Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/sjar

Spanish Journal of Agricultural Research 2008 6(2), 248-251
ISSN: 1695-971-X

Short communication. Inbreeding and homozygosity in Iberian pigs

E. Alves, C. Barragán and M. A. Toro*

Departamento de Mejora Genética Animal. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA). Ctra. A Coruña, km 7,5. 28040 Madrid. Spain

Abstract

Genetic markers could provide a tool to estimate inbreeding in absence of known pedigree. In the present work, 62 individuals from two related strains of Iberian pigs (32 from Guadyerbas and 30 from Torbiscal), with a deeply known pedigree, have been genotyped for 49 microsatellite loci. The inbreeding values, calculated from pedigree, have been compared to the molecular inbreeding values or homozygosity, calculated from molecular markers. Simulations were also carried out to facilitate the interpretation of the results. The correlation between genealogical and molecular inbreeding was negative for the Guadyerbas (–0.32), low for the Torbiscal (0.19) and substantial for all the animals together (0.69). In conclusion, it seems preferable to use pedigree information whenever available, and limiting the use of markers to verify, correct, complete or even implement pedigree recording.

Additional key words: Guadyerbas strain, microsatellites, molecular markers, Torbiscal strain.

Resumen

Comunicación corta. Consanguinidad y homocigosidad en cerdos Ibéricos

Cuando no se conoce el pedigrí los marcadores genéticos pueden proporcionar una estima de la consanguinidad. Se han genotipado 49 microsatélites en 62 individuos con genealogía conocida de dos estirpes relacionadas de cerdos Ibéricos (32 de la estirpe Guadyerbas y 30 de la Torbiscal). Se han comparado los valores de la consanguinidad, calculada a partir del pedigrí, con los valores de la consanguinidad molecular u homocigosis calculada a partir de los marcadores moleculares. También se han llevado a cabo simulaciones para facilitar la interpretación de los resultados. La correlación entre la consanguinidad genealógica y molecular fue negativa en Guadyerbas (–0,32), baja en Torbiscal (0,19) y sustancial para todos los animales en conjunto (0,69). Se concluye que es preferible utilizar la información genealógica siempre que esté disponible y limitar el uso de marcadores para verificar, complementar o incluso establecer el pedigrí.

Palabras clave adicionales: estirpe Guadyerbas, estirpe Torbiscal, marcadores moleculares, microsatélites.

Inbreeding and the associated inbreeding depression play a key role in practical applications of genetics. In the last decade, the great development of molecular markers has lead some authors to suggest that it would be possible to infer genealogical inbreeding from homozygosity of the markers or at least to determine if ranking individuals for molecular homozygosity will be equivalent to ranking them for the genealogical inbreeding coefficient (Pemberton, 2004).

In previous works, we have compared the genealogical and the molecular coancestry (Toro *et al*., 2002, 2003). In the present study, the inbreeding individual values of 62 pigs, calculated from pedigree, have been compared to the molecular inbreeding value or homozygosity, calculated from molecular markers. Animals belonging to two related strains of Iberian pigs, with deeply known pedigree, have been genotyped for 49 microsatellite loci. Simulations have also been carried out to clarify the results obtained.

The two strains of Iberian pigs considered in the present work, Guadyerbas and Torbiscal, belong to an early conservation programme established in 1945 at «El Dehesón del Encinar» (Oropesa, Toledo, Spain). Details are given in Odriozola (1976) and Toro *et al*. (2002). The complete genealogy of all animals is available back to 1945, with approximately 20 generations from the founders until the 62 animals of both

* Corresponding author: toro@inia.es
Received: 05-09-07; Accepted: 11-04-08.

sexes genotyped in this study (32 Guadyerbas and 30 Torbiscal).

Fourty nine microsatellite loci, distributed among the 18 autosomal chromosomes, were genotyped as described in Toro *et al*. (2002). The number of alleles per loci ranged from 2 to 9, with an average of 4.24 alleles and the heterozygosity values ranged from 0.19 to 0.79, with an average of 0.58.

Genealogical inbreeding was calculated tracing the pedigree back 20 generations to common ancestors. The molecular inbreeding or homozygosity was simply calculated as the proportion of homozygous loci (identical by state). There is a very well known relationship between the genealogical (F) and the molecular inbreeding (F_M):

$$1 - F_M = (1 - \sum p_i^2)(1 - F) \qquad [1]$$

where $1 - F_M$ and $1 - F$ are the heterozygosities by state and by descent in the actual population and $1 - \sum p_i^2$ is the expected heterozygosity by state in the founder population. The relationship simply indicates that $1 - F$ express the heterozygotes actual frequency relative to the heterozygotes frequency in the base population. This formula provides a simple way to inferring genealogical inbreeding from the molecular one as

$$F = \frac{F_M - \sum p_i^2}{1 - \sum p_i^2} \qquad [2]$$

However, there is a practical problem because the p_i values in the base population are unknown and must be estimated from its values in the actual population. Therefore several approaches have been developed to address this estimation (reviewed in Toro *et al*., 2002).

