Livestock Science 174 (2015) 10–17

Contents lists available at ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Genomic selection for boar taint compounds and carcass traits in a commercial pig population

Carolina Filardi de Campos [a], Marcos Soares Lopes [b], Fabyano Fonseca e Silva [a,*], Renata Veroneze [a], Egbert F. Knol [b], Paulo Sávio Lopes [a], Simone E.F. Guimarães [a]

[a] Department of Animal Science, Universidade Federal de Viçosa, Av. P.H. Holfs, 36570-000 Viçosa, Brazil
[b] TOPIGS Research Center IPG, Institute for Pig Genetics B.V., PO Box 43, 6640AA Beuningen, The Netherlands

ARTICLE INFO

Article history:
Received 5 August 2013
Received in revised form
30 October 2014
Accepted 21 January 2015

Keywords:
Genomic estimated breeding values
Androstenone
Skatole
Backfat

ABSTRACT

This study aimed to compare two different Genome-Wide Selection (GWS) methods (Ridge Regression BLUP – RR-BLUP and Bayesian LASSO – BL) to predict the genomic estimated breeding values (GEBV) of four phenotypes, including two boar taint compounds, i.e., the concentrations of androstenone (andro) and skatole (ska), and two carcass traits, i.e., backfat thickness (fat) and loin depth (loin), which were measured in a commercial male pig line. Six hundred twenty-two boars were genotyped for 2,500 previously selected single nucleotide polymorphisms (SNPs). The accuracies of the GEBV using both methods were estimated based on Jack-knife cross-validation. The BL showed the best performance for the andro, ska and loin traits, which had accuracy values of 0.65, 0.58 and 0.33, respectively; for the fat trait, the RR-BLUP accuracy of 0.61 outperformed the BL accuracy of 0.56. Considering that BL was more accurate for the majority of the traits, this method is the most favoured for GWS under the conditions of this study. The most relevant SNPs for each trait were located in the chromosome regions that were previously indicated as QTL regions in other studies, i.e., SSC6 for andro and ska, SSC2 for fat, and SSC11, SSC15 and SSC17 for loin.

1. Introduction

Most progress that has been made in pig breeding programs regarding quantitative traits has been a result of selection based on the estimation of genetic breeding values using pedigree information. However, with the development of molecular markers, such as single nucleotide polymorphisms (SNPs), new approaches, such as genome-wide selection (GWS) and genome-wide association studies (GWAS), have been proposed (Hayes and Goddard, 2010). In the pig, these approaches remain under development. The high-density Porcine SNP60 Genotyping BeadChip (Illumina Inc., San Diego, CA, USA, Ramos et al., 2009) was proposed using next-generation sequencing technologies for the mass identification of SNPs in regions of the genome that have not been previously sequenced, and this technology is currently widely used in the pig breeding industry.

With respect to phenotypes that have been used in GWAS studies, the phenotypes that are related to boar taint and carcass traits stand out because they are considered specialised phenotypes. Boar taint is the undesirable smell and taste of pork derived from uncastrated males, and its main associated compounds are androstenone and skatole (Gregersen et al., 2012). Duijvesteijn et al. (2010) attempted to determine the SNPs associated with androstenone levels in fat tissue, Ramos et al. (2011) reported an association

* Corresponding author.
E-mail address: fabyanofonseca@ufv.br (F.F. e Silva).

http://dx.doi.org/10.1016/j.livsci.2015.01.018
1871-1413/

study that aimed to identify the SNPs related to skatole levels in the pig carcass, and Rowe et al. (2014) presented an association study for both androstenone and skatol concentrations in Danish Landrace boars. In relation to carcass traits, Luo et al. (2012) conducted a GWAS study for meat quality; the results effectively narrowed down the associated regions compared with previous QTL studies and revealed haplotypes and candidate genes of SSC12 in pigs. Although GWAS studies have been conducted on boar taint and carcass traits in pigs, there are no references to GWS studies that aimed to estimate the genomic breeding values for these traits in commercial pig lines.

Since the initial paper by Meuwissen et al. (2001) was published, several studies have compared the efficiency of the simplest GWS method, the Ridge Regression BLUP (RR-BLUP) (Meuwissen et al., 2001), with more sophisticated methods, such as Bayesian LASSO (BL) (de los Campos et al., 2009). Because of the scarcity of GWS studies of boar taint and carcass traits in pigs, it is worthwhile to compare these methods to best predict the breeding values for these specialised phenotypes. In summary, the main difference between these two very popular GWS methods is that the RR-BLUP assumes, a priori, that each locus explains an equal amount of the genetic variation, whereas the BL assumes that each locus explains a unique amount of variation.

