ABSTRACT: Genome-wide information was used for detecting inbreeding depression in two reproductive traits (total number of piglets born and number of piglets born alive) in an ancient strain of Iberian pigs (the Guadyerbas strain). A total of 109 sows with phenotypic records were genotyped with the Illumina PorcineSNP60 BeadChip v1. Inbreeding depression was estimated using a bivariate animal model where the inbreeding coefficient was included as a covariate. We used two different measures of genomic inbreeding: inbreeding estimated on a SNP-by-SNP basis and inbreeding estimated from runs of homozygosity. We also performed the analyses using pedigree-based inbreeding. Significant inbreeding depression for both traits was detected using all inbreeding measures. Genome-wide information allowed to narrow down one region significantly associated to inbreeding depression that was located on chromosome 13 and spanned from 27 to 54 Mb. This region overlaps with a previously detected QTL region comprising the inter-alpha-trypsin inhibitor genes that seem to play an important role in embryo implantation. This study underlines the power of genomic inbreeding: inbreeding estimated on a SNP-by-SNP basis for detecting specific genome regions causing inbreeding depression and thus for identifying candidate genes.

Keywords: genomic inbreeding; inbreeding depression; runs of homozygosity; Iberian pig

Materials and Methods

Animals, pedigree and phenotypic data. Data for this study was originated from Guadyerbas animals maintained in a closed herd since 1944. The complete and very accurate genealogy is available since the foundation of the herd and comprises about 25 generations. The effective population size has been estimated to be about 10 animals when using both genealogical and SNP data (Saura et al., 2013). Data from 832 pedigreed sows with TNB and NBA records in successive parities were available. Mean values (and standard deviation) for TNB and NBA were respectively 7.39 (2.34) and 7.06 (2.25).

Genotypic data. Illumina PorcineSNP60 BeadChip v1 genotypes for 113 sows with phenotypic records were available. A larger sample of Iberian pigs (468 animals) was used to analyse the genotyping data in order to increase the power when clustering the genotypes. Quality control procedures (Call frequency > 0.99, GenTrainScore > 0.70, AB R Mean > 0.35, number of inconsistencies with the genealogy < 9) were applied for identifying SNPs and samples performing incorrectly. SNPs unmapped and those mapped on sex chromosomes according to Sscroфа10.2 were also excluded. The final numbers of SNPs and genotyped Guadyerbas females that satisfied the selection criteria were 51,127 and 109, respectively.

Inbreeding coefficients. Different estimates of $F$ were used in the inbreeding depression analysis:
1. Genealogical inbreeding coefficients ($F_{ped}$) obtained using all pedigree information available.
2. Genomic marker-by-marker inbreeding coefficients ($F_{snp}$) obtained based on the excess of SNP homozygosity, in an attempt to correct genomic inbreeding for the homozygosity in the base population (Keller et al. (2011)). Note that $F_{snp}$ can take negative values.
3. Inbreeding coefficients estimated from ROH ($F_{roh}$). For a particular individual $F_{roh(i)}$ was defined as the proportion of the genome that is in ROH (Lencz et al. (2007)). The following criteria were used for defining a ROH: i) a maximum of two missing genotypes and one heterozygous genotype within a particular ROH were permitted; ii) the minimum density was 1 SNP per 100 kb; iii) the maximum distance allowed between two consecutive homozygous SNPs in a run was 1 Mb; and iv) the minimum number of SNPs that constituted a ROH was 30. We also performed analyses based on short and long ROH. We defined $F_{roh_short}$ as the proportion of the autosome in ROH of length

Introduction

The standard approach for estimating inbreeding depression is to regress the phenotype of the trait of interest on the inbreeding coefficient ($F$) computed from pedigree data. However, the current availability of very large numbers of single nucleotide polymorphisms (SNPs) offers new opportunities for obtaining more accurate estimates of $F$ and more detailed approaches for detecting inbreeding depression (Keller et al. (2011)). Several potential advantages of using genomic $F$ rather than pedigree-based $F$ ($F_{ped}$) have been highlighted (Keller et al. (2011); McQuillan et al. (2012)) and they include to i) reflect accurately the actual percentage of the genome that is homozygous; ii) estimate inbreeding and inbreeding depression in specific genomic regions; and iii) incorporate homozygosity arising from very distant common ancestors and distinguish recent from ancestral inbreeding.

