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

and an average distance of 1,038 kb between markers. These markers can be seen as a special set of SNPs used to correct pedigree information in routine genetic evaluations of a pig breeding company (Lopes et al., 2013). Although reduced, this panel is economically feasible for large-scale genotyping, and can be better exploited when used to improve genetic predictions for certain complex traits that are difficult to be measured or are sex-limited, like carcass and boar taint traits, respectively.

2.3. GWS methods and estimation of the genetic parameters

With respect to the GWS, the phenotypic outcomes, which are denoted by y_i ($i=1, 2,\ldots$, and 622), were regressed on the marker covariates x_{ik} ($k=1, 2, \ldots$, and 2,500) following the regression model that was previously proposed by Meuwissen et al. (2001):

$$y_i = \mu + \sum_{k=1}^{2,500} x_{ik}\beta_k + e_i \quad (1)$$

where y_i is the phenotypic observation of animal i, μ is the general mean, β_k is the effect of marker k, and e_i is the residual term $e_i \sim N(0, \sigma_e^2)$. In this model, x_{ik} equals a value of 2, 1 or 0 that represents the SNP genotypes AA, Aa and aa, respectively. Thus, these values directly represent the allele copy number at each locus k. Under a matrix notation, the presented GWS model can be rewritten as follows:

$$\mathbf{y} = \mathbf{1}'\mu + I\sum_{k=1}^{2,500}\mathbf{X}_k\beta_k + \mathbf{e}, \quad (2)$$

where $\mathbf{1}'$ and $\boldsymbol{I}$ respectively represent a unit vector and an identity matrix with dimensions of 622; $\mathbf{y} = [y_1, y_2, \ldots, y_{622}]'_{622x1}$, $\mathbf{X}_k = [x_{1k}, x_{2k}, \ldots, x_{622k}]'_{622x1}$ and $\mathbf{e} = [e_1, e_2, \ldots, e_{622}]'_{622x1}$.

In the RR-BLUP (Meuwissen et al., 2001) method, β_k is the random marker effect $\beta_k \sim N(0, \sigma^2_{\beta_k})$, which assumes that $\sigma^2_{\beta_1} = \sigma^2_{\beta_2} = \ldots = \sigma^2_{\beta_{2,500}} = \sigma^2_{\beta}$ (i.e., each locus explains an equal amount of genetic variation). This method was implemented using the R software (R Development Core Team, 2011) package *rrBLUP* (Endelman, 2011), in which the additive genetic variance is given by $\sigma^2_a = 2\sigma^2_{\beta}\sum_{k=1}^{2,500} p_k(1-p_k)$.

The BL (de los Campos et al. 2009) method is a penalised Bayesian regression procedure whose general estimator is given by $\hat{\boldsymbol{\beta}} = \arg\min_{\beta}\{(\hat{\mathbf{y}} - X\boldsymbol{\beta})'(\hat{\mathbf{y}} - X\boldsymbol{\beta}) + \lambda\sum_{k=1}^{2,500}|\beta_k|\}$,where λ is the regularisation parameter. The BL method was implemented in the package *BLR* (de los Campos et al., 2009; Pérez et al., 2010) of the R software using 10,000 MCMC iterations, with the burn-in and thin with 4000 and 2 iterations, respectively. In this approach, the additive genetic variance is $\sigma^2_a = 2\sum_{k=1}^{2,500}\sigma^2_{\beta_k}p_k(1-p_k)$ because each locus presents a particular variance ($\sigma^2_{\beta_k}$).

For both the RR-BLUP and BL methods, the vector of genomic estimated breeding values (GEBV) were obtained as $\hat{\mathbf{u}} = \sum_{k=1}^{2,500}\mathbf{X}_k\hat{\beta}_k = X\hat{\boldsymbol{\beta}}$, and the heritability was defined as $h^2 = \sigma^2_{\mathbf{a}}/(\sigma^2_{\mathbf{a}} + \sigma^2_{\mathbf{e}})$. Furthermore, the hot carcass weight was used as a linear covariate in the analysis of the backfat thickness and loin depth. For the concentrations of androstenone and skatole, the hot carcass weight and age were used as linear covariates. The contemporary groups (month and year of slaughter) were used as qualitative fixed effects for all traits. The genetic correlations across the four traits were computed using the Pearson's correlation of the GEBVs in the most accurate method (RR-BLUP or BL). The effects of the markers were distributed throughout the chromosomes for each trait, and Manhattan plots were built using the package *ggplot2* in the R software (R Development Core Team, 2011).