Simulations were carried out using the gene dropping method. A 2125 cM pig genome was simulated with 49 markers located at their map positions or with 212 markers placed evenly across the autosomes (details are given in Toro *et al*., 2002). The number of simulation runs was always 100. Results were averaged across simulations.

The mean, minimum and maximum values of both genealogical and molecular inbreeding are given in Table 1. The mean value of the genealogical inbreeding was greater in the Guadyerbas than in the Torbiscal strains but its variability (measured by the standard deviation and maximum/minimum values) was lower. The value of molecular inbreeding or homozygosity was higher than the genealogical one as expected from formula [1]. An estimate of genealogical inbreeding could be obtained from [2] using the actual observed value of the expected heterozygosity $1 - \sum p_i^2 = 0.58$ of genotyped individuals, instead of the value in the founder population that, obviously, is not known. The estimates obtained were 0.19, 0.01 and 0.10 for the Guadyerbas, Torbiscal and the whole population, respectively. These values are remarkably lower than the observed values of 0.35, 0.15 and 0.25 and even include negative values for some individuals.

In the simulation results, the statistics of molecular inbreeding (49 loci) mimic reasonably well the empirical data (Table 1). When an estimate of genealogical inbreeding is obtained using $1 - \sum p_i^2$ (formula [2]) as a scale factor, with this value calculated from the genotyped population, we obtained the very biased values of 0.22, –0.03 and 0.10 for the Guadyerbas, Torbiscal and the whole population, respectively. On the other hand, when the true values of gene frequencies in the founder population are used we recover the correct values of 0.35, 0.15 and 0.25.

The correlation between the genealogical and the molecular inbreeding is given in Table 2. In the pig

Table 1. Statistics of genealogical and molecular inbreeding

Population	Inbreeding coefficient	Mean	Standard deviation	Maximum	Minimum
Guadyerbas	Pedigree	0.35	0.01	0.37	0.33
	Molecular	0.53	0.05	0.63	0.43
	Molecular (simulation)	0.56	0.07	0.70	0.43
Torbiscal	Pedigree	0.15	0.02	0.21	0.13
	Molecular	0.44	0.05	0.53	0.27
	Molecular (simulation)	0.43	0.07	0.57	0.29
All animals	Pedigree	0.25	0.10	0.37	0.13
	Molecular	0.48	0.07	0.63	0.27
	Molecular (simulation)	0.50	0.10	0.57	0.29

Table 2. Correlation coefficient between the genealogical and the molecular inbreeding

Population	Pig data	Simulation (49 loci)		Simulation (212 loci)	
		Mean	Standard deviation	Mean	Standard deviation
Guadyerbas	–0.32	0.13	0.16	0.25	0.14
Torbiscal	0.19	0.21	0.18	0.38	0.15
All population	0.69	0.70	0.11	0.90	0.03

data the value was negative (–0.32) for the Guadyerbas, low (0.19) for Torbiscal but substantial (0.69) when animals were taken altogether. In the simulations assuming 49 loci the observed values of correlation follow the same pattern and the standard deviation are big and, therefore, negative values are not unexpected. When the number of loci is increased to 212 the correlation values improve, mainly when all animals are considered.

The availability of molecular markers in the last decade has prompted their use to infer relative inbreeding among individuals especially in those situations where pedigrees are difficult to collect or are of questionable quality (Pemberton, 2004). However, there are problems in estimating genomic heterozygosity using only a few molecular markers. This is well known since Nei (quoted by Chakraborty, 1981) showed that the expected correlation between the heterozygosity of the genome and of a sample is approximately $\sqrt{r/n}$, where n is the number of unlinked loci in the genome and r the number of loci assayed. For example a sample of 20 markers from a genome of 20,000 genes would lead to an expected correlation of 0.03. This simple relationship seems to have been overlooked until the recent empirical reviews and theoretical update of Pemberton (2004) where she pointed out that the critical factor influencing the correlation value is the variance of the genealogical inbreeding.