The aim of this study was to use genomic data for detecting inbreeding depression on two reproductive traits (total number of piglets born, TNB, and number of piglets born alive, NBA), in a highly inbred strain of Iberian pigs (the Guadyerbas strain).
Inbreeding depression analyses. Inbreeding depression was estimated by regressing the phenotype of each reproductive trait on $F$. This regression was performed by including $F$ as a covariate in a bivariate animal model. The model equation for both traits was:

$$
\mathbf{y} = \mathbf{X}\mathbf{\beta} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_p\mathbf{p} + \mathbf{e},
$$

where $\mathbf{y}$ is the vector of observations, $\mathbf{\beta}$ is the vector of fixed effects, including the comparison of farrowing periods and facilities (eight levels), parity (four levels), strain of boar (two levels: Guadyerbas or Torbiscal) and the (linear) regression on $F$, $\mathbf{a}$ is the vector of additive genetic effects, $\mathbf{p}$ is the vector of permanent environmental effects associated with the sows, $\mathbf{e}$ is the vector of random residual effects, and $\mathbf{X}$, $\mathbf{Z}_a$, and $\mathbf{Z}_p$ are incidence matrices relating the fixed and random effects to the observations. The variances and covariances of the random effects were assumed to be $V(\mathbf{a}) = A_N\sigma^2_a$, $V(\mathbf{p}) = I_m\sigma^2_p$ and $V(\mathbf{e}) = I_m\sigma^2_e$, where $A_N$ is the pedigree-based numerator relationship matrix of order $N$ (number of animals in the pedigree), $I_N$ and $I_m$ are identity matrices of order $m$ (number of sows with litter size records), and $n$ (number of records), respectively, and $\sigma^2_a$, $\sigma^2_p$, and $\sigma^2_e$ are the variances of additive genetic effects, permanent environmental effects, and residual effects, respectively. The analyses were performed using the VCE-PEST software and implementing REML optimization (Groeneveld et al. (1990)).

Different analyses were performed by varying the inbreeding coefficient used in the model ($F_{ped}$, $F_{snp}$, $F_{roh}$, $F_{roh_short}$ and $F_{roh_long}$). The analysis using $F_{ped}$ was carried out using all available performance and pedigree records. There were 823 sows with reproductive records of 2,712 litters, and the total pedigree file contained 1,032 animals. Analyses using $F_{snp}$ and $F_{roh}$ included only records for the 109 genotyped females and used estimates of variance and covariance components obtained from the $F_{ped}$ analysis. Three different analyses were implemented with genomic inbreeding coefficients: i) using average coefficients over the whole genome; ii) using average coefficients for each chromosome; and iii) using average coefficients for specific regions within chromosomes.

Results and Discussion

A significant reduction in both NBA and TNB with an increasing $F_{ped}$ was observed when performing the genealogical analysis using all data available (records from 823 sows). Estimates of inbreeding depression were −0.20 (SE 0.09) for NBA and −0.21 (SE 0.10) for TNB per 10% increase in $F_{ped}$. Although the relationship of molecular $F$ at the whole-genome level (i.e., using $F_{snp}$ and $F_{roh}$) and litter size was not significant, a significant inbreeding depression was found for both traits in chromosome SSC13 when the molecular analyses were carried out at the autosomal level. Figure 1 shows the regression coefficients for the inbreeding depression analyses performed for each autosome using $F_{snp}$ for NBA. Results for TNB (not shown) were very similar to those for NBA. Only SSC13 showed a significant effect ($p < 0.0001$). The regression coefficient for this chromosome was −1.2 for both traits which translates in a decrease of 0.12 NBA and TNB per 10% increase in $F_{snp}$. A similar result was obtained when $F_{roh}$ was used instead (data not shown).

![Figure 1: Regression coefficients and standard errors for inbreeding depression in NBA across autosomes. Values of $-2 \log(p)$ for the analyses are represented below.](image)

The reductions in number of piglets per 10% increase in $F_{snp}$, $F_{roh}$ and $F_{roh_long}$ were all significant and of the same order of magnitude than those derived from the genealogical analysis (Table 1). There was no significant effect of increasing $F_{roh_short}$ for neither NBA nor TNB suggesting that purging of deleterious alleles in ancient generations may have occurred.