2.4. Cross-validation by Jack-knife

To compare the RR-BLUP and BL methods, Jack-knife (leave-one-out) cross-validation was used. Without a loss of generality for each method and each phenotype, the original dataset with 622 animals was divided into 622 training datasets ($\mathbf{D}_{-i}$) of 621 individuals, $\mathbf{D}_{-1}$, $\mathbf{D}_{-2}$, …, $\mathbf{D}_{-622}$, and each dataset contained the marker and phenotype information for all animals except for animal $-i$. In these analyses, the predicted genomic breeding value of animal i for each trait was calculated by $\hat{u}*_i = \mathbf{X}_i\hat{\boldsymbol{\beta}}_{-i}$, where $\mathbf{X}_i$ denotes the SNP genotype vector of animal i, and $\hat{\boldsymbol{\beta}}_{-i}$ denotes the estimated marker effects vector from the analysis that considered all animals except for animal i. All codes that are related to the leave-one-out cross-validation implemented for the RR-BLUP and BL methods are available as Supplementary material.

The vector that contained all predicted values was denoted by $\hat{u}* = [\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_{622}]$, and the accuracy ($r$) that was used to measure the efficiency of RR-BLUP and BL was given by $r = r_{y\hat{u}*}/\sqrt{h^2}$, where $r_{y\hat{u}*}$ is the correlation between the observed phenotype ($\boldsymbol{y}$) and $\hat{\mathbf{u}}*$, and h^2 is the estimated heritability (Resende Jr et al., 2012).

3. Results

The accuracy (Fig. 1) was used to support the choice of the best method for genomic selection. For the traits concentrations of androstenone (andro) and skatole (ska), backfat thickness (fat) and loin depth (loin), the accuracy values with BL were 0.65, 0.58, 0.56 and 0.33, respectively; in contrast, the RR-BLUP accuracy values were 0.63, 0.57, 0.61 and 0.26, respectively. Thus, for three traits (andro, ska and loin), the BL reached a higher accuracy compared with the RR-BLUP method; in contrast, for fat, the RR-BLUP reached a higher accuracy. With respect to the correlations (Fig. 1), it is possible to note that the methods showed smaller than accurate values because there is no square root of the heritability in the denominator, but the performance of the methods in the comparative study was the same as that observed for the accuracy analysis. In Fig. 1, the 95% confidence intervals (using Student's t distribution) were also plotted for the correlation coefficients that were calculated for each method in each trait. As the accuracy was a function of the correlation coefficients, there was no theoretical justification to use the same Student's t distribution to obtain intervals for this quantity.

Table 1 shows the heritability estimates and genetic correlations (Pearson's correlation between GEBVs), as well as the correlations between the marker effect estimates; the

