Genome-wide Association Study of Integrated Meat Quality-related Traits of the Duroc Pig Breed

Taechon Lee, Dong-Hyun Shin, Seoaee Cho1, Hyun Sung Kang2, Sung Hoon Kim3, Hak-Kyo Lee3, Heebal Kim4, and Kang-Seok Seo5*

Department of Agricultural Biotechnology, Animal Biotechnology Major, Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Korea

ABSTRACT: The increasing importance of meat quality has implications for animal breeding programs. Research has revealed much about the genetic background of pigs, and many studies have revealed the importance of various genetic factors. Since meat quality is a complex trait which is affected by many factors, consideration of the overall phenotype is very useful to study meat quality. For integrating the phenotypes, we used principle component analysis (PCA). The significant SNPs refer to results of the GRAMMAR method against PC1, PC2 and PC3 of 14 meat quality traits of 181 Duroc pigs. The Genome-wide association study (GWAS) found 26 potential SNPs affecting various meat quality traits. The loci identified are located in or near 23 genes. The SNPs associated with meat quality are in or near five genes (ANKI, BMP6, SHH, PIP4K2A, and FOXN2) and have been reported previously. Twenty-five of the significant SNPs also located in meat quality-related QTL regions, these result supported the QTL effect indirectly. Each single gene typically affects multiple traits. Therefore, it is a useful approach to use integrated traits for the various traits at the same time. This innovative approach using integrated traits could be applied on other GWAS of complex-trait including meat-quality, and the results will contribute to improving meat-quality of pork. (Key Words: Genome-wide Association Study, Principle Component Analysis, Meat Quality, Duroc Pig)

INTRODUCTION

As consumers seek tastier, healthier, safer and more nutritious pork, the quality of the meat has been improved. As pork quality is influenced by many genetic and non-genetic factors, many studies have been focused on revealing the genetic background and its various genetic factors (Røsvold and Andersen 2003; Hu et al., 2013). These studies have shown that selective breeding of pigs and using DNA markers could play important roles in enhancing pork quality. Meat quality traits are also affected by many genes that are mapped in the quantitative trait loci (QTL) regions. Therefore, a great deal of research related to farm animals has focused on mapping and characterization of trait loci, and genome research (Andersson 2001; Andersson and Georges, 2004). Genetic variants as a part of genetic factors are also used to identify contribution to complex traits. Even within breeds, there is considerable genetic variation in relation to meat quality traits such as pH and intramuscular fat. Part of this variation is due to genetic differences. It therefore follows that DNA marker technology is important in improving meat quality (De Vries et al., 2000).

Genome-wide association study (GWAS) uses genetic variants with interesting traits. Since meat quality is a complex trait affected by many factors, consideration of the
several traits is vital. Although there have been many GWAS for pork quality traits, there has been few studies of multiple traits of meat quality (Luo et al., 2012; Ma et al., 2013). The few studies of multiple traits, used several traits, but each trait was analyzed separately then the results were considered together. Most research has been performed on single traits, where we performed a multiple trait analysis to find candidate genes of good meat quality by integrating the phenotypes. The overall represented phenotype related to meat quality obtained by principal component analysis (PCA) in this study is defined as the integrated phenotype. The integrated phenotypes are representative of overall meat-quality phenotypes and are non-redundant information. However, some unique information obtained from several phenotypes may be important, since the phenotypes are correlated and partially redundant. Therefore a study of the common genetic associations among these phenotypes may provide a more complete understanding of the mechanisms involved in the production of good-quality meat. PCA can extract non-redundant information among the correlated traits with the goal being to extract as much variance with the fewest components. PCA makes new variables, principal components (PCs), to describe the data in a more concise or convenient way. PCA explains most of the variance of the data by extracting PCs from a specific linear combination of the original phenotypes and reducing the dimensionality of a data set which are correlated. An earlier study using PCA to integrate phenotypes successfully found the variants associated with osteoporosis (Karask et al., 2012).

To detect candidate genetic variants associated with various meat quality traits, we considered meat quality in relation to all phenotypes together, and reduced the dimensions of these phenotypes by PCA in the the Duroc breed, which has been identified as a superior genetic source for improving the eating qualities of pork according to the recent National Pork Producers’ Council (NPPC) Terminal Sire-Line Genetic Evaluation Program (Goodwin and Burroughs, 1995).

MATERIALS AND METHODS

Samples and phenotypes

In this study, we used a two-generation resource population of Duroc pigs. Duroc boars were mated with 106 Duroc sows. The 181 female Duroc pigs from four farms were humanely killed at 184±16.3 (mean±SD) days in 10 batches (slaughter groups). The genomic DNA samples were genotyped using the Illumina Porcine 60K SNP BeadChip (Illumina Inc., USA). After slaughtering, meat quality traits of the longissimus muscle were evaluated. All muscle samples were taken from the same position of longissimus dorsi, between 10th and 11th rib, from one slaughter batch. The information of meat quality measurement is described in Supplementary Table S1, S2. We measured 14 meat quality traits, including meat pH 24 h post-mortem (pH[24]), meat color (CIE-L*, CIE-a*, CIE-b*), shear force (SF), drip loss, heat loss, fat (intramuscular fat content), water-holding capacity (WHC), back-fat thickness (BF), protein, cholesterol, moisture and ash contents. PCA was conducted for these 14 phenotypes using R, and correlations between the phenotypes were calculated with the corplot R package (Friendly, 2002).