In order to detect specific genomic regions causing inbreeding depression, chromosomes were fragmented in different segments of equal size (3, 5 and 8 segments) and three further analyses per chromosome using $F_{snp}$ were carried out for the different segments. When chromosomes were divided into 3 segments, the inbreeding depression analysis was only significant for the first region of SSC13 (0.0 - 73 Mb) for both traits. When they were divided in 5 segments, the analyses for the first (0 - 44 Mb) and second
(44 - 88 Mb) regions were again significant for both traits in SSC13. Finally, when chromosomes were divided into 8 fragments, only the analysis for the second region of this chromosome was significant for both traits. This region has a size of 32.4 Mb and is located between 27 to 55 Mb. Significance was still maintained for NBA even after strict Bonferroni correction.

Table 1. Estimates of inbreeding depression in chromosome 13 for total number of piglets born (TNB) and number of piglets born alive (NBA) expressed as the change in phenotypic mean per 10% increase in $F_s$.  

<table>
<thead>
<tr>
<th>NBA</th>
<th>TNB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (S.E.)</td>
<td>Mean (S.E.)</td>
</tr>
<tr>
<td>$F_{snps}$</td>
<td>$-0.121 (0.047)^*$</td>
</tr>
<tr>
<td>$F_{roh}$</td>
<td>$-0.230 (0.087)^{**}$</td>
</tr>
<tr>
<td>$F_{roh\ short}$</td>
<td>$0.340 (0.380)$</td>
</tr>
<tr>
<td>$F_{roh\ long}$</td>
<td>$-0.181 (0.074)^*$</td>
</tr>
</tbody>
</table>

* $p < 0.05$, ** $p < 0.01$

One of the first genome-wide scans for prolificacy traits was performed by Noguera et al. (2009) who, using data from a Guadyerbas x Meishan F2 intercross and microsatellite markers, detected a QTL region located on SSC13 with effects on both NBA and TNB. This QTL region extended from about 38 to 194 Mb and overlaps with the region identified here. Specifically, the region detected in the inbreeding depression analysis is shorter (it spans from 27 to 54 Mb) and overlaps with the first part of the QTL region detected by Noguera et al. (2009). We examined the gene content of this common region by using the porcine genome annotation Sscrofa10.2 in BioMart tool of Ensembl (ensembl.org/biomart) and the Ensembl Genes 69 database and found 271 genes annotated. Interestingly three of these genes, the inter-alpha-trypsin inhibitor heavy chains 1, 3 and 4 ($ITIH1$, $ITIH3$ and $ITIH4$), mapped at 38 Mb on SSC13, play several important roles in maintaining the uterine surface glycocalyx during placental attachment in pigs (Geisert et al., 2003). Even more, these genes have been previously associated with NBA and TNB (Balcells et al. (2011)). Using the same material as Noguera et al. (2009), Balcells et al. (2011) sequenced the porcine $ITIH1$, $ITIH3$ and $ITIH4$ genes and analyzed endometrial gene expression in order to identify polymorphisms that could explain differences in prolificacy of sows. Their results revealed significant associations with NBA and TNB for two SNPs within $ITIH1$, four SNPs within $ITIH3$, and four SNPs within $ITIH4$. These studies support thus our findings, as genes affecting the studied traits are located in the region identified on SSC13.

Conclusion

These results show the effectiveness of using genome-wide SNP information for identifying specific regions of the genome responsible for inbreeding depression. The specific region identified here for inbreeding depression in litter size of Iberian pigs overlaps with a previously detected QTL regions and contains genes implicated in embryo implantation.

Acknowledgements

This work was funded by the Ministerio de Economía y Competitividad and INIA, Spain (grants CGL2012-39861-C02-02 and RZ2010-00009-00-00). M.S. was supported by a Juan de la Cierva fellowship (CI-2011-10896) from the Ministerio de Economía y Competitividad, Spain. We thank Prof M.A. Toro for useful comments, M.C. Rodríguez and L. Silió for facilitating samples and useful discussions and F. García and R. Benítez for technical support.

Literature Cited