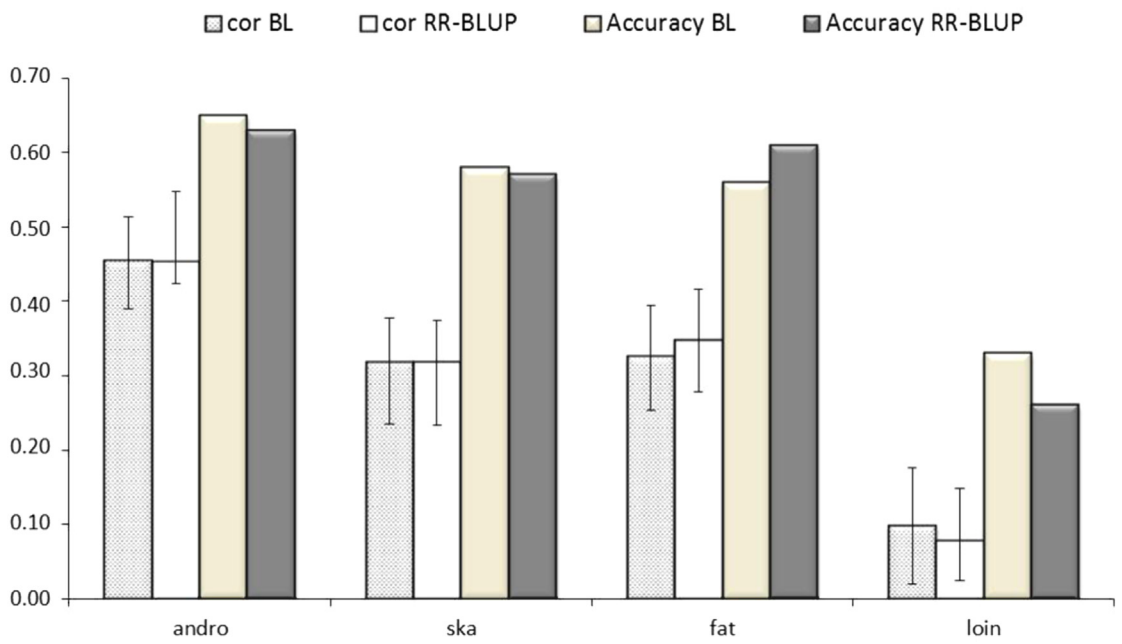

Fig. 1. Correlation between observed phenotype and predicted values and accuracies of the methods BL and RR-BLUP. BL and RR-BLUP indicate estimates from Bayesian LASSO and Ridge Regression BLUP, respectively.

Table 1
Heritabilities estimates on the diagonal, genetic correlations on the upper (right triangle) and correlations between estimated marker effects on the lower (left triangle).

	Phenotypes			
	Andro (BL)	Ska (BL)	Fat (RR-BLUP)	Loin (BL)
Phenotypes				
andro	0.46	0.24*	−0.01	−0.05
aka	0.36*	0.26	0.05	−0.14*
fat	0.10*	0.03	0.32	0.07
loin	−0.09*	−0.07*	0.01	0.10

Andro: concentration of androstenone, ska: concentration of skatole, fat: backfat thickness, loin: loin depth. BL and RR-BLUP indicate estimates from Bayesian LASSO and Ridge Regression BLUP, respectively.

* $P < 0.01$.

BL method was used for three traits (andro, ska and loin), and the RR-BLUP was used for fat.

The effects of the markers were distributed throughout the chromosomes for the traits. The Manhattan plots are presented in Fig. 2 for the boar taint and carcass traits; the BL method was used for three traits (andro, ska and loin), and RR-BLUP was used for fat. An empirical threshold (dotted line) was assumed in order to facilitate the visualization of the most relevant SNPs for all traits, since in GWS approach there are no *p*-values to indicate significance of SNPs like in GWAS.

Considering the distribution of the effects throughout the chromosomes, the markers that showed the highest peaks represent the most important markers controlling the traits. Thus, for the concentration of androstenone, the highest peaks were identified on chromosomes SSC1 (SNP ALGA0106999 at position 309,125,803 bp), SSC6 (SNP ASGA0030401 at position 56,833,928 bp), SSC13 (SNP MARC0069512 at position 123,984,079 bp) and SSC16 (SNP ASGA0074756 at position 82,981,442 bp). The Pig QTL database (http://www.animalgenome.org/cgi-bin/QTLdb/SS/index) includes 31 (SSC1), 5 (SSC6), 1 (SSC13) and 1 (SSC15) QTLs for this trait. Although some chromosomes, such as SSC9, SSC12, SSC13, SSC14 and SSC15 do not have the greatest effects, they may show regions of QTLs for this trait.

Regarding the concentration of skatole, the highest peaks were observed on chromosomes SSC6 (SNP ASGA0082907 at position 3,503,039 bp), SSC7 (SNP M1GA0009471 at position 5,307,401 bp), SSC11 (SNP H3GA0032391 at position 76,377,968 bp) and SSC12 (SNP MARC0048672 at position 17,733,364 bp).The Pig QTL database includes 12 QTLs that are associated with this trait on chromosomes SSC4, SSC6, SSC7, SSC13 and SSC14. Although such has not got the greatest effect, SSC 13 and SSC 16 may show regions of QTLs for this trait, but there are no reported QTLs

With respect to the carcass traits, the highest peaks for backfat thickness were identified on SSC1 (SNP ASGA0005243 at position 186,106,230 bp), SSC2 (SNP ASGA0009363 at position 18,229,708 bp), SSC4 (SNP MARC0095077 at position 105,126,097 bp), SSC12 (SNP ALGA0108041 at position 23,531,215 bp) and SSC14 (SNP ALGA0080832 at position 11,711,232 bp). The Pig QTL database includes 81, 100, 83, 7 and 1 QTLs, respectively, that are associated with this trait on these chromosomes. The chromosomes SSC6 and SSC7 may show regions of QTLs for this trait; however, these chromosomes do not have the greatest effects. For the loin depth trait, the highest peaks were identified on SSC9 (SNP ALGA0051570 at position 14,512,022 bp), SSC11 (SNP ALGA0061430 at position 23,790,063 bp), SSC15 (SNP H3GA0044020 at position