Comparisons have also been made between coancestry estimated from markers and their corresponding genealogical values (Toro *et al*., 2002; Álvarez *et al*., 2007; Royo *et al*., 2007). Their conclusion, confirmed by the results of the present paper, is that the attempt to infer coancestries of a deep genealogy from molecular markers gives results severely biased even if many loci (up to 200) are used. The reason is that the inference (formula [2]) requires information on the true allelic frequencies of markers in the base population that are not known and therefore they are usually substituted by the corresponding values in the actual genotyped population. In our situation, due to the genetic drift accumulated over more than fifty years, the actual frequencies of marker alleles are probably very different to those of the base population. On the other hand, as the simulation results evidence, when the true values of gene frequencies in the founder population are used we recover the correct genealogical values.

The correlation between the genealogical and the molecular inbreeding was negative for the Guadyerbas (–0.32), low for the Torbiscal population (0.19) and substantial for all the animals together (0.69). From a practical point of view the high correlation between molecular and genealogical inbreeding coancestry obtained when the two strains are considered together will allow to predict with accuracy which strain an animal comes from. On the other hand, if the population we are dealing with has been maintained closed with a reasonable amount of mixing, predicting inbreeding even with 49 markers will be poor as it is shown by the results of the much lower correlation within strains. Finally, the previous values are lower than those obtained when genealogical and molecular coancestries are compared indicating, as pointed out by Álvarez *et al*. (2007), that the last must be a better criteria for monitoring populations.

In the simulations run with 49 loci the correlation values are again low for Guadyerbas (0.13) and Torbiscal (0.21) but substantial for all animals (0.70). If the number of loci is increased up to 212 and both populations are considered the correlation approaches to one, but when the populations are considered separately it increases only up to 0.25 (Guadyerbas) and 0.38 (Torbiscal) indicating, again, that the low variability in the genealogical inbreeding values limits the discrimination ability of the markers. Finally, although the standard deviations of correlations are high it is hard to attribute to chance the negative value of –0.32 in the Guadyerbas population. In this strain inbreeding has risen to the considerable value of 35% and inbreeding depression has been detected (Fernández *et al*., 2002). This prevents

homozygosity for genes underlying fitness traits and indirectly homozygosity for neutral markers in linkage disequilibrium with that. That's why molecular inbreeding could be below its expected value under strict neutrality.

In conclusion, it seems preferable to use pedigree information whenever available, and limiting the use of markers to verify, correct, complete or even implement pedigree recording.

Acknowledgements

We thank to the CIA «Dehesón del Encinar» (Oropesa, Toledo) for providing biological samples from pigs and to L. Silió and A. López for their collaboration. This research was funded by INIA RZ03-010 grant.

References

ÁLVAREZ I., ROYO L.J., GUTIÉRREZ J.P., FERNÁNDEZ I., ARRANZ J.J., GOYACHE F., 2008. Relationship between genealogical and microsatellite information characterising losses of genetic variability: empirical evidence from the rare Xalda sheep breed. Livestock Sci 115, 80-88. doi:10.1016/j.livsci.2007.06.009.

CHAKRABORTY R., 1981. The distribution of the number of heterozygous loci in an individual in natural populations. Genetics 98, 461-466.

FERNÁNDEZ A., RODRIGÁÑEZ J., TORO M.A., RODRÍGUEZ M.C., SILIÓ L., 2002. Inbreeding effects on the parameters of the growth function in three strains of Iberian pig. J Anim Sci 80, 2267-2275.

ODRIOZOLA M., 1976. Investigación sobre los datos acumulados en dos piaras experimentales. IRYDA, Madrid. [In Spanish].

PEMBERTON J., 2004. Measuring inbreeding depression in the wild: the old ways are the best. TREE 19, 613-615. doi:10.1016/j.tree.2004.09.010.

ROYO L.J., ÁLVAREZ I., GUTIÉRREZ J.P., FERNÁNDEZ I., GOYACHE F., 2007. Genetic variability in the endangered Asturcón pony assessed using genealogical and molecular information. Livestock Sci 107, 162-169. doi:10.1016/j.livsci.2006.09.010.

TORO M.A., BARRAGÁN C., ÓVILO C., RODRIGÁÑEZ J., RODRÍGUEZ C., SILIÓ L., 2002. Estimation of coancestry in Iberian pigs using molecular markers. Conserv Genet 3, 309-320.

TORO M.A., BARRAGÁN C., ÓVILO C., 2003. Estimation of genetic variability of the founder population in a conservation scheme using microsatellites. Anim Genet 34, 226-228. doi:10.1046/j.1365-2052.2003.00988.x.