Research Article

Fine mapping and single nucleotide polymorphism effects estimation on pig chromosomes 1, 4, 7, 8, 17 and X

André M. Hidalgo[1], Paulo S. Lopes[1], Débora M. Paixão[1], Fabyano F. Silva[2], John W.M. Bastiaansen[3], Samuel R. Paiva[4], Danielle A. Faria[4] and Simone E.F. Guimarães[1]

[1]*Departamento de Zootecnia, Universidade Federal de Viçosa, Viçosa, MG, Brazil.*
[2]*Departamento de Estatística, Universidade Federal de Viçosa, Viçosa, MG, Brazil.*
[3]*Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands.*
[4]*Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, Brazil.*

Abstract

Fine mapping of quantitative trait loci (QTL) from previous linkage studies was performed on pig chromosomes 1, 4, 7, 8, 17, and X which were known to harbor QTL. Traits were divided into: growth performance, carcass, internal organs, cut yields, and meat quality. Fifty families were used of a F2 population produced by crossing local Brazilian Piau boars with commercial sows. The linkage map consisted of 237 SNP and 37 microsatellite markers covering 866 centimorgans. QTL were identified by regression interval mapping using GridQTL. Individual marker effects were estimated by Bayesian LASSO regression using R. In total, 32 QTL affecting the evaluated traits were detected along the chromosomes studied. Seven of the QTL were known from previous studies using our F2 population, and 25 novel QTL resulted from the increased marker coverage. Six of the seven QTL that were significant at the 5% genome-wide level had SNPs within their confidence interval whose effects were among the 5% largest effects. The combined use of microsatellites along with SNP markers increased the saturation of the genome map and led to smaller confidence intervals of the QTL. The results showed that the tested models yield similar improvements in QTL mapping accuracy.

Keywords: Bayesian LASSO, Piau breed, pig genetics, QTL.

Received: March 11, 2013; Accepted: August 26, 2013.

Introduction

Quantitative trait loci (QTL) mapping efforts often result in detection of genomic regions that explain part of the quantitative trait variation. However, these regions are usually so large that they do not allow accurate identification of the responsible genes or variants. By using single nucleotide polymorphism (SNP) in the analysis, the genome can be saturated with more markers and the interval of these QTL may become narrowed. Making QTL regions as small as possible is a first step in the process towards the identification of the relevant gene(s) and the respective causative mutation(s).

Previous studies from our research group were conducted on the same population and detected QTL by means of microsatellite markers. A combined total of 40 QTL for growth performance, meat quality, internal organs, cut yield, and carcass composition were found in studies by Paixão *et al.* (2008, 2012, 2013), Silva *et al.* (2008), and

Sousa *et al.* (2011): five on chromosome 1 (SSC1), 12 on chromosome 4 (SSC4), nine on chromosome 7 (SSC7), eight on chromosome 8 (SSC8), three on chromosome 17 (SSC17), and three on chromosome X (SSCX). The sparse genetic maps that were used led to the detection of QTL with large confidence intervals. Combining the microsatellite genotypes with new information from SNP markers in these regions will allow fine mapping and reveal the true positions of these QTL.

One important issue when performing fine mapping is how to analyze the data and combine the resulting information, since linkage mapping and genome-wide association (GWA) use different statistical approaches. Linkage mapping is based on simple linear regression using line of origin probabilities of the genotypes (Haley *et al.*, 1994). In contrast, certain GWA methods are based on multiple regression models where the allele substitution effects for all SNPs are estimated simultaneously (*e.g.* Meuwissen *et al.*, 2001). The multiple regression models can have estimation problems that stem from multicollinearity between markers, requiring some special statistical treatment. The Bayesian LASSO regression (BLR) (Park and Casella, 2008)

Send correspondence to Paulo S. Lopes. Departamento de Zootecnia, Universidade Federal de Viçosa, Av. PH Rolfs s/n, 36570-000 Viçosa, MG, Brazil. E-mail: plopes@ufv.br.

combines desirable features of variable selection and regularization via shrinkage of the regression coefficients (de los Campos *et al.*, 2009).

The objective of this study was to (i) fine map chromosomes that had QTL on our pig population: SSC1, SSC4, SSC7, SSC8, SSC17 and SSCX with increased marker coverage; and (ii) to use complementary information from SNP marker effects estimated by BLR to determine the concordance between the positions of SNPs with the top 5% estimates and the regions covered by a QTL confidence interval.

Material and Methods

Experimental population and phenotypic data

All procedures with animals were carried out in accordance with the Ethics Statements of the Department of Animal Science, Federal University of Viçosa (UFV), MG, Brazil.

A three-generation resource population was created and managed as described by Band *et al.* (2005a). Briefly, two local Piau breed grandsires were crossed with 18 granddams, composed of Large White, Landrace, and Pietrain breeds, producing the F1 generation from which 11 F1 sires and 54 F1 dams were randomly selected (Peixoto *et al.*, 2006). These F1 individuals were crossed to produce 627 F2 animals. Piau is a local unimproved breed with high level of fatness (Serão *et al.*, 2011).

Phenotype collection was described in Band *et al.* (2005a,b). The 54 analyzed traits were divided into five categories: (a) Growth performance: slaughter age (SA), birth weight (BW), total teat number (TN), body weight at 21, 42, 63, 77 and 105 days (W21, W42, W63, W77, W105), slaughter weight (SW), average daily gain from 77 to 105 days (ADG), feed intake from 77 to 105 days (FI), feed conversion from 77 to 105 days (FC); (b) Meat quality: drip loss (DL), cooking loss (CL), total loss (TL), shear force (SF), pH 45 min after slaughter (pH45), pH 24 hours after slaughter (pH24), redness (A), yellowness (B), saturation (C), lightness (L), hue angle (H), intramuscular fat percentage (IMF); (c) Internal organs: lung weight (LUNG), heart weight (HEART), liver weight (LIVER), small intestine length (SIL), spleen weight (SPLEEN), kidney weight (KIDNEY); (d) Cut yields: abdominal fat weight (AF), bacon weight (BCW), boneless loin weight (LW), total loin weight (TLW), total ham weight (THW), trimmed ham weight (TRIMHW), total Boston shoulder weight (TBSW), trimmed Boston shoulder weight (TRIMBSW), jowl weight (JW), total picnic shoulder weight (TPSW), trimmed picnic shoulder weight (TRIMPSW), rib weight (RW), sirloin weight (SLW); (e) Carcass: loin eye area (LEA), carcass length by the Brazilian carcass classification method (CLBRA), carcass length by the American carcass classification method (CLUSA), carcass yield including feet and head (CY), the thickest backfat thickness on the

