ABSTRACT: This study aimed at identifying genetic markers with potential pleiotropic effects for growth, feed intake and backfat in Duroc pigs by performing multivariate and univariate genome-wide association (GWA) analysis. Linear mixed models were used to perform genome-wide association analysis for three traits collected on 1,022 boars with 35,042 SNP genotyped on 18 autosomes. Heritability estimates for average daily gain (ADG), average daily feed intake (ADFI) and ultrasound backfat (BF) were 0.40, 0.63 and 0.56. Estimates for genetic correlation were 0.48, 0.40 and 0.64, for ADG and ADFI, ADG and BF and ADFI and BF, respectively. One SNP ALGA0006684 on chromosome 1, located near gene SOCS6 and DOX6, was significant (1.4×10^{-8}) within three-trait GWA, showing potential pleiotropic effect. This SNP was also significant in single-trait GWA for ADG (2.2×10^{-8}) and BF (2.0×10^{-5}) and was approaching significance for ADFI (4.1×10^{-6}).

Keywords: Duroc pigs; GWAS; Pleiotropic

Introduction

Maximizing profits within the swine industry requires consideration of system inputs and outputs. Pigs that utilize feed more efficiently, grow faster and produce more lean meat are desirable in terms of reducing costs and increasing products (McGlone et al. (2003)). Therefore, simultaneous improvement of these production traits with economic interests becomes a major goal for swine breeders. Genetic improvements for feed efficiency, growth as well as backfat thickness in pigs have been proven to be successful in certain breeding programs (Rothschild and Ruvinsky (2010)), however, knowledge of correlations for these traits on a molecular level remains unknown.

Genomic improvement of feed efficiency in swine relies on identification of genomic variants associated with one or more traits. The availability of the Porcine60K BeadChip has greatly facilitated whole-genome association studies; contributes to successfully identified genomic variants with major effects on a particular trait (Sahana et al. (2013); Do et al. (2014)). However, the majority of the SNPs reported to be associated with correlated traits were identified from genome-wide association analysis of each single trait, separately. Less attention has been devoted to the fact that correlated traits may be evaluated simultaneously to identify genetic variants with potential pleiotropic effects.

An effective strategy to tackle the challenge of detecting pleiotropic genes is to analyze correlated traits simultaneously via a multivariate GWA approach (Korte et al. (2012)). This study aimed at identifying genetic markers with potential pleiotropic effects for growth, feed intake and backfat in Duroc pigs by performing multiple-trait and single-trait genome-wide association (GWA) analysis.

Materials and Methods

Animal and phenotype data. Data used for this study were provided by Smithfield Premium Genetics (Rose Hill, NC). This data included 1022 Duroc boars within a single nucleus farm. Individual feed intake and serial body weight records of the boars obtained from FIRE (Feed Intake Recording Equipment, Osborne Industries, Inc., Osborne, KS, USA) were edited to compute growth and feed intake. Feed intake data were recorded for boars, starting at an average age of 85 days and ending after approximately 45 days. Detail of feed intake and growth data editing can be found in Jiao et al. (2014). The ultrasound images of all animals were captured over the last three ribs via an Aloka 500 machine (Corometrics Medical Systems, Wallingford, CT).

Genotype data. Genomic data from 1022 boars were obtained using the Illumina PorcineSNP60K BeadChip genotyping. The SNP with call rates ≤ 0.90, MAF (minor allele frequency) ≤ 0.002, and p-value < 0.0001 from a chi-square test for Hardy-Weinberg equilibrium was excluded from the genotype dataset. Additionally, individuals with call rate ≤ 0.90 were also excluded. Missing SNP genotypes were imputed for all available boars using Alphalmpute v1.1 (Hickey et al., 2011). After quality control 35,042 SNP on 1 to 18 autosomes for 1,022 boars remained in the dataset.

Statistical analyses. Phenotypes of the 1,022 boars were pre-adjusted for corresponding fixed effects and covariates before genome-wide association analysis. Fixed effects for ADFI included batch (contemporary effect) and on test body weight. Batch and on-test age were included as fixed effect and covariate for ADG. Fixed effect of batch and linear covariate on-test body weight were pre-adjusted for BF.

A multivariate linear mixed model (similar to GBLUP) was used to perform the three-trait GWA analysis:
\[ Y = 1\beta_0 + x\beta + U + E \]

where \(Y\) is a \(n\) by \(d\) matrix of pre-adjusted phenotypes (\(d = 3, n = 1022\)); \(\beta_0\) is a scalar for general mean; \(x\) is a \(n\) by 1 vector of genotypes for a particular SNP (0 for AA, 1 for Aa and 2 for aa); \(\beta\) is a 1 by \(d\) vector for SNP effect for the \(d\) phenotypes; \(U\) is a \(n\) by \(d\) matrix of random effects with \(U \sim MVN(0, G\sigma_u^2)\); \(V_\beta\) is a \(d\) by \(d\) matrix of genetic variance component and \(G\) is the \(n\) by \(n\) genomic relationship matrix; \(E\) is a \(n\) by \(d\) matrix of random error with \(E \sim MVN(0, I\sigma_e^2)\); \(V_e\) is a \(d\) by \(d\) matrix of residual variance component and \(I\) is an identity matrix. The GEMMA software (Zhou et al. (2012)) was used for the multivariate (three-trait) GWA analysis and the \(p\)-value from log likelihood ratio test (LRT) were obtained to test the global \(H_0: \beta = 0\).

The univariate GWA model is similar to the above model except with reduced dimension.

Significance of SNP was declared with \(p\) value less than the genome-wise type I error rate, adjusted with the Bonferroni correction (Devlin et al. (1999)) by using \(\alpha/K\), where \(\alpha = 0.05\) and \(K = 35042\) (number of SNP). So the genome-wise significance threshold was set to be \(1.427 \times 10^{-6}\).

Results and Discussion

Genetic parameters. Table 1 shows the genetic covariance, heritability and genetic correlation estimates from a three-trait GBLUP model (Null model, without fitting any SNP into the model as covariate). Estimated heritabilities for ADG, ADFI and BF using this model were 0.40, 0.63 and 0.56, respectively in the present study which were similar to the estimates using traditional animal model (Jiao et al. (2014)) and were consistent with estimates of 0.48, 0.49, reported by Hoque et al. (2008; 2009) for ADG and ADFI in Duroc pigs and the estimates (0.54 for BF) published by Lo and collegues (1992) for crossbreds of Landrace and Duroc pigs. Moderate to high genetic correlations between growth and feed intake (0.48), growth and back fat (0.40) as well as positive genetic correlation between feed intake and backfat (0.64) were found in the study, which were slightly higher than estimates reported by Lo et al. (1992) and Hoque et al. (2008).

Table 1. Genetic covariance (above diagonal), heritability (diagonal) and genetic correlation (below diagonal) estimates\(^1\) for average daily gain (ADG), average daily feed intake (ADFI) and ultrasound backfat thickness (BF).

<table>
<thead>
<tr>
<th>Trait</th>
<th>ADG</th>
<th>ADFI</th>
<th>BF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADG</td>
<td>0.400</td>
<td>0.008</td>
<td>0.008</td>
</tr>
<tr>
<td>ADFI</td>
<td>0.480</td>
<td>0.630</td>
<td>0.068</td>
</tr>
<tr>
<td>BF</td>
<td>0.400</td>
<td>0.640</td>
<td>0.560</td>
</tr>
</tbody>
</table>

\(^1\)Standard errors of estimates were not provided by the program GEMMA.

Multivariate and univariate GWA analysis. The Manhattan plots for minus log \(p\)-value of each SNP against its genomic position were shown in Figure 1. Multivariate GWAS for ADG, ADFI and BF was fitted first using a multivariate linear mixed model (Figure 1D). As illustrated in the Manhattan plot (Figure 1D), we can observe that one SNP marker on top of a peak on chromosome 1 exceeding the genome-wise significance level with \(p\)-value \(1.4 \times 10^{-6}\). The test for this SNP rejected the null hypothesis, which may have potential pleiotropic effect for three of the traits. Estimates of the coefficient for this SNP from the three-trait GWA model were 0.018, 0.072 and 0.066 for ADG, ADFI and BF, respectively, indicating positive effects on the three traits.
GWA analysis. Additionally, several SNPs not detected in multivariate GWA analysis were identified significantly associated with the three traits in univariate GWA analysis, representing in Table 2.

![Table 2](Image)

<table>
<thead>
<tr>
<th>SNP_ID</th>
<th>Multi.</th>
<th>Uni.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARC0013872</td>
<td>4.1×10^-6</td>
<td>4.3×10^-5</td>
</tr>
<tr>
<td>ALGA0006684</td>
<td>1.4×10^-6</td>
<td>2.2×10^-7</td>
</tr>
<tr>
<td>H3GA0003149</td>
<td>4.2×10^-4</td>
<td>1.3×10^-6</td>
</tr>
<tr>
<td>ALGA0006690</td>
<td>4.2×10^-4</td>
<td>1.3×10^-6</td>
</tr>
<tr>
<td>H3GA0011666</td>
<td>2.0×10^-5</td>
<td>1.0×10^-6</td>
</tr>
<tr>
<td>MARC0104059</td>
<td>1.1×10^-3</td>
<td>9.3×10^-2</td>
</tr>
<tr>
<td>INRA0055214</td>
<td>1.7×10^-3</td>
<td>1.4×10^-6</td>
</tr>
</tbody>
</table>

P-value for log likelihood test statistic (LRT) for significant SNP was shown in Table 2, with p-value equal or less than genome-wise threshold (1.4×10^-6) bolded.

Several bivariate analysis for GWA analysis were reported in human and livestock (Korte et al. (2012); Han et al. (2012); Serao et al. (2013)), reported potential gains of using multivariate GWA model to fit correlated traits simultaneously for testing of potential pleiotropic effects. However, multivariate analyses has its own challenges that rejecting the null hypothesis from multivariate analysis does not immediately indicate which phenotypes are driving the association (Zhou et al. (2012)). Thus, after observing a significant SNP in multivariate analysis for ADG, ADFI and BF, we also fitted univariate GWA model for each trait separately to examine whether the SNP would remain significant within a univariate model. From the above analysis, the SNP identified in Multivariate analysis was consistent with results from univariate GWA analyses. This SNP ALGA0006684 on chromosome 1 locates near gene SOCS6 and DOX6, which were involved in insulin receptor signaling (Gupta et al. (2011); Kurotsuchi et al. (2010)) and were previously identified (Jiao et al., 2014).

Conclusion

Moderate or high genetic correlation estimates among ADG, ADFI and BF suggest there might be some genomic variants with pleiotropic effects for these traits. To test this hypothesis, we performed multivariate and univariate GWA analyses in which we identified a SNP ALGA0006684 on chromosome 1, exceeding genome-wise significant threshold, implying potential pleiotropic effects for growth, feed intake and ultrasound backfat thickness. Further investigation of this SNP is needed to validate the effectiveness of this SNP as a genetic predictor for feed efficiency, growth and backfat in a larger population.

Literature Cited

Han, Y., Pei, Y., Liu, Y. et al. (2012). Bone 51:1000-1007.
Rothschild, M. F., and Ruvinsky, A. (2010). The genetics of the pig. 2nd ed. CABI, Cambridge, MA.