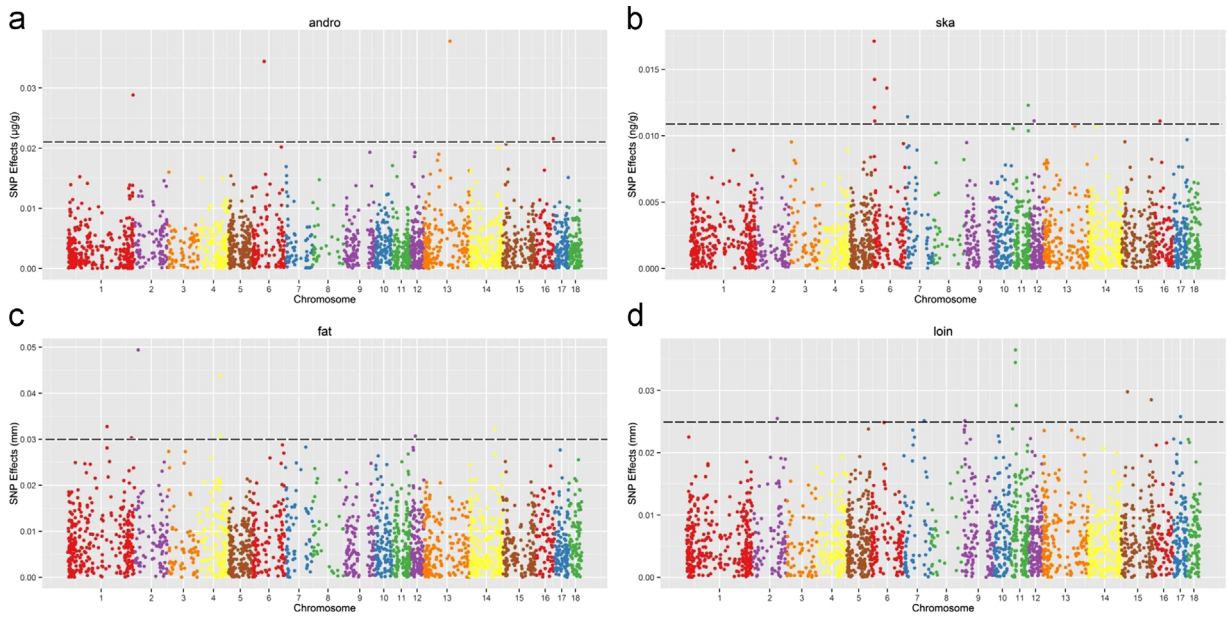

Fig. 2. Distribution and plot of markers absolute effects for boar taint and carcass traits. (a) Concentration of androstenone, (b) concentration of skatole, (c) backfat thickness (RR-BLUP method) and (d) loin depth. Effects for (a), (b) and (d) were achieved by Bayesian LASSO method. Effects for (c) were achieved by RR-BLUP method.

30,870,509 bp) and SSC17 (SNP H3GA0048589 at position 37,207,235 bp). The Pig QTL database includes 13, 2, 9 and 3 QTLs, respectively, that are associated with this trait on these chromosomes. The chromosomes SSC2, SSC6 and SSC7 do not have the greatest effects, but may show regions of QTLs for this trait.

4. Discussion

Although the overlapping of the confidence intervals for the correlation coefficients from BL and RR-BLUP indicated no significant differences between these methods (Fig. 1) for the evaluated traits, the BL method showed higher accuracy values compared with the RR-BLUP method for three traits (andro, ska and loin); these findings indicate that for these traits, the BL property of assuming a priori the different variance for different markers ensures a higher genomic selection accuracy. These results are consistent with Ogutu et al. (2012) who used a simulated dataset to demonstrate that LASSO-type regressions were more efficient compared with RR-BLUP for genomic selection because these regressions provided more-accurate and less-biased predictions. Usai et al. (2009) also supported this result by stating that Bayesian methods are better in the presence of markers with larger effects because they indicate non-linear predictions, thereby ensuring the description of the genetic architecture of the traits. However, the RR-BLUP was superior for fat, assuming that this trait has a polygenic effect; in general terms, these findings indicated that all markers contribute equally to genetic variation as previously described by Hayes et al. (2009).