shoulder region (SBF), midline thinnest backfat thickness above the last lumbar vertebrae (LBF), midline backfat thickness between last and penultimate lumbar vertebrae (PBF), midline backfat thickness immediately after the last rib (LRBF), backfat thickness at P2 site (last rib, 6.5 cm from the midline) (P2BF), bacon depth (BD), loin depth (LD). Estimates of heritabilities and phenotypic and genetic correlations of the studied traits are shown in Mendonça *et al.* (2012).

DNA extraction, SNP selection and genotyping

DNA was extracted at the Animal Biotechnology Laboratory of the Department of Animal Science at the Federal University of Viçosa. Genomic DNA was extracted from white blood cells of grand parental, F1, and F2 animals, described in Band *et al.* (2005b).

The 384 SNPs used for fine mapping and estimation of marker effects were selected from the Illumina Porcine SNP60 Beadchip (San Diego, CA, USA, Ramos *et al.*, 2009). SNPs spanned the entire chromosomes where QTL had previously been detected in our population. The distribution over chromosomes was as follows: SSC1 (85), SSC4 (71), SSC7 (84), SSC8 (42), SSC17 (36), and SSCX (66). Genotyping of the 384 SNPs was performed by the Golden Gate/VeraCode technology, using the BeadXpress reader from Illumina. From these 384 SNPs, 66 were discarded due to non-amplification. From the remaining 318 SNPs, 81 were discarded due to a minor allele frequency (MAF) smaller than 0.05. These final 237 SNPs were distributed as follows: SSC1 (56), SSC4 (54), SSC7 (59), SSC8 (31), SSC17 (25), and SSCX (12). Genotypes from 37 microsatellite markers were available from previous studies (Paixão *et al.*, 2008, 2012, 2013; Silva *et al.*, 2008; Sousa *et al.*, 2011) and included in the analyses (Table 1).

Statistical analysis - QTL mapping

Genetic distance between markers was extrapolated from the physical distance (1Mb = 1 cM) (Amaral *et al.*, 2008) to build the combined map of microsatellites and SNP. The combined genotypic data was used in linkage analysis using the regression method described by Haley *et al.* (1994), implemented in GridQTL (Seaton *et al.*, 2006).

Table 1 - Number of SNP and microsatellite markers, chromosome length (cM) and average distance between markers (cM) in *Sus scrofa* chromosomes 1, 4, 7, 8, 17, and X.

SSC	Nº SNP	Nº microsatellites	Chromosome length	Average distance
1	56	5	290	4.75
4	54	6	128	2.13
7	59	6	133	2.05
8	31	7	118	3.50
17	25	7	67	2.09
X	12	6	132	7.33

The statistical model assumed that the putative QTL is diallelic, with alternative alleles fixed in each of the grand parental breed. The probability that each F2 individual carries each of the three QTL genotypes was calculated according to the genotype of the markers at 1 cM intervals along the chromosome. From these probabilities the additive and dominance coefficients were calculated and used to regress the individual phenotypes for each animal.

Chromosome-wide significance thresholds ($\alpha = 0.05$ and $\alpha = 0.01$) were obtained by permutation, with 10,000 permutations per test (Churchill and Doerge, 1994). The genome-wide significance thresholds ($\alpha = 0.05$ and $\alpha = 0.01$) were obtained using a Bonferroni correction (Knott *et al.*, 1998). The 95% confidence interval (95% CI) for the QTL location was calculated using the chi-square approximation (χ^2) as described by Pérez-Enciso *et al.* (2000).

The following statistical model was adopted:

$$y_{ijkl} = S_i + L_j + H_k + (C_{ijkl} - \overline{C})b + c_a a + c_d d + e_{ijkl} \qquad (1)$$

where y_{ijk} = phenotype; S_i = fixed effect of sex i; L_j = fixed effect of batch j, $j = 1, 2, 3, 4, 5$; H_k = fixed effect of the halothane genotype k, $k = 1$ (*NN*), 2 (*Nn*); $(C_{ijkl} - \overline{C})b$ = adjustment for covariates; $c_a = P(QQ/Mi) - P(qq/Mi)$; $c_d = P(Qq/Mi)$; e_{ijkl} = residual errors.

The halothane genotype was included as a fixed effect since Band *et al.* (2005a,b) reported significant effects of the Hal[1843] mutation on performance, carcass, and meat quality traits in our population. Carcass weight at slaughter was included as a covariate for carcass and internal organ traits; age at slaughter was included for meat quality traits. Litter size was included as a covariate for birth weight; litter size at weaning was included for weight at 21, 42, 63, 77, 105 days, and slaughter weight; weight at 77 days was included for feed conversion, feed intake and average daily gain.

The F ratio was calculated at each position, comparing the model with a QTL to the equivalent model without QTL. Estimates for *a* and *d* were calculated at the best estimated position with the highest F ratio. The additive fraction of phenotypic variance (h^2_Q) in the F2 generation explained by a given QTL was computed according to Pérez-Enciso *et al.* (2000). The conditional probability functions of the QTL given the genotype of the markers (c_a and c_d) were estimated according to Haley *et al.* (1994).

