Identification of quantitative trait transcripts for growth traits in the large scales of liver and muscle samples

Xinwei Xiong, Hui Yang, Bin Yang, Congying Chen, and Lusheng Huang

Key Laboratory for Animal Biotechnology of Jiangxi Province and the Ministry of Agriculture of China, Jiangxi Agricultural University, Nanchang, China

Submitted 20 January 2015; accepted in final form 21 April 2015

GROWTH-RELATED TRAITS ARE economically important traits to the pig industry. More and more studies have focused on the genetic improvement of pig growth. Furthermore, considering that pigs possess greater similarity with humans in nutritional and metabolic physiology than other animal models (32), dissecting the genetic architecture of porcine growth traits not only benefits the pig industry but also may provide improved understanding of human growth retardation. However, it is difficult to uncover the genetic background of growth-related traits because these traits are complex and influenced by multiple interacting factors, including diet, age, sex, and environmental factors.

The heritability of growth traits in pigs is estimated at ~0.2 (41). Since Andersson et al. (3) reported the first genome-wide scan of quantitative trait loci (QTLs) for growth traits in a wild boar × large white intercross, several QTLs related to pig growth have been identified on Sus scrofa chromosomes (SSC) 1, 2, 4, 6, 7, and X (6, 27, 29, 39, 45). More recently, several positional candidate genes corresponding to growth traits have been investigated, such as high mobility group AT-hook 1 (HMGA1) (22) and melanocortin 4 receptor (MC4R) (10). However, to our knowledge, except for the IGF2 gene (44), no other causative gene for growth traits has been identified in pigs.

The confidence intervals of the most of QTLs for growth traits are still too large to identify the causative genes and mutations. In all cases, the road to identification of the quantitative trait nucleotides remains formidable. In recent years, with the development of high-throughput sequencing and high-density single nucleotide polymorphism (SNP) genotyping platforms, system biology has been confirmed to be an efficient approach for identifying causative genes and mutations by integrative analysis of genome-wide association studies (GWAS), gene expression QTL (eQTL), and gene coexpression network. Schadt et al. (40) and Zhu et al. (48) have indicated that an integrative genomic approach can facilitate the identification of causative genes leading to human diseases. Ponsuksili et al. (37) identified 663 genes with fatness-associated expression in porcine liver and mapped their eQTL. More recently, Wimmer et al. (47) characterized AHNAK, SLC3A2, and MAP4K4 as candidate genes for meat drop loss by integrating data of gene expression, eQTL, and phenotypic QTL.

In this study, we used tag-based RNA sequencing to analyze expression profiles of whole genome transcripts in two porcine tissues, liver and muscle, a metabolically active tissue that is critical to pig growth and development and a main organ for meat production. More importantly, the large sample size allowed us to accurately evaluate the correlation between gene expression and growth traits (quantitative trait transcript, QTT). Moreover, whole genome gene expression and QTT analyses facilitated the identification of candidate genes for growth traits. This study provides useful information on the genetic architecture of pig growth.

MATERIALS AND METHODS

Experimental population and phenotype measurements. The experiment in this study included a large-scale population of White Duroc × Erhualian F2 intercross. This F2 population was constructed as described previously (17). In brief, two White Duroc boars were mated to 17 Erhualian sows to produce F1 animals. Nine F1 boars and 59 F1 sows were then intercrossed to produce a total of 1,912 F2 animals in six batches. The F2 animals were weaned at 46 days of age, and the males were castrated at 90 days of age. The fattening pigs were then housed under consistent indoor conditions at the experimental farm of Jiangxi Agricultural University and slaughtered at 240 ± 3 days after an overnight fast with free access to water (~12 h). All procedures involving animals followed the guidelines for the care and use of experimental animals approved by the State Council of the People’s Republic of China. The ethics committee of Jiangxi Agriculture University specially approved this study. In this study, a total of nine...
growth-related traits were phenotyped, including body weight at birth (BodyWt_D0) and at d21 (BodyWt_D21), 46 (BodyWt_D46), 120 (BodyWt_D120), 210 (BodyWt_D210) and 240 (BodyWt_D240) and average daily gain (ADG_D0-120, ADG_D120-210, and ADG_D210-240) by the methods described previously (2).

RNA extraction and whole genome gene expression analysis. We harvested 497 liver samples and 586 longissimus dorsi muscle samples from F2 animals and used them to extract total RNA with Trizol (Invitrogen) following the manufacturer’s manual. RNA quantity and integrity were assessed by using a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific) and a 2100 Bioanalyzer (Agilent). Expression profiles of whole genome transcripts were assayed by digital gene expression analysis (34). In brief, mRNA was isolated from total RNA with the magnetic oligo (dT) beads (Invitrogen). Double-stranded cDNA was synthesized with oligo (dT) primers and then digested with NlaIII and MmeI enzyme (NEB). The digested cDNA was ligated to specific adapters.Polymerase chain reaction was performed to enrich cDNA library. After purification and denaturation, the single chain cDNA library was sequenced on a GAII sequencer (Illumina). Data processing was conducted by the method described previously (8, 9). In brief, we first constructed the reference transcript set by downloading the transcripts from the database of PEDE (Pig Expression Data Explorer; http://pede.dnaaffrc.go.jp/) and pig unigene in the National Center for Biotechnology Information (ftp://ftp.ncbi.nih.gov/repository/UniGene/Sus_scrofa/). The redundant transcripts overlapping between the two databases were removed from the reference transcript set. The positions of the reference transcripts in reference genome assembly 10.2 were determined with Bowtie (24). For monitoring the mapping events on both sense and antisense strands, a virtual sense and antisense tag sequence database (all possible reference tags, 17 bp sequences next to an NlaIII restriction site) was generated for both full gene and cDNA sequences using in-house Perl scripts. The raw tags were first filtered to produce clean tag data. The clean tag sequences were then BLASTed with the reference tags using SOAP2 (26), allowing up to one mismatch in 21 bp tag sequences. The number of clean tags that were uniquely mapped to the reference transcript sequence was calculated and then normalized to TPM (number of tags mapped to each gene per million clean tags) as the expression level of each transcript.

Statistical analyses. The transcripts expressed in >20% of samples were used for subsequent analysis. The gene expression profiles and phenotypic data in F2 animals were further adjusted for sex, batch, and kinship with a robust linear regression model by using the polygenic function of GenABEL package in R software (28). Sex and batch were considered as fixed effects; the polygenic effect among samples was accounted for with IBS-matrix in the GenABEL package with genocontrol, a total of 20,108 transcripts in liver and 23,728 transcripts in muscle were used for further QTT analysis. We used a regression model to evaluate the association between gene expression level and phenotypic value of growth traits at the significance threshold of P < 0.0005. The results are presented in Table 1, Table 2, and Supplemental Table S1.1

Table 1. Candidate genes identified in liver for growth traits

<table>
<thead>
<tr>
<th>Trait</th>
<th>QTT</th>
<th>Position, bp</th>
<th>Annotated Gene</th>
<th>r Value</th>
<th>P Value</th>
<th>Phenotype in Knockout Mice</th>
</tr>
</thead>
<tbody>
<tr>
<td>BodyWt_D21</td>
<td>BKKL1_0059_D0</td>
<td>SSC1:100972507-10106424</td>
<td>MYO6</td>
<td>0.191</td>
<td>2.12E-05</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D21</td>
<td>gnt UG Scc#S31133723</td>
<td>SSC14:134677755-143680316</td>
<td>ZFP36</td>
<td>0.159</td>
<td>4.25E-04</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D21</td>
<td>PBL01_0007_D0</td>
<td>SSC18:6629467-6635745</td>
<td>GIMAP6</td>
<td>−0.176</td>
<td>9.47E-04</td>
<td>abnormal body weight</td>
</tr>
<tr>
<td>BodyWt_D46</td>
<td>BKKL1_0059_D0</td>
<td>Scc:100972507-10106424</td>
<td>MYO6</td>
<td>0.190</td>
<td>2.27E-05</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D46</td>
<td>gnt UG Scc#S32776821</td>
<td>SSCX:139064924-139065816</td>
<td>IDS</td>
<td>0.158</td>
<td>4.51E-04</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D46</td>
<td>gnt UG Scc#S20116789</td>
<td>GL594574.1:1629-22640</td>
<td>SERPINE1</td>
<td>0.166</td>
<td>2.33E-04</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D210</td>
<td>UTC01_0022_H02</td>
<td>GL5899280.1:6143-16328</td>
<td>PDXK</td>
<td>0.199</td>
<td>2.39E-04</td>
<td>increased body weight</td>
</tr>
<tr>
<td>ADG_D120-210</td>
<td>gnt UG Scc#S40444212</td>
<td>SSC7:10929836-109308540</td>
<td>DIO2</td>
<td>−0.352</td>
<td>6.25E-06</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>ADG_D120-240</td>
<td>gnt UG Scc#S35614999</td>
<td>SCC12:5489685-54863767</td>
<td>ADVAL</td>
<td>−0.209</td>
<td>1.25E-04</td>
<td>increased body weight</td>
</tr>
<tr>
<td>ADG_D210-240</td>
<td>LVRM1_0096_D0</td>
<td>SSC15:6342620-63437375</td>
<td>PPP1R3B</td>
<td>0.194</td>
<td>3.80E-04</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>ADG_D210-240</td>
<td>SSKN1_0041_G08</td>
<td>SSC5:18846917-18857897</td>
<td>RARG</td>
<td>0.191</td>
<td>4.66E-04</td>
<td>decreased body weight</td>
</tr>
</tbody>
</table>

1Chromosomal location of quantitative trait transcript (QTT) according to Sus Scrofa Build 10.2 assembly; 2annotated gene of the QTT; 3The coefficient between gene expression and phenotype value; 4Phenotype in knockout mice were found in MGI (http://www.informatics.jax.org/).
associated with more than one trait. For examples, the expression of transcript gnl|UG|Ssc#S35166996 was significantly correlated with BodyWt_D210 ($P = 4.71E-05$, $r = 0.220$) and BodyWt_D240 ($P = 4.04E-04$, $r = -0.160$); the transcript PBL01_0086_D08 was associated with BodyWt_D21 ($P = 1.09E-04$, $r = -0.174$), BodyWt_D210 ($P = 1.67E-04$, $r = -0.203$), and BodyWt_D240 ($P = 1.98E-04$, $r = -0.168$); the transcript gnl|UG|Ssc#S6091219 was correlated with BodyWt_D21 ($P = 1.09E-05$, $r = 0.174$), BodyWt_D210 ($P = 3.57E-04$, $r = 0.203$), and BodyWt_D240 ($P = 3.24E-04$, $r = 0.168$).