Regarding other studies that compared traditional BLUP and Bayesian methods, Legarra et al. (2011) reported that BL is an attractive candidate for genomic selection, and that its exponential distribution reasonably reflects the nature of the QTL effects. This information could be useful in identifying regions of the genome that encode the traits in question. Legarra et al. (2011) compared BL with other LASSO and BLUP regressions and concluded that BL is appropriate for genomic selection, with generally the highest accuracies and less inflation of GEBVs compared with other methods. Guo et al. (2012) stated that the RR-BLUP method is a good and simple method for estimating marker effects in GWS and provided prediction accuracies that are comparable with Bayes A and Bayes B when assuming polygenic effects as a result of its advantages, such as being computationally easier.

In summary, the traits considered in the present study characterised as the traits in which the GWS is useful to increase the accuracy of selection and shorten the generation intervals (Daetwyler et al., 2012) because traits such as carcass composition and meat quality (fat and loin) are difficult to measure in live animals, and traits such as andro and ska are sex limited. Other point that deserves be highlighted is the cross-validation strategy based on Jackknife method proposed in the current study, which has not yet been used to compare GWS methods. We believe this method is the best one in terms of choosing training and validation populations, because it maximizes the number of animals in the training phase ensuring the ideal scenario to methods comparison.

The heritability estimate for the backfat thickness in this study is consistent with estimates that have been identified in the literature. van Wijk et al. (2005), who studied a very similar population compared with this study and used traditional restricted maximum likelihood (REML), observed a value of 0.45, while Akanno et al. (2013) identified a value of 0.44 for the estimate of heritability for backfat thickness on carcasses; however, different values have been observed depending on the structure of the population and environmental conditions. For the loin depth, the heritability value

in this study is consistent with the value of 0.13 of van Wijk et al. (2005) and Edwards et al. (2006), who identified heritabilities that ranged from 0.05 to 0.73 for the loin muscle area.

For the boar taint traits, the heritability of androstenone concentration in this study is consistent with the estimates that ranged from 0.25 to 0.88 according to Sellier et al. (2000) and 0.54 reported by Windig et al. (2012). Lower heritabilities of 0.19, 0.41 and 0.55 have been reported for the skatole concentration by Pederson (1998), Windig et al. (2012) and Tajet et al. (2006), respectively. The value in this study is consistent with these findings. It is worth mentioning that all previously reported studies used only phenotypic and pedigree data for the estimations. The genomic selection of the boar taint traits becomes a superior option to male piglet castration because this latter practice has been banned in some countries as a result of welfare concerns and a reduction in the feed conversion efficiency and carcass trait values (Claus et al., 1994). Moreover, in agreement with Duijvesteijn et al. (2010), in the near future, uncastrated males will no longer exist because of animal welfare concerns; thus, high values for boar taint traits must be prevented, and genetic selection is one of the most important tools to execute this prevention.

Genetic correlations were computed across the GEBVs of the four traits. In this study, the genetic correlation between the backfat thickness and loin depth was zero. However, Tomiyama et al. (2009) reported genetic correlation estimates of −0.40 and −0.23 between the loin eye area at 60 days of age and the backfat thickness at 60 days of age, respectively. Between the loin eye area at finish and the backfat thickness at 60 days of age in Berkshire pigs, van Wijk et al. (2005) described a genetic correlation between the backfat thickness and loin weight equal to −0.60 in pigs; these values are not consistent with the present study potentially because heritabilities and genetic correlations may vary according to the genetic constituents and breeding structure of a population (Tomiyama et al., 2009). The results between the concentrations of androstenone and skatole showed a genetic correlation of 0.24. Genetic correlations of 0.36 (Tajet et al., 2006) and 0.37 (Windig et al., 2012) have been reported for different Landrace populations.

As shown in Table 1, the correlations that were calculated between the vectors of the SNP effects were useful in demonstrating the relevance of the markers for each trait. Therefore, it can be assumed that the markers that explain the boar taint traits (i.e., concentrations of androstenone and skatole) had a positive correlation between their vectors of effects (0.36), indicating that some marker loci have influence on both traits simultaneously. This result is interesting and can be exploited under a biological approach in order to identify metabolic functions behind these markers.