Statistical analysis - SNP effect estimates

In the GWA analysis the phenotypic outcomes, y_i ($i = 1, 2, ..., 345$), were regressed on marker covariates x_{ik} ($k = 1, 2, ..., 237$) and the same fixed effects and covariates were used as in the linkage analysis following the regression model proposed by Meuwissen *et al.* (2001):

$$y_i = \mu + \mathbf{Z}_i \phi_i \sum_{k=1}^{237}(x_{ik}\beta_k) + e_i \qquad (2)$$

where y_i is the phenotypic observation of animal i, μ is the general mean, ϕ_i is a set of fixed effects relative to each animal i with incidence matrix $\mathbf{Z}_i$, β_k is the effect of marker k, and e$_i$ the residual term, $e_i \sim N(0, \sigma_e^2)$. In this model, x_{ik} take the values 0, 1, and 2 for the SNP genotypes *AA*, *Aa*, and *aa* at each locus k, respectively. Under a matrix notation, the presented GWA model can be rewritten as:

$$\mathbf{y} = \mathbf{1'}\mu + \mathbf{Z}\phi + \mathbf{I}\sum_{k=1}^{237}\mathbf{X}_k b_k + \mathbf{e} \qquad (3)$$

where $\mathbf{1'}$, $\mathbf{I}$, and $\mathbf{Z}$ are, respectively, a unit vector, an identity matrix, and a fixed effect incidence matrix ($\mathbf{Z}_{345 \times Nf}$), being Nf the number of fixed effects; $\mathbf{y} = [y_1, y_2, ..., y_{345}]\mathbf{'}_{345 \times 1}$, $\mathbf{X}_k = [x_{1k}, x_{2k}, ..., x_{345k}]\mathbf{'}_{345 \times 1}$, and $\mathbf{e} = [e_1, e_2, ..., e_{345}]\mathbf{'}_{345 \times 1}$.

Solutions from model 3 were obtained using BLR (de los Campos *et al.*, 2009), assuming that each locus explains its own amount of the variation. The BLR is a penalized Bayesian regression procedure whose general estimator is given by

$$\hat{\beta} = \arg \min{}_\beta \left\{ \left[\hat{\mathbf{y}} - (\mathbf{1'}\mu + \mathbf{Z}\phi + \mathbf{I}\sum_{k=1}^{237}\mathbf{X}_k\beta_k)\right]\mathbf{'} \left[\hat{\mathbf{y}} - (\mathbf{1'}\mu + \mathbf{Z}\phi + \mathbf{I}\sum_{k=1}^{237}\mathbf{X}_k\beta_k)\right] + \lambda\sum_{k=1}^{237}|\beta_k| \right\}$$

where λ is the regularization parameter. When $\lambda = 0$ there is no regularization, and when $\lambda > 0$ there is a shrinkage of the marker effects toward zero, with the possibility of setting some redundant effects (β's) identically equal to zero, resulting in a simultaneous estimation and variable selection procedure. The BLR package (de los Campos *et al.*, 2009; Pérez *et al.*, 2010) of R (R Development Core Team, 2011) was used, which assumes that the joint prior distribution of marker effects (β_1, β_2, ..., β_{237}) is $\prod_{k=1}^{237} N(0, \sigma_{\beta_k}^2)$, being $\sigma_{\beta_k}^2 = \sigma_e^2 \tau_k^2$, where σ_e^2 is the residual variance, with a scaled inverse χ^2 prior distribution, and τ_k^2 is the scale parameter related to each marker. In turn, the BLR also assumes that the joint prior distribution for the scale parameters ($\tau_1^2, \tau_2^2, ... \tau_{237}^2$) is the product of Exponential distributions, and that the λ prior distribution is $\Gamma(\nu_1, \nu_2)$. The BLR was implemented using 50,000 MCMC iterations with a burn-in equal to 10,000 iterations. Chain length was validated for each MCMC separately using Geweke convergence diagnostics implemented in the BOA package (Bayesian Output Analysis, Smith, 2007).

The large effects SNPs were identified for each trait as those SNPs with absolute values within the top 5%. The genome positions of these markers with large effects were used to analyze overlap with QTL regions.

Results

QTL analysis

In total, 32 QTL were detected that surpassed the 5% chromosome-wide significance level (CWL) (Table 2). From these 32 QTL, 12 surpassed the 1% CWL and 7 the 5% genome-wide level (GWL). The 32 QTL included 7 that were found in previous studies on this F2 population, and 25 novel QTL were detected by applying linkage analysis with the increased marker coverage.

Confirmation of previously known QTL

None of the previously found QTL on SSC1 surpassed the 5% CWL in the current analysis. The QTL for

Table 2 - Evidence for significant QTL for various traits. Number of animals analyzed for each trait (N), phenotypic mean (standard deviation), location, confidence interval at 95% (CI), maximum F statistics (Fmax), estimates of additive and dominance effects, standard errors (SE) and % of phenotypic variance in the F2 explained by each QTL (h^2_Q) for significant traits on *Sus scrofa* chromosomes 1, 4, 7, 8, 17 and X.