GRAMMAR: step 1

We performed PCA analysis and got three integrated phenotypes (PCs) related to meat quality, which were then used to estimate meat quality. Residual in Genome-wide rapid association using mixed model and regression (GRAMMAR) (Aulchenko et al., 2007) was used as a phenotype that reflected the genetic factors.

\[ Y = Wv + Za + e, \]

where \( Y \) = the vector of each observations (3 PCs), \( v \) = the vector of fixed effect as age on the slaughter day, \( a \) = the vector of the random effect as relationship matrix and \( e \) = the vector of residual error. \( W \) and \( Z \) are design matrices for \( Y \) and \( a \), respectively.

Observations comprised a total of three PCs of 181 pigs. As all 181 pigs were female, we only used age as a covariate. The relationship matrix \( Z \), based on pedigree information, included the parents of the 181 pigs. Residual \( e \) contains the environmental effect along with a major part of the genetic effect. All parameters were estimated using the ASREML program (Gilmour et al., 2009) facilitated by the AIREML algorithm for a single-trait animal model. Using this model, we calculated the residual for the phenotype related to meat quality used in this GWAS.

GRAMMAR: step 2

The Illumina Porcine 60K SNP BeadChip contains 62,163 SNPs that are distributed across the 18 pig autosomes and the X chromosome. We excluded all SNPs with a miss rate of >5%, a minor allele frequency (MAF) of <0.05, and a Hardy-Weinberg equilibrium (HWE) test p-value of <10^{-6}. After quality control, we retained 45,620 SNPs for further analysis. Association analysis of the SNPs with a residual for PC1, PC2, and PC3 were conducted on the basis of a linear regression using PLINK 1.07 (Purcell et al., 2007). The SNPs were mapped on the Sus scrofa genome 10.2 assembly, version 2 from NCBI FTP, using BLAT (Kent, 2002).

Genome-wide significance was defined based on the LD-adjusted Bonferroni method (Duggal et al., 2008) and
simple Bonferroni method to correct p-value thresholds of significance: significant association of 0.05 false positives per GWAS was used as a genome-wide significance. Duroc pigs are estimated to have 6,948 “independent” tests based on the Gabriel-type block (Gabriel et al., 2002), as calculated by Haploview (Barrett et al., 2005). Therefore, 7.2E-06 (0.05/6,948) has genome-wide significance. An overview of the p-values of tests using Manhattan plots was produced by R. Significant SNPs were mapped on to the meat-quality QTL regions of the pig genome, assembly version 10.2 (Hu et al., 2013).

RESULTS

PCA of meat-quality related phenotypes

Summary of phenotypes measurement are described in Supplementary Table S1, S2. There was a correlation between phenotypes (Figure 1a), which shows the results of a PCA of the meat quality-related phenotypes. The three PCA components jointly comprised 52.3% of the total variability of meat quality-related traits (Supplementary Figure S1). PC1, which accounted for 21.9% of the total variance, includes: pH, moisture, fat content, and meat color components. PC1 weights were very similar in these factors (Figure 1b). PC2, which accounted for 15.5% of the variance, includes SF, BF and protein components (Figure 1b). PC3, which accounted for 14.8% of the total variance, includes: moisture, drip loss, and heat loss components (Figure 1b). To fully consider pork quality, we used PC1, PC2, and PC3, because these components explained a large proportion of the total variability in meat quality-related traits.

SNP-association studies for integrated phenotypes

Using the GRAMMAR method (Aulchenko et al., 2007), we identified the top 26 and 11 SNPs associated with integrated meat quality in the Duroc breed at the significance level of 7.2E-06, which equals an LD-adjusted Bonferroni p-value of 0.05 and a simple Bonferroni significance level of 1.096E-06 (Table 1). The inflation factor λ for PC1, PC2 and PC3 of pig were 1.646, 1.691, and 1.173, respectively (Supplementary Figure S2). These were corrected results for structure/familial effects by using the G matrix. The number of significant SNPs in the genic regions was roughly equal to those in the inter-genic regions. Fourteen significant SNPs were associated with PC1 (six in genic regions, eight in intergenic regions), eleven SNPs were associated with PC2 (six in genic regions, five in intergenic regions), and two SNPs were associated with PC3 (one in a genic region, one in an intergenic region). The coverage of genic and intergenic regions were roughly 1:3 (25.8 to 74.2%; 637 to 1,838Mb). When considering the coverage, the number of significant SNPs in genic regions was significantly greater than that in the
intergenic regions. This may suggest that intergenic regions, although containing important regulatory sequences, have the smallest ratio of functional to total DNA.

Whether the SNPs significantly associated with integrated meat-quality phenotypes were located in meat quality-related QTL regions was determined (Figure 2 and Table 1). Twenty-five of the 26 SNPs were located in meat quality-related QTL regions, such as back fat thickness, intramuscular fat, protein, shear force, meat color (L*, a*, b*), pH, heat loss, and drip loss (Supplementary Table S3) (Hu et al., 2013). The exceptions were MARCO0070937 (ENSSSCG00000026984, SSCX: 44930506) and MARCO0096007 (near SHH, SSC18: 2334006). