, whereas genotyping of sets IV and V was done at the Wageningen University (WU) on an ABI373 sequencer. Although overall failure rates vary slightly between the different typing laboratories this seems to be somewhat higher at WU. The data shown in **figure 1** suggest that the DNA quality is the major factor influencing the success

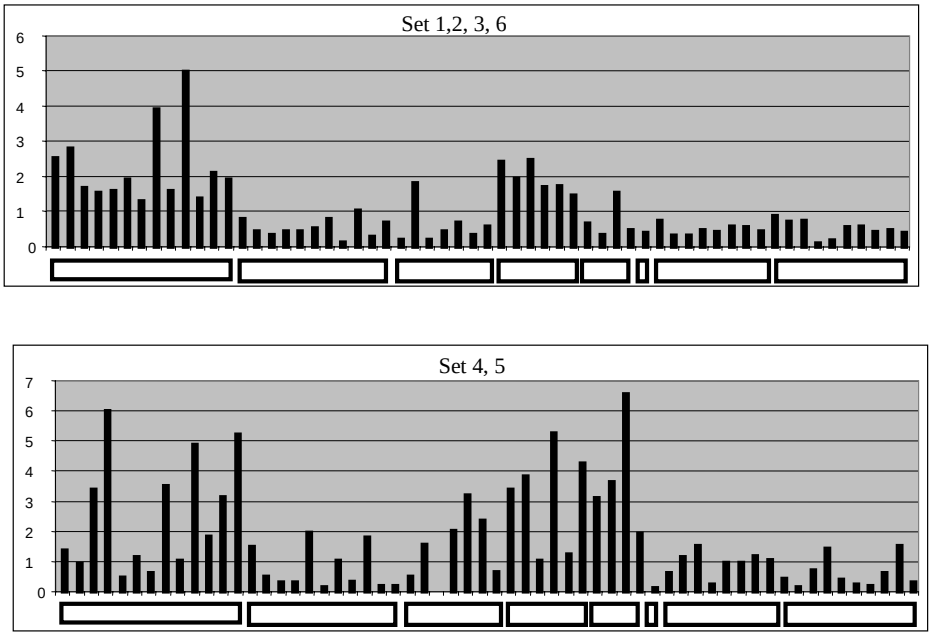

Figure 1. *Overview of the failure rates for the populations typed by Labogena (sets I, II, III and VI) and Wageningen University (sets IV and V). The percentage of the animals that were not genotyped is indicated on the y-axis. Populations are grouped along the x-axis according to the participant responsible for the blood collection and DNA isolation.* (Una revisión de las tasas de fallo para las poblaciones tipificadas en Labogena (series I,II,III y VI) y la Universidad de Wageningen (Series IV y V). El porcentaje de animales que no fueron tipificados están representados en el eje de las Y. Las poblaciones se agrupan a lo largo del eje de las X de acuerdo al participante responsable del muestreo de sangre y el aislamiento de ADN).

rate for typing as this is very dependent upon the origin of the samples. This is also in agreement with the DNA typing results using AFLP (see Plastow *et al*., 2003, in this Proceeding). For example, the average failure rate of the genotypes for sets IV and V was 14.9 percent but for individual populations this varied from 0.1 percent for the Italian Duroc line 1 (ITSSDU01) to 65.3 percent for the Spanish Negro Canario (ESSSNC01). This clearly demonstrates the key importance of the procedures for collecting, handling and storing of the blood and isolation of the DNA. Based on other projects the most critical factor is the collection of the blood. We also observed clear differences in the performance of the different markers used, in particular those in sets IV to VI. As has been outlined above, the markers with the highest quality had already been used for sets I to III and it was known that the overall quality of sets IV to VI was somewhat lower. These factors probably are responsible for the difference seen between the two laboratories. For the markers in sets IV and V, failure rates varied from 6.5 percent for SW2008 to 32 percent for S0301. However, it has to be realised that these numbers are slightly biased because the failure rate is also influenced by the quality of the DNA samples and markers that are less robust are more susceptible to poor DNA quality.

MICROSATELLITE GENOTYPING ON DNA POOLS

Genotyping of microsatellites on pooled samples will result in a drastic reduction in the number of genotypes that have to be performed and thus in a concomitant reduction in time and money involved. DNA typing in pools has been successfully applied in chicken (Khatib *et al*., 1995; Crooijmans *et al*., 1996) and human (Pacek *et al*., 1993; LeDuc *et al*., 1995) but had not yet been tested in pigs. Within the EU PigBioDiv project we tested whether it also would be feasible to use this approach for pig microsatellites. For this approach, DNA from a number of animals (1 population) is combined and genotyping is performed on the pooled sample. The PCR signal of the individual alleles is directly correlated with the frequency of that particular allele. The frequency of a particular allele is calculated based upon the area under the peak of that allele relative to the total of the area under the peak for all alleles.

An aliquot of 200 ng of DNA from all the individual animals from each of the breeds was combined and analysed as a pool with all the 50 microsatellites of set I to VI. PCR reactions for the pools were essentially the same as for individual typings, with a total of 80 ng of the pooled DNA per reaction. Previously, it was observed that abundant alleles (high signal) can result in some read through in the other dyes complicating the analysis of the other markers. Therefore, for the pooled analysis, the 50 markers were combined into 15 sets containing 3 to 4 markers to avoid overlap of alleles of different markers. The total number of breeds analysed using the pools was 72, since in addition to the 60 original breeds it included one Italian breed (MR01) and 11 breeds that were used previously in the PiGMaP pilot project (Laval *et al.*,

2000). Although all 50 markers were analysed on the pooled samples, eventually, we were only able to generate useful genotyping data for 20 microsatellites. This in part is due to the problems with the quality of the DNA for some of the samples but the main reason is the fact that the markers in sets I to VI were selected for genotyping on individual samples. For genotyping of microsatellite markers on pooled DNA samples, the markers have to fulfil additional criteria (Crooijmans *et al.*, 1996) which are not met by the majority of the markers in sets I to VI. In particular the degree of stuttering, the sharpness of the peaks on the automatic sequencer and the absence of alleles that differ only by 1 bp are additional important criteria for using pooled DNA samples. If these criteria are not met, it is often not possible to identify individual alleles and the allele frequency based on the area under the peak is not calculated correctly. In **figure 2** and **3** respectively examples are shown of markers that are suitable (S0217 and SW787) and a marker that is not suitable (S0101) for analysis on DNA pools. In **figure 3**, the genotyping results are shown for two populations with similar allele

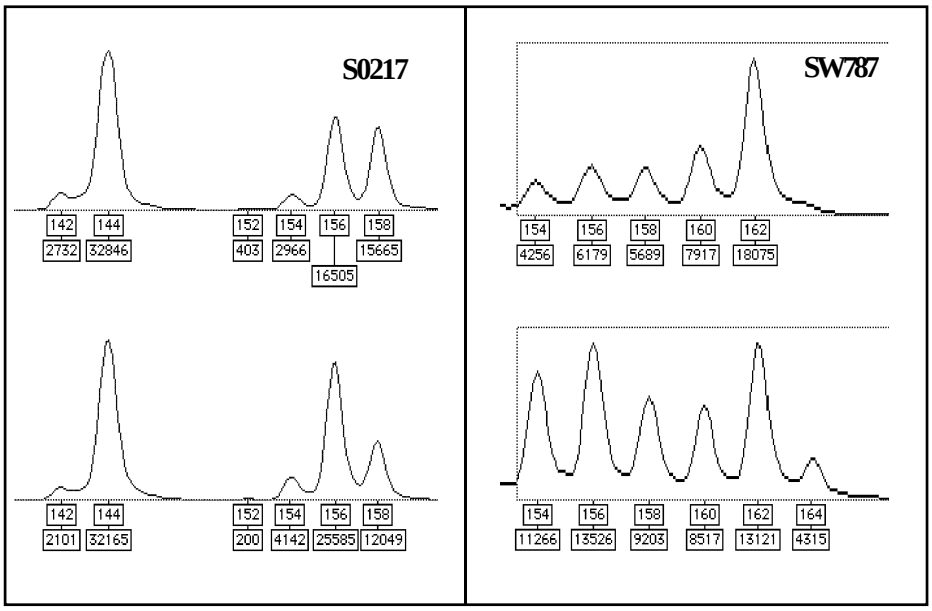

Figure 2. *Example of typing results on DNA pools for successful microsatellite markers. The results for two populations for the markers S0217 and Sw787 are shown.* (Ejemplo de los resultados de tipificación sobre los pools de ADN para los marcadores microsatélites satisfactorios. Se muestran los resultados de dos poblaciones para los marcadores S0217 y Sw787).

frequencies. However, in one of the populations shown, the peaks for alleles 210 and 212 are not recognised as individual peaks by the GENOTYPER software and as a result these are included in the area for the peaks for

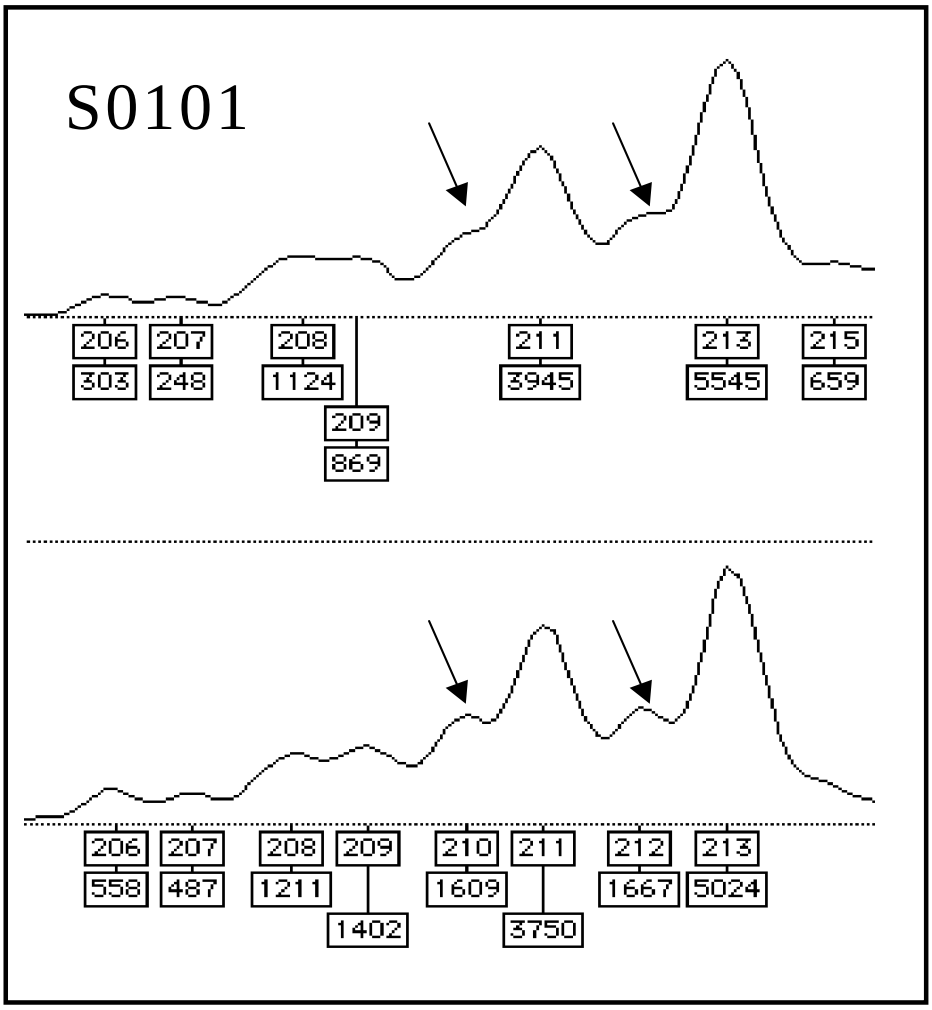

Figure 3. *Example of typing results on DNA pools for an unsuccessful marker (S0101). The arrows indicate alleles that are not correctly identified in the population at the top.* (Ejemplo de los resultados de tipificación para un marcador insatisfactorio (S0101). Las flechas, en la parte superior, muestran los alelos que no son correctamente identificados en la población).

Table II. *Pig microsatellites typed on pooled DNA samples. The number of populations that could be analysed successfully for these markers is given. The total number of populations genotyped was 72, considering the French and British samples of* Meishan *as two different populations.* (Microsatélites de cerdo tipificados sobre muestras mezcladas de ADN. Se ofrece el número de poblaciones que pudieron ser analizadas exitosamente con esos marcadores. El número total de poblaciones genotipadas fue 72, considerando las muestras francesas e inglesas de Meishan como dos poblaciones diferentes).

Marker	Populations Analysed
SW911	69
S0002	58
S0070	62
S0090	61
S0097	56
S0217	56
S0218	52
S0227	68
SW1828	59
SW2008	68
SW2406	55
SW2476	51
SW72	67
SW787	67
SW830	67
SW857	62
SW902	66
SW936	68
SW951	68
SWR1004	49

alleles 211 and 213.

Table II gives an overview of the markers for which it was possible to analyse the majority of the populations. In total 85 percent of the population-marker combinations could be analysed successfully for these 20 microsatellites.

Taking all of the results together, the main outcome of the genotyping on pooled DNA samples is that this method seems to be feasible, but that it requires a selection of a different (sub) set of microsatellites. Furthermore, it is to be expected that the use of capillary sequencers rather then gel based sequencers will further improve the resolution, and thus will make genotyping on pooled samples a feasible cost-effective alternative for an initial scan of the populations.

ACKNOWLEDGEMENTS

We acknowledge Dennis Milan for his help in defining the panel of the 23 additional microsatellites and Louis Ollivier and Graham Plastow for critically reading the manuscript. This research was financially supported by the European Union (contract Bio4-CT98-0188).

REFERENCES

Archibald A.L., J.F. Brown, S. Couperwhite, H.A. McQueen, D. Nicholson, C.S. Haley, W. Coppieters, A. van der Weghe, A. Stratil, A-K. Wintero, M. Fredholm, N.J. Larsen, V.H. Nielsen, D. Milan, N. Woloszyn, A. Robic, M. Dalens, J. Riquet, J. Gellin, J-C. Caritez, D. Hue, G. Burgaud, L. Ollivier, J-P. Bidanel, M. Vaiman, C. Renard, H. Gelderman, R. Davoli, D. Ruyter, E.J.M. Verstege, M.A.M. Groenen, W. Davies, B. Hoyheim, A. Keiserud, L. Andersson, H. Ellegren, M. Johansson, L. Marklund, R.J. Miller, D.V. Anderson Dear, E.

Signer and A.J. Jeffreys. 1995. The PiGMaP Consortium linkage map of the pig (*Sus scrofa*). *Mamm. Genome,* 6: 157-175

Barker, J.S.F., W.G. Hill, D. Bradley, M. Nei, R. Fris and R.K. Wayne. 1998. Measurement of domestic animal diversity (MoDAD): original working group report. FAO Rome.

Crooijmans, R.P.M.A., A.F. Groen, J.A. Van Kampen, S. Van der Beek, J.J. Van der Poel and M.A.M. Groenen. 1996. Microsatellite polymorphism in commercial broiler and layer lines estimated using pooled blood samples. *Poultry Sci.,* 75: 904-909.

Khatib, H., A. Darvasi, Y. Plotski and M. Soller. 1995. Determining relative microsatellite allele frequencies in pooled DNA samples. *PCR Methods Applic.,* 4: 13-18.

Laval, G., N. Iannucelli, C. Legault, D. Milan, M.A.M. Groenen, E. Giuffra, L. Andersson, P.H. Nissen, C.B. Jorgensen, P. Beeckman, H. Geldermann, J.L. Foulley, C. Chevalet and L. Ollivier. 2000. Genetic diversity of eleven European pig breeds. *Genet. Sel. Evol.,* 32: 187-203.

LeDuc, C., P. Miller, J. Lichter and P. Parry. 1995. Batched analysis of genotypes. *PCR Methods Applic.,* 4: 331-336.

Martínez, A.M., J.V. Delgado, A. Rodero and J.L. Vega-Pla. 2000. Genetic structure of the Iberian pig breed using microsatellites. *Anim. Genet.,* 31: 295-301.

Pacek, P., A. Sajantil and C. Syvanen. 1993. Determination of allele frequencies of loci with length polymorphism by quantitative analysis of DNA amplified from pooled samples. *PCR Methods Applic.,* 2: 313-317.

Plastow, G., K. Siggens, M. Bagga, B. Brugmans, H. Heuven and J. Peleman. 2003. Utilization of AFLP® for genetic distance analysis in pigs. *Arch. Zootec.*, 52: 157-164.

Rohrer, G.A., L.J. Alexander, J.W. Keele, T.P. Smith and C.W. Beattie. 1994. A microsatellite linkage map of the porcine genome. *Genetics,* 136: 231-245.

Rohrer, G.A., L.J. Alexander, Z. Hu, T.P. Smith, J.W. Keele and C.W. Beattie. 1996. A comprehensive map of the porcine genome. *Genome Research,* 6: 371-391.

Rohrer, G.A., P. Vögeli, G. Stranzinger, L.J. Alexander and C.W. Beattie. 1997. Mapping 28 erythrocyte antigen, plasma protein and enzyme polymorphisms using an efficient genomic scan of the porcine genome. *Anim. Genet.,* 28: 323-330.

Van Zeveren, A., Y. Bouquet, A. Van de Weghe and W. Coppieters. 1990. A genetic blood marker study on 4 pig breeds. I-Estimation and comparison of within-breed variation. II- Genetic relationship between the populations. *J. Anim. Breed. Genet.,* 107: 104-118.

Van Zeveren, A., L. Peelman, A. Van de Weghe and Y. Bouquet. 1995. A genetic study of four Belgian pig populations by means of seven microsatellite loci. *J. Anim. Breed. Genet.,* 112: 191-204.