Trait	N	Mean (SD)	SSC	Position (CI)	Fmax	Additive ± SE	Dominance ± SE	h^2_Q
Previously known QTL confirmed								
Heart weight (HEART), kg	425	0.235 (0.030)	4	75 (71-80)	8.41**	-0.01 ± 0.01	-0.01 ± 0.00	3.81
Abdominal fat (AF), kg	537	0.457 (0.160)	7	45 (42-48)	7.19*	-0.04 ± 0.01	0.01 ± 0.01	4.06
Carcass length MBCC (CLBRA)[1], cm	425	86.01 (4.139)	7	44 (41-49)	9.68***	1.12 ± 0.26	-0.58 ± 0.37	6.01
Carcass length MLC (CLUSA)[2], cm	424	71.64 (3.421)	7	47 (42-50)	9.17***	0.93 ± 0.22	-0.31 ± 0.31	6.34
L Backfat (LBF)[3], mm	425	23.23 (6.110)	8	15 (12-20)	9.69***	2.04 ± 0.47	-0.09 ± 0.65	7.30
Weight at 77 days (W77), kg	427	21.387 (4.334)	17	66 (64-66)	8.26**	0.96 ± 0.33	-1.44 ± 0.47	2.68
Redness (A)	424	0.693 (0.722)	X	102 (95-104)	8.92*	0.19 ± 0.06	-	3.30
New QTL detected								
Slaughter age (SA), days	424	148.212 (10.497)	1	240 (232-259)	6.50*	-3.10 ± 0.95	-1.72 ± 1.50	5.03
Birth weight (BW), kg	415	1.201 (0.273)	1	15 (10 - 38)	7.06*	0.11 ± 0.04	-0.28 ± 0.09	8.72
Drip loss (DL), %	427	3.157 (1.743)	1	240 (237 - 254)	7.33*	0.18 ± 0.16	0.90 ± 0.26	0.60
Loin eye area (LEA), cm²	390	26.43 (4.034)	1	200 (198-210)	8.70***	1.13 ± 0.27	0.02 ± 0.44	5.60
Cooking loss (CL), %	426	32.46 (2.512)	4	110 (106-117)	6.31*	0.09 ± 0.2	1.09 ± 0.31	0.07
Shear force (SF), kg/cm²	391	5505.95 (958.73)	4	0 (0-3)	6.11*	144.10 ± 77.91	339.72 ± 122.01	1.08
Lung weight (LUNG), kg	422	0.448 (0.077)	4	69 (65-74)	8.80***	-0.03 ± 0.01	-0.001 ± 0.01	6.54
L Backfat (LBF)[3], mm	425	23.23 (6.110)	4	99 (94-104)	7.16*	1.68 ± 0.45	-0.01 ± 0.69	4.97
Shear force (SF), kg/cm²	391	5505.95 (958.73)	7	4 (2-8)	7.15*	27.35 ± 90.64	492.42 ± 137.05	0.04
Liver weight (LIVER), kg	422	1.262 (0.149)	8	24 (20-29)	8.06**	-0.04 ± 0.01	-0.02 ± 0.02	3.90
Small intestine length (SIL), m	426	18.38 (1.870)	8	6 (5-8)	8.31**	-0.48 ± 0.01	-0.39 ± 0.19	3.83
P Backfat (PBF)[4], mm	425	28.34 (5.902)	8	15 (11-21)	6.90*	1.69 ± 0.46	-0.05 ± 0.65	5.23
LR Backfat (LRBF)[5], mm	427	19.61 (4.819)	8	21 (16-24)	6.87*	0.95 ± 0.36	0.96 ± 0.51	2.45
S Backfat (SBF)[6], mm	426	40.57 (5.607)	8	10 (9-15)	9.13***	1.46 ± 0.41	1.06 ± 0.57	4.55
Abdominal fat (AF), kg	537	0.457 (0.160)	8	21 (19-24)	9.34***	0.04 ± 0.01	0.03 ± 0.01	3.97
Birth weight (BW), kg	415	1.201 (0.273)	17	7 (6-9)	7.93**	0.02 ± 0.02	0.12 ± 0.03	0.42
Total teat number (TN)	426	13.11 (1.271)	17	7 (5-8)	7.13*	-0.22 ± 0.09	0.47 ± 0.15	1.47
Weight at 63 days (W63), kg	422	16.245 (3.402)	17	66 (64-66)	7.50*	0.57 ± 0.26	-1.27 ± 0.38	1.50
pH24 hours (pH24), pH	417	5.704 (0.126)	17	49 (47-50)	7.54*	0.02 ± 0.01	0.06 ± 0.01	1.41
Small intestine length (SIL), m	426	18.38 (1.870)	X	106 (103-110)	8.88*	0.46 ± 0.15	-	3.40
Bacon weight (BCW), kg	538	2.683 (0.480)	X	45 (29-59)	9.71*	0.10 ± 0.03	-	5.61
Boneless loin weight (LW), kg	535	1.022 (0.183)	X	89 (62-102)	8.10*	-0.04 ± 0.02	-	4.82
Trimmed HW (TRIMHW)[7], kg	538	4.998 (0.631)	X	99 (86-102)	8.65*	-0.10 ± 0.03	-	3.42
Total Boston SW (TBSW)[8], kg	537	2.326 (0.339)	X	117 (108-124)	9.38*	0.05 ± 0.02	-	4.16
Trimmed Boston SW (TRIMBSW)[9], kg	538	1.679 (0.265)	X	132 (129-132)	10.64*	0.06 ± 0.02	-	4.04

*, ** and *** significant at the 5% chromosome-wide level, 1% chromosome-wide level, and at the 5% genome-wide level, respectively; Positive additive effects indicate that Piau alleles increased the trait and negative, that commercial alleles increased it.
[1]CLBRA, carcass length by the Brazilian carcass classification method; [2]CLUSA, carcass length by the American carcass classification method; [3]LBF, midline lower backfat thickness above the last lumbar vertebrae; [4]PBF, midline backfat thickness between last and penultimate lumbar vertebrae; [5]LRBF, midline backfat thickness immediately after the last rib; [6]SBF, higher backfat thickness on the shoulder region; [7]TRIMHW, trimmed ham weight; [8]TBSW, total Boston shoulder weight; [9]TRIMBSW, trimmed Boston shoulder weight.

HEART on SSC4 was confirmed at the 1% CWL and explained 3.81% of phenotypic variance. On SSC7 the QTL for AF was significant at the 5% CWL, explaining 4.06% of the phenotypic variance. The QTL for CLBRA and CLUSA (SSC7) were significant at the 5% GWL and explained 6.01% and 6.34% of the phenotypic variance, respectively. On SSC8, the QTL for LBF was significant at the 5% GWL, explaining 7.30% of the phenotypic variance. On SSC17, the QTL for W77 was significant at the 1% CWL and explained 2.68% of the phenotypic variance. On SSCX, the QTL for A was significant at the 5% CWL and explained 3.30% of the phenotypic variance. The peaks of the QTL confirmed in the current analysis deviated on average by 33.8 cM from the QTL positions obtained in the previous analysis.