The GWS methods are typically compared using cross-validation techniques, which are useful when evaluating the predictive ability of genomic breeding values. However, because of the varying degrees of relationships in animal breeding applications, it is difficult to obtain independent training and testing sets. Therefore, the training–testing partitions have a significant effect on the cross-validation results (Pérez-Cabal et al., 2012). In this context, although the Jack-knife (leave-one-out) partition is computationally intensive, it maximises the training population size (Resende Jr et al., 2012), thereby representing the best option for use in cross-validation analyses.

Considering that genomic selection for traits such as androstenone and skatole concentrations, backfat thickness and loin depth have not been published for commercial pig lines to date, the main objective of this study was to compare the RR-BLUP and BL methods in relation to their efficiencies in predicting genomic breeding values using the Jack-knife method for optimal cross-validation analysis. We also aimed estimate heritabilities and genetic correlations, besides to identify the most relevant SNPs for each trait to associate the chromosomal region of these markers with previously reported QTLs for these phenotypes.

2. Materials and methods

2.1. Phenotypic data

The field experiment was conducted strictly in line with Dutch law regarding the protection of animals. All boars were animals from a composite Duroc-based line; the animals were related, raised under the same conditions, and obtained from a traditional selection program. Six hundred twenty-two boars from a farm in the Netherlands were phenotyped for the following traits: concentrations of androstenone (andro) and skatole (ska), backfat thickness (fat) and loin depth (loin). The average and standard deviations for the andro, ska, fat and loin phenotypes were 0.2 (0.82) μg/g, 4.08 (0.77) ng/g, 14.33 (2.93) mm, and 61.74 (6.88) mm, respectively.

For the measurements of the backfat thickness and loin depth, a Hennessy Grading Probe (HGP) was used. The back of the carcass was penetrated with a needle to identify the tissue interfaces, and the phenotypic measurements were produced according to the site (http://www.hennessy-technology.com/grading.html). Samples were collected from the neck fat of the animal carcass's left side and were stored under vacuum at −20 °C until phenotypic analysis, when the concentrations of androstenone and skatole were measured. Additional information regarding the collection and phenotype processing can be found in Duijvesteijn et al. (2010).

The phenotypic values for the concentrations of androstenone and skatole were not normally distributed and were, therefore, subjected to a logarithmic transformation as previously described by Duijvesteijn et al. (2010) and Ramos et al. (2011). After the transformation, the Shapiro−Wilk test for normality it was applied to validate the efficiency of log-transformation. The *p*-values for androstenone and skatole were equal to 0.098 and 0.136, respectively. Since the alternative hypothesis is given by absence of normality, the *p*-values imply that these traits follow a normal distribution at 5% level of significance.

2.2. Genotypic data

The animals were genotyped using the Illumina PorcineSNP60 BeadChip (San Diego, CA, USA, Ramos et al., 2009). The DNA was prepared from ethylenediaminetetraacetic acid (EDTA) blood, hair roots or meat samples using the Gentra Puregene DNA Preparation Kit (Minneapolis, MN) according to the manufacturer's instructions. The extraction was performed using a standard phenol-chloroform method as previously described (Sambrook and Russell, 2006). The DNA concentration and purity (absorbance ratios of 260/280 and 260/230, respectively) were measured using the Nanodrop ND-1000 Spectrophotometer (NanoDrop Technologies, LLC, Wilmington, Delaware). Following a quality check, 10,210 SNPs were removed because of low quality scores (GenCall score <0.7). A threshold of 30 or more pedigree errors was applied, and 190 SNPs were removed. In addition, 20,736 SNPs were excluded from the analyses because of a minor allele frequency (MAF) <0.05 in at least one of the three lines. An additional 374 markers with a call rate of $<95\%$ were also excluded. A total of 3,982 SNPs that were located in one of the sex chromosomes were also excluded. Additional details regarding the DNA preparation and genotyping process can be found in Duijvesteijn et al. (2010).

The set of 2,500 SNPs that were used in this study comprised a subset that was previously identified by Lopes et al. (2013) using the same dataset. These authors tested six subsets with different numbers of markers ($n=500$, 1,000, 1,500, 2,000, 2,500 and 3,000 SNPs) and concluded that the subset of 2,500 SNPs represented an optimal number for estimating genomic relatedness because these markers showed the same results that were obtained using 47,897 SNPs. The 2,500 selected SNPs were distributed throughout the genome with an average of 131 SNPs per chromosome