Regarding the androstenone concentration, Duijvesteijn et al. (2010) described 37 SNPs that affected the androstenone levels that were located on pig chromosomes SSC1 and SSC6 in the same population. In the latter, a larger region of 33 to 44.9 Mb was associated or potentially involved with androgens. In this larger region, the authors described candidate genes, such as hydroxyl steroid sulfotransferase A1, hydroxy steroid sulfotransferase B1 and several cytochrome P450 genes. These genes are involved in androgen and oestrogen synthesis. In this study (Fig. 2), the SNP with a larger effect in SSC6 is located close to the region that was described by Duijvesteijn et al. (2010). In relation to SSC1, our larger SNP effect is also within the region that was described by Duijvesteijn et al. (2010) in this same population. The correspondence between both studies validates the regions in SSC1 and SSC6 as important for the boar taint traits. To reinforce this importance, Szyda et al. (2003) described a QTL in SSC6 for the smell intensity of the meat in a F_1 and F_2 Duroc × Norwegian Landrace population. In general terms, QTL detection studies are valid for the same population, although dense SNP panels ensure good precision when extrapolated for other related populations (Toosi et al., 2010). In the case of the present study, we believe that comparison with results from Duijvesteijn et al. (2010) and Szyda et al. (2003) make sense because both populations are composite Duroc-based. Thus it is recommended comparisons involving the results obtained here in pig populations originated from Duroc crossings. There are no literature reports of QTL or significant SNP effects in SSC13 or SSC15 for this trait in Duroc-derived pig population. However, Rowe et al. (2014) reported significant QTL for androstenone in SSC13 in a Danish Landrace boar population.

Ramos et al. (2011) reported a genome-wide association study that revealed 16 SNPs located in the proximal region of SSC6 that were significantly associated with skatole levels similar to a Duroc based-line. One hundred forty-three SNPs were located in the region that encompasses the initial 6 Mb of SSC6. However, only 16 markers displayed significant associations with skatole in the analysed pig population and showed three separate regions. The first region contained two SNPs that were located at 0.63 and 0.65 Mb and seven SNPs that were located in the second region, while the third region contained five significant SNPs. The SNPs with highest effects in this study were close to the SNPs of the third region (approximately 3.3 – 4 Mb). For chromosome SSC7, Grindflek et al. (2001) reported a QTL that was associated with smell intensity and was located at 40 cM near the SNP that was identified in this study (Fig. 2), which was located at the beginning of SSC7. For SSC12, there are no reports for the concentration of skatole.

For the carcass traits, including backfat thickness and loin depth, there is a large number of QTLs that are distributed throughout the genome described in the Pig QTL database. For the backfat thickness, Sanchez et al. (2006) described a QTL on SSC1 located 143 cM; two SNPs with higher effects in this study were located close to this region. For SSC2, de Koning et al. (2001) determined that a QTL that was located at 5 cM and the SNP in this study were within a confidence interval. Sanchez et al. (2014) reported a QTL not described previously at position 139 Mb on SSC14 for backfat thickness in a commercial population of Large White pigs. For the loin depth, van Wijk et al. (2006) identified a QTL for loin weight that was located on SSC11 at 8.8 cM, and the SNP was very close to the QTL peak location.

The presence of these QTLs reported in the Pig QTL database and in other studies is useful when validating the regions that have markers with major effects (i.e., absolute values of the estimates) on the traits. The magnitude of the

marker effect estimates throughout the chromosomes provides information for future studies on the candidate genes and supports the implementation of genomic selection in pigs.

Even if the differences between RR-BLUP and BL were not too high in this study, the accuracy values from Jackknife cross-validation indicated the BL method as the best choice to evaluate traits including the concentrations of androstenone and skatole, and loin depth under a GWS scenario. However, because RR-BLUP performed better for backfat thickness, we cannot state that Bayesian methods are always the best option.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

We acknowledge financial support from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior)/NUFFIC (Netherlands Organization for International Cooperation in Higher Education) and FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.livsci.2015.01.018.

References

Akanno, E.C., Schenkel, F.S., Quinton, V.M., Friendship, R.M., Robinson, J.A.B., 2013. Meta-analysis of genetic parameter estimates for reproduction, growth and carcass traits of pigs in the tropics. Livest. Sci. 152, 101–113.

Claus, R., Weiler, U., Herzoq, A., 1994. Physiological aspects of androstenone and skatole formation in the boar—a review with experimental data. Meat Sci. 38 (2), 289–305.

Daetwyler, H.D., Swan, A.A., van der Werf, J.H.J., Hayes, B.J., 2012. Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation. Genet. Sel. Evol. 44, 33.

de Koning, D.J., Rattink, A.P., Harlizius, B., Groenen, M.A.M., Brascamp, E.W., van Aredonk, J.A.M., 2001. Detection and characterization of quantitative trait loci for growth and reproduction traits in pigs. Livest. Prod. Sci. 72, 185–198.

de los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., Weigel, K., Cotes, J.M., 2009. Predicting quantitative traits with regression models for dense molecular markers. Genetics 182, 375–385.

Hayes, B.J., Goddard, M.E., 2010. Genome-wide association and genomic selection in animal breeding. Genome 53 (11), 876–883.

Duijvesteijn, N., Knol, E., Merks, J., Crooijmans, R., Groenen, M., Bovenhuis, H., Harlizius, B., 2010. A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6. BMC Genet. 20, 11–42.

Edwards, D.B., Tempelman, R.J., Bates, R.O., 2006. Evaluation of Duroc- vs. Pietrain-sired pigs for growth and composition. J. Anim. Sci. 84 (2), 266–275.

Endelman, J.B., 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome 4, 250–255.

Gregersen, V.R., Conley, L.N, Sorensen, K.K, Guldbrandtsen, B., Velander, I. H., Bendixen, C., 2012. Genome-wide association scan and phased haplotype construction for quantitative trait loci affecting boar taint in three pig breeds. BMC Genomics 13, 22.

Grindflek, E., Szyda, J., Liu, Z., Lien, S., 2001. Detection of quantitative trait loci for meat quality in a commercial slaughter pig cross. Mamm. Genome 12, 299–304.

Guo, Z., Tucker, D.M., Lu, J., Kishore, V., Gay, G., 2012. Evaluation of genome-wide selection efficiency in maize nested association mapping populations. Theor. Appl. Genet. 124, 261–275.

Hayes, B.J., Visscher, P.M., Goddard, M.E., 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genet. Res. 91, 47–60.

Hennesy Grading Probe. ⟨http://www.hennessytechnology.com/porkgrading.html⟩ (accessed 20.08.2011).

Legarra, A., Robert-Granié, C., Croiseau, P., Guillaume, F., Fritz, S., 2011. Improved Lasso for genomic selection. Genet. Res. 93, 77–87.

Lopes, M.S., Silva, F.F., Harlizius, B., Duijvesteijn, N., Lopes, P.S., Guimarães, S.E.F., Knol, E.F., 2013. Improved estimation of inbreeding and kinship in pigs using optimized SNP panels. BMC Genet. 14, 92.

Luo, W., Cheng, D., Chen, S., Wang, L., Li, Y., Ma, X., Song, X., Liu, X., Li, W., Liang, J., Yan, H., Zhao, K., Wang, C., Wang, L., Zhang, L., 2012. Genome-wide association analysis of meat quality traits in a porcine large white × minzhu intercross population. Int. J. Biol. Sci. 8 (4), 580–595.

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829.

Ogutu, J.O., Schulz-Streeck, T., Piepho, H.P., 2012. Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. BMC Proc. 6, S10.

Pederson, B., 1998. Heritability of skatole in back fat. In: Jensen, W. (Ed.), Skatole and Boar Taint, Danish Meat Research Institute, Roskilde, pp. 129–136.

Pérez, P., de los Campos, G., Crossa, J., Gianola, D., 2010. Genomic-enabled prediction based on molecular markers and pedigree using the Bayesian linear regression package in R. Plant Genome 3 (2), 106–116.

Pérez-Cabal, M.A., Vazquez, A.I., Gianola, D., Rosa, G.J.M., Weigel, K.A., 2012. Accuracy of genome-enabled prediction in a dairy cattle population using different cross-validation layouts. Front. Genet. 3 (27), 1–7.

PIGQTL database. ⟨http://www.animalgenome.org/cgi-bin/QTLdb/SS/index⟩ (accessed 12.03.2013).

R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing3-900051-07-0 ⟨http://www.R-project.org⟩ (URL).

Ramos, M.A., Crooijmans, R.P.M.A., Affara, N.A., Amaral, A.J., Archibald, A. L., Beever, J.E., Bendixen, C., Churcher, C., Clark, R., Dehais, P., Hansen, M.S., Hedegaard, J., Hu, Z.-L., Kerstens, H.H., Law, A.S., Megens, H.J., Milan, D., Nonneman, D.J., Rohrer, G.A., Rothschild, M.F., Smith, T.P.L., Schnabel, R.D., Van Tassell, C.P., Taylor, J.F., Wiedmann, R.T., Schook, L. B., Groenen, M.A.M., 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PLoS One 4, e6524.

Ramos, M.A., Duijvesteijn, N., Knol, E.F., Merks, J.W.M, Bovenhuis, H., Crooijmans, R.P.M.A., Groenen, M.A.M., Harlizius, B., 2011. The distal end of porcine chromosome 6p is involved in the regulation of skatole levels in boars. BMC Genet. 12, 35.

Resende Jr., M.F.R., Muñoz, P., Resende, M.D.V., Garrick, D.J., Fernando, R.L., Davis, J.M., Jokela, E.J., Martin, A., Peter, G.F., Kirst, M., 2012. Accuracy of genomic selection methods in a standard data set of Loblolly Pine (*Pinustaeda L.*). Genetics 190, 1503–1510.

Rowe, S.J., Karacaören, B., Dirk-Jan de Koning, D.-J., Lukic, B., Hastings-Clark, N., Velander, I., Haley, C.S., Archibald, A.L., 2014. Analysis of the genetics of boar taint reveals both single SNPs and regional effects. BMC Genomics 15, 424.

Sambrook, J., Russell, D.W., 2006. Purification of nucleic acids by extraction with phenol:chloroform. Cold Spring Harb. Protoc. http://dx.doi.org/10.1101/pdb.prot4455.

Sanchez, M.P., Riquet, J., Iannuccelli, N., Gogué, J., Billon, Y., Demeure, O., Caritez, J.C., Burgaud, G., Fève, K., Bonnet, M., Péry, C., Lagant, H., Le Roy, P., Bidanel, J.P., Milan, D., 2006. Effects of quantitative trait loci on chromossomes 1,2,4, and 7 on growth, carcass, and meat quality traits in backcross Meishan × Large White pigs. J. Anim. Sci. 84, 526–537.

Sanchez, M.P., Tribout, T., Iannuccelli, N., Bouffaud, M., Servin, B., Thenge, A., Dehais, P., Muller, N., Schneider, M.P.D., Mercat, M.J., Rogel-Gaillard, C., Milan, D., Bidanel, J.P., Gilbert, H., 2014. A genome-wide association study of production traits in a commercial population of Large White pigs: evidence of haplotypes affecting meat quality. Genet. Sel. Evol. 46, 12.

Sellier, P., Le Roy, P., Fouilloux, M., Gruand, J., Bonneau, M., 2000. Responses torestricted index selection and genetic parameters for

fat androstenonelevel and sexual maturity status of young boars. Livest. Prod. Sci. 63, 265–274.
Szyda, J., Grindflek, E., Liu, Z., Lien, S., 2003. Multivariate mixed inheritance models for QTL detection on porcine chromosome 6. Genet. Res. Camb. 81, 65–73.
Tajet, H., Andresen, O., Meuwissen, T., 2006. Estimation of genetic parameters of boar taint; skatole and androstenone and their correlations with sexual maturation. Acta Vet. Scand. 48 (Suppl. 1), S9.
Tomiyama, M., Oikawa, T., Hoque, M.A., Kanetani, T., Mori, H., 2009. Influence of early postweaning traits on genetic improvement of meat productivity in purebred Berkshire pigs. J. Anim. Sci. 87, 1613–1619.
Toosi, A., Fernando, R.L., Dekkers, J.C.M., 2010. Genomic selection in admixed and crossbred populations. J. Anim. Sci. 88 (1), 32–46.
Usai, M.G., Goddard, M.E., Hayes, B.J., 2009. LASSO with cross-validation for genomic selection. Genet. Res. 91, 427–436.
van Wijk, H.J., Arts, D.J.G., Matthews, J.O., Webster, M., Ducro, B.J., Knol, E. F., 2005. Genetic parameters for carcass composition and pork quality estimated in a commercial production chain. J. Anim. Sci. 83, 324–333 .
van Wijk, H.J., Dibbits, B., Baron, E.E., Brings, A.D., Harlizius, B., Groenen, M.A.M., Knol, E.F., Bovenhuis, H., 2006. Identification of quantitative trait loci for carcass composition and pork quality traits in a commercial finishing cross. J. Anim. Sci. 84, 789–799.
Windig, J.J., Mulder, H.A., ten Napel, J., Knol, E.F., Mathur, P.K., Crump, R.E., 2012. Genetic parameters for androstenone, skatole, indole, and human nose scores as measures of boar taint and their relationship with finishing traits. J. Anim. Sci. 90, 2120–2129.