New QTL detected

On average there were four new QTL detected on each of the targeted chromosomes. On SSC1 a QTL for LEA was found surpassing the 5% GWL, explaining 5.60% of the phenotypic variance. In addition three new QTL were found surpassing the 5% CWL on SSC1 for SA, BW, and DL, explaining 5.03%, 8.72%, and 0.60% of the phenotypic variance, respectively. On SSC4 a QTL for LUNG was found surpassing the 5% GWL, explaining 6.54% of the phenotypic variance. In addition three new QTL were found surpassing the 5% CWL on SSC4 for CL, SF and LBF, explaining 0.07%, 1.08%, and 4.97% respectively. On SSC7 a single new QTL was found for SF that was significant at the 5% CWL, explaining 0.04% of the phenotypic variance. On SSC8 six new QTL were found at different thresholds. Two QTL for AF and SBF were significant at the 5% GWL, explaining 3.97% and 4.55% of the phenotypic variance, respectively. Two additional QTL for LIVER and SIL were significant at the 1% CWL, explaining 3.90% and 3.83% of the phenotypic variance, respec-

tively. The final two QTL on SSC8 for PBF and LRBF were significant at the 5% CWL and explained 5.23% and 2.45%, respectively. On SSC17 one QTL for BW was significant at the 1% CWL, explaining 0.42% of the phenotypic variance, and three additional QTL for TN, W63, and pH24 were significant at the 5% CWL, explaining 1.47%, 1.50%, and 1.41% of the phenotypic variance, respectively. On SSCX all six new QTL for SIL, BCW, LW, TRIMHW, TBSW and TRIMBSW were significant at the 5% CWL and respectively explained 3.40%, 5.61%, 4.82%, 3.42%, 4.16%, and 4.04%, of the phenotypic variance.

Confidence interval

Of the 32 QTL described in the current analysis, most (23 QTL) were mapped with a 95% confidence interval of 10 cM or less. Only 4 of the QTL were mapped with a confidence interval larger than 20 cM. Confidence intervals of QTL that were detected previously and confirmed in this study were reduced by 23.9 cM on average (Figure 1).

Top 5% SNPs with the largest effect

Six of the seven QTL that were significant at the 5% GWL had one top 5% SNP within their QTL confidence intervals. For 11 of the 32 QTL described in the current analysis at least one marker from the top 5% SNPs was located within their QTL confidence interval. Seven of these 11 QTL harbored exactly one of the top 5% SNP: QTL affecting CL, LUNG and HEART on SSC4, AF and CLBRA on SSC7, LBF on SSC8, and BCW on SSCX. The other four QTL each harbored exactly two top 5% SNPs within their confidence interval: QTL affecting LEA on SSC1, CLUSA on SSC7, AF and SBF on SSC8.

Discussion

A QTL mapping study was carried out and QTL confidence intervals were inspected for harboring the positions
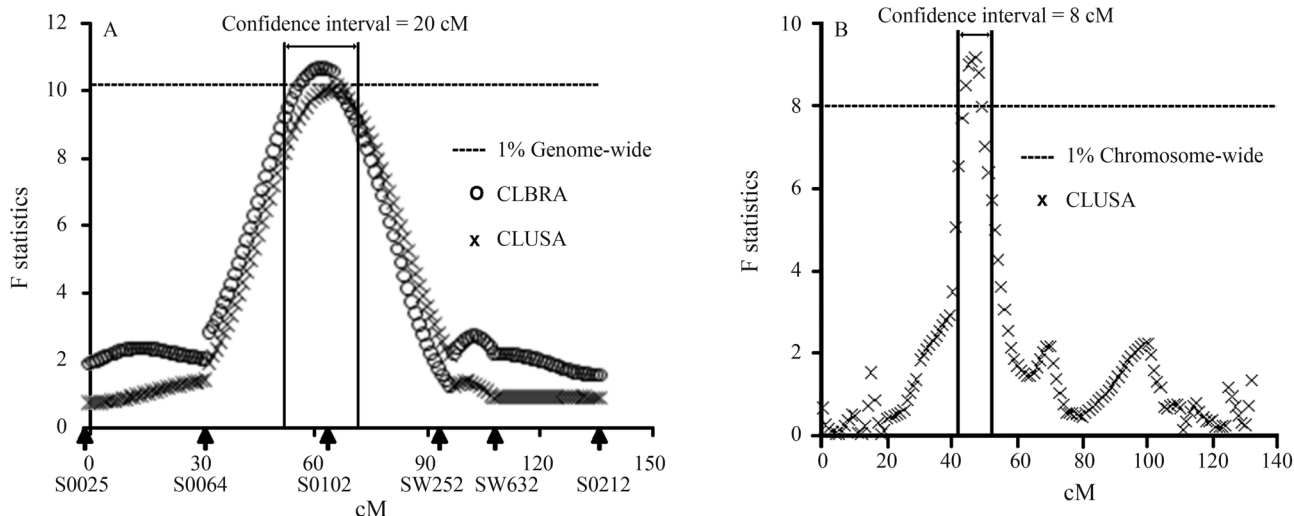

Figure 1 - Example of confidence interval reduction from A (QTL for CLUSA from Sousa *et al.* (2011)) to B (QTL for CLUSA from the current study).

of any of the top 5% SNPs by means of the Bayesian LASSO method. Using the QTL regression approach, 32 QTL were detected at the 5% CWL using a combined genetic map. Eight of these had not been reported in the consulted literature: SF and LUNG on SSC4; SBF, PBF and LRBF on SSC8; BW on SSC17; SIL and BCW on SSCX (PigQTLdb). Compared to previous studies that used the same F2 population and relied only on microsatellite markers (Paixão *et al.*, 2008, 2012, 2013; Silva *et al.*, 2008; Sousa *et al.*, 2011), only seven of 40 QTL were confirmed in the present analysis. For these confirmed QTL the confidence intervals were narrowed down on average by 23.9 cM using the dataset with increased marker coverage. Given the increase in power in the present study, from adding SNP markers to the existing microsatellite map, we infer that QTL that were not confirmed in the current study are most likely false positives.

Confirmation of previously known QTL

On SSC4 a QTL associated with HEART (Silva *et al.*, 2008) was confirmed. The new analysis positions the QTL in the same interval, between the S0001 and S0217 microsatellite markers, but Silva *et al.* (2008) reported a much larger confidence interval (68 cM) than obtained in the present study (9 cM), showing an increase in mapping precision.

On SSC7, QTL were confirmed for AF, CLBRA and CLUSA, previously reported in this population by Sousa *et al.* (2011). The QTL for AF and CLBRA were located between microsatellite markers S0064 and S0102, whereas the QTL for CLUSA was located in the neighboring interval, between microsatellite markers S0102 and SW252. The current results place the SSC7 QTL in the same microsatellite intervals as before, but again with much smaller confidence intervals. A QTL on SSC7 was found by Mikawa *et al.* (2011), affecting the number of vertebrae, which would increase carcass length. The marker SW252 is also flanking a QTL in their study and the *vertnin* gene is the suspected cause of variation in the number of vertebrae in commercial populations. This gene could affect other traits as there are other QTL near this region, an example is the QTL for AF that we confirmed. The QTL affecting CLBRA and CLUSA were detected close together on SSC7, suggesting that the same gene possibly affects these two similar traits. The Piau alleles at CLBRA and CLUSA were associated with longer carcasses, which is not expected, as an increased carcass length would be expected from the larger commercial breed. These sources of cryptic variation are, however, known to exist (Abasht *et al.*, 2006).

On SSC8 the QTL for LBF detected by Sousa *et al.* (2011) was confirmed. The QTL was located between the microsatellite markers SW905 and S0017. The current study narrowed the confidence interval from 12 cM to 8 cM, which is not as dramatic as for some other QTL, but this is mainly because the interval was already quite narrow. Piau alleles were associated with higher values of backfat thickness as measured by LBF, which is expected since Piau is a breed with a high level of fatness.

On SSC17 the QTL for W77 detected by Paixão *et al.* (2008) was also confirmed. The QTL was located between the microsatellite markers S0359 and SW2427. The original confidence interval of 35 cM was dramatically reduced to 2 cM in the present study. Piau alleles were associated with higher W77, which was not expected from the phenotypic means of the grand parental populations, and like the results for carcass length on SSC7 this again indicates the presence of cryptic variation.

On SSCX a QTL associated with meat color, A, identified by Paixão *et al.* (2012) was confirmed. The peak of the QTL remained located between markers SW1943 and S0218. Our confidence interval was much smaller (9 cM) than theirs (33 cM). Piau alleles were related to an increase in A, following the expectation, as values for A were higher in Piau than in the commercial breed.

New QTL detected

On SSC1 four new QTL were detected from this resource family. The QTL affecting SA was considered a new QTL because the SA QTL detected by Paixão *et al.* (2013) on this chromosome was located at a different position, in a different marker interval. We detected a QTL affecting BW near a QTL previously reported by Knott *et al.* (1998) and Beeckmann *et al.* (2003), located in the proximal region of the chromosome. A QTL associated with DL, as detected here, was also previously reported, but at a different position (*e.g.* Ponsuksili *et al.*, 2008). The QTL affecting LEA may coincide with the one detected by Malek *et al.* (2001) and Grapes and Rothschild (2006). Even though it was considered a cryptic effect, alleles originating from the Piau breed were shown to increase LEA, this corresponding with results in the literature where Berkshire alleles, which is considered a fatter breed, were found to increase growth and leanness (Malek *et al.*, 2001; Grapes and Rothschild 2006).

On SSC4, QTL for SF and LUNG were detected for which we did not find any previous reports in the literature. The QTL affecting CL was also detected by Große-Brinkhaus *et al.* (2010) in a Duroc x Pietrain cross. The QTL for LBF that we found was also detected by Silva *et al.* (2008) and Malek *et al.* (2001). Piau alleles were estimated to increase backfat thickness as expected. On SSC7 we detected a QTL affecting SF, for which a QTL has also been reported by Edwards *et al.* (2008) on this chromosome, but at different location.

On SSC8, QTL for SBF and PBF on SSC8 have not previously been reported. Other previously mapped QTL, such as the one affecting LIVER, was mapped in the same interval by Beeckmann *et al.* (2003), and QTL associated with AF (Knott *et al.*, 1998; Sousa *et al.*, 2011), and LRBF

(Fan *et al.*, 2011) have also been reported. For all backfat traits, the Piau alleles would cause an increase in backfat thickness. The QTL associated with SIL was previously reported by Knott *et al.* (1998) in a cross between European wild pigs and Large White, by Gao *et al.* (2010) in a Duroc x Erhualian cross, and by Sousa *et al.* (2011) in the current reference population. These other reported QTL were located at some distance from the current QTL. The commercial breed alleles were associated with longer SIL (0.48), supporting the hypothesis that small intestine length increased in response to selection and domestication, as proposed by Andersson *et al.* (1994)

On SSC17, a QTL affecting BW was detected and has not been reported before. QTL associated with TN (Guo *et al.*, 2008), with W63 (Paixão *et al.*, 2008) and with pH24 (Wimmers *et al.*, 2007) were previously reported, but in different chromosomal regions. BW is increased by the Piau alleles, and commercial breed alleles increase TN, which were expected effects in view of the higher fatness of the Piau breed and higher number of piglets per litter in commercial breeds. Piau alleles were related to an increase in pH24, following the expectation of higher pH24 for Piau than for commercial breeds.

On SSCX a QTL for BCW was detected at 45 cM, and a QTL for SIL at 106 cM for which no previous reports were found in the literature. Piau alleles were associated with longer SIL, which was against expectation (Andersson *et al.*, 1994), and different from the other QTL effects detected for SIL in this study. Three of the four remaining new QTL, for LW, TRIMHW, and TRIMBSW, were previously reported by Milan *et al.* (2002). The QTL for TRIMHW was also reported by Cepica *et al.* (2003) who additionally reported a QTL for TBSW. The estimated additive effect of the QTL affecting BCW, TRIMBSW and TBSW implied that Piau alleles increase the phenotype for these traits, and that LW and TRIMHW are increased by commercial breed alleles.

The Piau breed has never undergone strong selection for lean growth, as is common in current commercial breeding programs, explaining the higher carcass fatness of the Piau. For QTL related to fatness, such as QTL for backfat on SSC8, for AF on SSC8, and for BCW on SSCX, the Piau breed alleles were expected to result in more fat. For QTL related to growth and meat weight, like LW and TRIMHW on SSCX, the Piau breed alleles were expected to result in less growth. Many of the new QTL detected in this study did not follow this expectation. Instead, many new QTL showed cryptic effects where the alleles of the Piau breed increased growth or decreased fatness. QTL with cryptic effects were BW and LEA on SSC1, BW and W63 on SSC17, and TRIMBSW, TBSW and SIL on SSCX. While cryptic QTL effects are unexpected, they are not uncommon. In other studies on pigs (Yue *et al.*, 2003), as well as studies on different species (Abasht *et al.*, 2006), cryptic QTL effects have been shown.