Table 2. Candidate genes identified in muscle for growth traits

<table>
<thead>
<tr>
<th>Trait</th>
<th>QTT</th>
<th>Position, bp</th>
<th>Annotated Gene</th>
<th>$r$ Value</th>
<th>$P$ Value</th>
<th>Phenotype in Knockout Mice</th>
</tr>
</thead>
<tbody>
<tr>
<td>BodyWt_D0</td>
<td>gnl UG Ssc#S50196300</td>
<td>SSC18: 20855084-20884021</td>
<td>FLNB</td>
<td>0.162</td>
<td>8.92E-05</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D21</td>
<td>SMG01_0071_A04</td>
<td>SSC2: 6867951-6878759</td>
<td>ESRRRA</td>
<td>0.182</td>
<td>1.01E-05</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D24</td>
<td>SMG01_0071_A04</td>
<td>SSC2: 6867951-6878759</td>
<td>ESRRRA</td>
<td>0.149</td>
<td>3.25E-04</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D46</td>
<td>gnl UG Ssc#S23768253</td>
<td>SSC6: 15060487-15160499</td>
<td>ZFHX3</td>
<td>0.168</td>
<td>5.00E-05</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D46</td>
<td>gnl UG Ssc#S35172260</td>
<td>SSC6: 145323643-145450640</td>
<td>USP24</td>
<td>0.148</td>
<td>3.57E-04</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D210</td>
<td>gnl UG Ssc#S34178450</td>
<td>SSC12: 51468686-51493212</td>
<td>ASPA</td>
<td>0.195</td>
<td>8.18E-05</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D210</td>
<td>gnl UG Ssc#S35173098</td>
<td>SSC9: 79239885-79480047</td>
<td>CDK6</td>
<td>−0.192</td>
<td>9.96E-05</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D210</td>
<td>gnl UG Ssc#S18377808</td>
<td>SSC1: 14630493-146312662</td>
<td>BUB1B</td>
<td>−0.173</td>
<td>4.72E-04</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D240</td>
<td>gnl UG Ssc#S43178450</td>
<td>SSC12: 51468686-51493212</td>
<td>ASPA</td>
<td>0.153</td>
<td>2.33E-04</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D240</td>
<td>gnl UG Ssc#S39835694</td>
<td>SSC7: 5162417-5316551</td>
<td>BMP6</td>
<td>0.149</td>
<td>3.24E-04</td>
<td>decreased body weight</td>
</tr>
<tr>
<td>BodyWt_D240</td>
<td>gnl UG Ssc#S46877694</td>
<td>GL894492.1: 7767-21848</td>
<td>ETS1</td>
<td>0.148</td>
<td>3.45E-04</td>
<td>decreased body weight</td>
</tr>
</tbody>
</table>

For definitions, see Table 1 footnotes.

Fig. 1. Functional annotation of quantitative trait transcripts (QTTs) for growth traits in liver and muscle by gene ontology (GO) analysis. The bar plot represents the gene counts within each GO category. All function or process terms listed have enrichment of corrected $P$ values $< 0.05$. 

A: QTTs for growth traits in liver. 
B: QTTs for growth traits in muscle.
Mutations within the FLNB gene are implicated in a variety of genetic disorders characterized by skeletal malformation, including spondyloepiphysyeal dysplasia tarda (12, 15), Larsen syndrome (4, 23), atelosteogenesis types I and III (11, 23), and boomerang dysplasia (5). In this study, the expression level of the FLNB gene in muscle was correlated with body weight at birth ($P = 8.92E-05$, $r = 0.162$). The ASPA gene has been reported to be associated with growth in mice and humans. Adipocyte-specific genetic deletion of ASPA gene in mice reduces body weight (30, 31, 43). In the present study, the expression of ASPA gene was significantly associated with BodyWt_D210 ($P = 8.18E-05$, $r = 0.195$) and BodyWt_D240 ($P = 2.33E-04$, $r = 0.153$). Fontanesi et al. (13) identified two SNPs within the IGSF3 gene that were significantly associated with average daily gain through GWAS and candidate gene analysis. In this study, the expression level of the IGSF3 gene was negatively associated with body weight at birth ($P = 6.74E-06$, $r = -0.186$).

**Comparative analysis of QTTs identified in liver and muscle.** The expression profiles of whole genome transcripts were obtained in both liver and muscle. This allowed us to perform comparative analysis of QTTs between two tissues. In fact, we did not identify any QTTs that were significantly associated with the same growth trait in both tissues, suggesting the tissue specificity of QTTs identified in this study. For each growth trait, a different pattern of QTT numbers was identified for liver and muscle (Fig. 2). For example, significantly more QTTs were identified in muscle than in liver for BodyWt_D120, suggesting that genes regulating the muscle physiology may play more important roles in pig growth at day 120. The same condition was also identified for BodyWt_D0 and ADG_D0-120. However, for BodyWt_D21, BodyWt_D46, BodyWt_D210, and ADG_D120-210, significantly more QTTs were detected in liver. This result suggests that growth at different ages stems from different systems/tissues at variable shares and that the proper tissue should be selected when gene expression analysis is performed.

**Identification of SDR16C5 as a candidate gene for body weight at days 240 by integrative genomic approach.** In our previous study, we identified 15 genome-wide significant SNPs...
on nine chromosomal regions that were significantly associated with six growth-related traits in the White Duroc × Erhualian F₂ intercross, including 10 SNPs at the 1% genome-wide significance level and five at the 5% genome-wide significance level (2). In this study, we used an integrative genomic approach to identify candidate genes related to porcine growth. For all QTLs identified previously, we first searched for the QTTs identified in this study and located within 2.5 Mb region around the strongest trait-associated SNPs (high linkage disequilibrium existed in the F₂ population, so 2.5 Mb was used). As a result, except for the QTL on SSC4, we did not identify any QTTs that were located within QTL regions on other chromosomes. We focused on the QTL on SSC4 for BodyWt_D240. The genome-wide significant SNPs were located within SSC4: 82.25–83.03 Mb (2). Interestingly, we identified the gnl|UG|Ssc#S46879708 that was exactly the QTT for BodyWt_D240 in the muscle samples ($r = 0.157, P = 1.47E-04$). This transcript was annotated to the short chain dehydrogenase/reductase family 16C, member 5 (SDR16C5) gene and located at SSC4: 82.51–82.52 Mb. The SDR16C5 gene has been reported to influence human height (16, 25, 46). To determine whether a cis-regulator existed for the SDR16C5 gene in the QTL region, we further analyzed the associations of the SNPs within QTL region with expression level of the SDR16C5 gene (cis-eQTL, $P = 9.09E-03$). These results suggested that SDR16C5 gene is the candidate for BodyWt_D240. We further retrieved the interacting genes of SDR16C5 through STRING 9.1 (http://string-db.org/). The result is shown in Fig. 3B. The SDR16C5 gene was the node for interaction network of 10 genes.

Fig. 3. The SDR16C5 gene was identified as a candidate gene for porcine growth trait by an integrative genomic approach. A: the positions of trait-associated single nucleotide polymorphisms (SNPs) for BodyWt_D240 and SDR16C5 gene on Sus scrofa chromosome (SSC)4. The x-axis indicates the chromosomal locations of SNPs and candidate gene; the y-axis on the left shows the $-\log_{10}(P$ value) of SNPs obtained from genome-wide association studies, the y-axis on the right shows the $-\log_{10}(P$ value) of candidate gene obtained from QTT analysis. B: the interaction network of SDR16C5 gene identified through STRING 9.1 (http://string-db.org/).
genes. Interestingly, five out of 10 interacting genes, ACACA, ACACB, CHD7, PENK, and ALDH5A1, are associated with body weight, body height, and body mass index in humans (1, 14, 18, 25). This result provides more evidence in support of SDR16C5 as a candidate growth gene. However, further experiments should be performed to confirm the causality of this gene with pig growth.

Conclusions

In summary, in this study, we identified a total of 169 and 168 QTTs for nine growth-related traits in porcine liver and muscle, respectively. Some of these QTTs have been reported as functional genes affecting growth in humans and mice. Through an integrative genomic approach, we identified SDR16C5 as an important candidate gene for porcine growth. These findings provide new insights into the genetic basis of growth traits in pigs.

ACKNOWLEDGMENTS

We thank BGI-Shenzhen for assistance in tag-based RNA sequencing. We are grateful to colleagues in Key Laboratory for Animal Biotechnology of Jiangxi Province and the Ministry of Agriculture of China, Jiangxi Agricultural University for sample collection.

GRANTS

This project was supported by National High-Tech Research and development program of China (2013AA102502).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: X.X. and B.Y. analyzed data; X.X. prepared figures; X.X. drafted manuscript; H.Y. performed experiments; C.C. and L.H. conception and design of research; C.C. edited and revised manuscript; L.H. approved final version of manuscript.

REFERENCES