The length of the intestine, for which a QTL was found on SSCX (SIL), is an important factor affecting the potential to grow, possibly by influencing the nutrient absorption efficiency and digestion (Gao *et al.*, 2010). It was expected that alleles from commercial breeds would be associated with longer intestine length, but the opposite was found. A similar cryptic allelic effect was also reported by Gao *et al.* (2010) for intestine length on a different chromosome, SSC7, using a White Duroc X Chinese Erhualian intercross resource population. In addition to cryptic QTL, other QTL were found by Gao *et al.* (2010), where the alleles for higher intestine length came from the commercial White Duroc breed. We speculate that alleles from local breeds cause an increase in SIL due to their adaption to low quality feed, which requires better digestion and higher absorption efficiency. This can be achieved by an increased time of digestion of the feed provided a longer small intestine length.

Marker effects

Eleven out of the 32 QTL confidence intervals covered at least one of the top 5% SNP from the BLR analysis. Six of the seven QTL that surpassed the 5% GWL each contained one top 5% SNP within their confidence intervals. The only genome-wide significant QTL without a top 5% SNP was found on SSC4 for LUNG (CI = 65 cM - 74 cM). Nonetheless, the marker ALGA0025795, located at 70 cM, had the largest effect in the region (0.00024) and was immediately below the significance threshold for inclusion in the top 5%. Out of the remaining 25 QTL that were significant at the 5% CWL, only four contained a top 5% SNP within their confidence intervals. The smaller proportion of overlap with a top 5% SNP for the chromosome-wide significant QTL is probably due to the smaller amount of variance explained by these chromosome-wide significant QTL. The overlap between results from linkage mapping and effects of individual markers based on association analysis corroborated to some extent the QTL found by the two models, especially for the genome-wide significant QTL. Given the increase in power in the present study compared to the previous analyses of this resource population, we infer that QTL that were found in previous, but not in the current study, were false positives. On the other hand, there is also a chance that new QTL detected at CWL, most of which did not present a top 5% SNP within their confidence interval, are also false positives.

In summary, the addition of more markers and animal genotypes increased the statistical power for QTL detection compared to previous studies and lead to QTL with much smaller confidence intervals. Seven previously discovered QTL were confirmed, 25 novel QTL were identified, and 33 QTL that were detected in previous studies were lost. Most of the genome-wide significant QTL contained at least one of the top 5% SNP effects estimated by the Bayesian approach, corroborating the QTL found by the re-

gression method and showing that both models can be used to refine QTL mapping results. With decreasing SNP genotyping costs, updating existing QTL studies with low density SNP genotypes can be a fruitful approach to improve statistical power to detect QTL and reduce confidence intervals.

Acknowledgments

References

Abasht B, Dekkers JCM and Lamont SJ (2006) Review of quantitative trait loci identified in the chicken. Poultry Sci 85:2079-2096.

Amaral AJ, Megens HJ, Crooijmans RPMA, Heuven HCM and Groenen MAM (2008) Linkage disequilibrium decay and haplotype block structure in the pig. Genetics 179:569-579.

Andersson L, Haley CS, Ellegren H, Knott SA, Johansson M, Andersson K, Ansersson-Eklund L, Edfors-Lilja I, Fredholm M, Hansson I, *et al.* (1994) Genetic mapping of quantitative trait loci for growth and fatness in pigs. Science 263:1771-1774.

Band GDO, Guimarães SEF, Lopes PS, Schierholt AS, Silva KM, Pires AV, Benevenuto Júnior AA and Gomide LAM (2005a) Relationship between the Porcine Stress Syndrome gene and pork quality traits of F2 pigs resulting from divergent crosses. Genet Mol Biol 28:88-91.

Band GDO, Guimarães SEF, Lopes PS, Peixoto JDO, Faria DA, Pires AV, Figueiredo FC, Nascimento CS and Gomide LAM (2005b) Relationship between the Porcine Stress Syndrome gene and carcass and performance traits in F2 pigs resulting from divergent crosses. Genet Mol Biol 28:92-96.

Beeckmann P, Moser G, Bartenschlager H, Reiner G and Geldermann H (2003) Linkage and QTL mapping for *Sus scrofa* chromosome 8. J Anim Breed Genet 120(suppl 1):66-73.

Cepica S, Stratil A, Kopecny M, Blazkova P, Schroffel J, Davoli R, Fontanesi L, Reiner G, Bartenschlager H, Moser G, *et al.* (2003) Linkage and QTL mapping for *Sus scrofa* chromosome 4. J Anim Breed Genet 120(suppl 1):28-37.

Churchill GA and Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963-971.

de los Campos G, Naya H, Gianola D, Crossa J, Legarra A, Manfredi E, Weigel K and Cotes JM (2009) Predicting quantitative traits with regression models for dense molecular markers and pedigree. Genetics 182:375-385.

Edwards DB, Ernst CW, Raney NE, Doumit ME, Hoge MD and Bates RO (2008) Quantitative trait locus mapping in an F2 Duroc x Pietrain resourece population: II. Carcass and meat quality traits. J Anim Sci 86:254-266.

Fan B, Onteru SK, Du ZQ, Garrick DJ, Stalder KJ and Rothschild MF (2011) Genome-wide association study identifies loci for body composition and structural soundness traits in pigs. PLoS One 6:e14726.

Gao J, Ren J, Zhou LH, Ren DR, Li L, Xiao SJ, Yang B and Huang LS (2010) A genome scan for quantitative trait loci affecting the length of small intestine in a White Duroc x Chinese Erhualian intercross resource population. J Anim Breed Genet 127:119-124.

Grapes L and Rothschild MF (2006) Investigation of a QTL region for loin eye area and fatness on pig chromosome 1. Mamm Genome 17:657-668.

Große-Brinkhaus C, Jonas E, Buschbell H, Phatsara C, Tesfaye D, Jüngst H, Looft C, Schellander K and Tholen E (2010) Epistatic QTL pairs associated with meat quality and carcass composition traits in a porcine Duroc x Pietrain population. Genet Sel Evol 42:39

Guo YM, Lee GJ, Archibald L and Haley CS (2008) Quantitative trait loci for production traits in pigs: A combined analysis of two Meishan x Large White populations. Anim Genet 39:486-495.

Haley CS, Knott SA and Elsen JM (1994) Mapping quantitative trait loci in crosses between outbred lines using least squares. Genetics 136:1195-1207.

Knott SA, Marklund L, Haley CS, Andersson K, Davies W, Ellegren H, Fredholm M, Hansson I, Hoyheim B, Lundström K, *et al.* (1998) Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and large white pigs. Genetics 149:1069-1080.

Malek M, Dekkers JC, Lee HK, Baas TJ and Rothschild MF (2001) A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. I. Growth and body composition. Mamm Genome 12:630-636.

Mendonça PT, Lopes PS, Braccini Neto J, Carneiro PLS, Torres RA, Guimarães SEF and Veroneze R (2012) Estimação de parâmetros genéticos de uma população F2 de suínos. Rev Bras Saúde Prod Animal 13:330-343.

Meuwissen TH, Hayes BJ and Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819-1829.

Mikawa S, Sato S, Nii M, Morozumi T, Yoshioka G, Imaeda N, Yamaguchi T, Hayashi T and Awata T (2011) Identification of a second gene associated with variation in vertebral number in domestic pigs. BMC Genetics 12:e5.

Milan D, Bidanel JP, Iannuccelli N, Riquet J, Amigues Y, Gruand J, Le Roy P, Renard C and Chevalet C (2002) Detection of quantitative trait loci for carcass composition traits in pigs. Genet Sel Evol 34:705-728.

Paixão DM, Silva Filho MID, Pereira MS, Lopes MS, Barbosa L, Souza KRS, Lopes PS and Guimarães SEF (2008) Quantitative trait loci for carcass, internal organ and meat quality traits on porcine chromosomes 16, 17 and 18. Genet Mol Biol 31:898-901.

Paixão DM, Carneiro PLS, Paiva SR, Sousa KRS, Verardo LL, Braccini Neto J, Pinto APG, Hidalgo AM, Nascimento CS, Périssé IV, *et al.* (2012) Mapeamento de QTL nos cromossomos 1,2,3,12, 14,15 e X em suínos: Características de carcaça e qualidade de carne. Arq Bras Med Vet 64:974-982.

Paixão DM, Braccini Neto J, Paiva SR, Carneiro PLS, Pinto APG, Sousa KRS, Nascimento CS, Verardo LL, Hidalgo AM, Lopes PS, *et al.* (2013) Detecção de locos de características

quantitativas nos cromossomos 1, 2, 3, 12, 14, 15 e X de suínos: Características de desempenho. Arq Bras Med Vet 65:213-220.

Park T and Casella G (2008) The Bayesian Lasso. J Am Stat Assoc 103:681-686.

Peixoto JO, Guimarães SEF, Lopes PS, Soares MAM, Pires AV, Barbosa MV, Torres RA and Silva MA (2006) Associations of leptin gene polymorphisms with production traits in pigs. J Anim Breed Genet 123:378-383.

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

Pérez-Enciso M, Clop A, Noguera JL, Ovilo C, Coll A, Folch JM, Dabot D, Estany J, Oliver MA, Diaz I, *et al.* (2000) A QTL on pig chromosome 4 affects fatty acid metabolism: Evidence from an Iberian by Landrace intercross. J Anim Sci 78:2525-2531.

Ponsuksili S, Jonas E, Murani E, Phatsara C, Srikanchai T, Walz C, Schwerin M, Schellander K and Wimmers K (2008) Trait correlated expression combined with expression QTL analysis reveals biological pathways and candidate genes affecting water holding capacity of muscle. BMC Genomics 9:367.

R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, *et al.* (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.

Seaton G, Hernandez J, Grunchec JA, White I, Allen J, De Koning DJ, Wei W, Berry D, Haley C and Knott S (2006) GridQTL: A grid portal for QTL mapping of compute intensive datasets. In: Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil.

Serão NVL, Veroneze R, Ribeiro AMF, Verardo LL, Braccini Neto J, Gasparino E, Campos CF, Lopes PS and Guimarães SEF (2011) Candidate gene expression and intramuscular fat content in pigs. J Anim Breed Genet 128:28-34.

Silva KDM, Paixão DM, Silva PV, Solero BP, Pereira MS, Lopes PS and Guimarães SEF (2008) Mapping of quantitative trait loci and confirmation of the FAT1 region on chromosome 4 in an F2 population of pigs. Genet Mol Biol 31:475-480.

Smith BJ (2007) BOA: An R package for MCMC output convergence. J Stat Softw 21:1-37.

Sousa KRS, Guimarães SEF, Filho MIS, Lopes MS, Gomes APP, Verardo LL, Braccini Neto J and Lopes PS (2011) Mapeamento de locos de características quantitativas nos cromossomos 5, 7 e 8 de suínos. Rev Bras Zootec 40:115-123.

Wimmers K, Murani E, Pas MFWT, Chang KC, Davoli R, Merks JWM, Henne H, Muraniova M, da Costa N, Harlizius B, *et al.* (2007) Associations of functional candidate genes derived from gene-expression profiles of prenatal porcine muscle tissue with meat quality and muscle deposition. Anim Genet 38:474-484.

Yue G, Stratil A, Cepica S, Schroffel J, Schroffelova D, Fontanesi L, Cagnazzo M, Moser G, Bartenschlager H, Reiner G, *et al.* (2003) Linkage and QTL mapping for *Sus scrofa* chromosome 7. J Anim Breed Genet 120(suppl. 1):56-65.

Internet Resources

Pig QTL data base (PigQTLdb)
http://www.animalgenome.org/cgi-bin/QTLdb/SS/index (June 19, 2011).

Appendix

Supplementary information on SNP names and locations are available from authors upon request.

Associate Editor: Alexandre Rodrigues Caetano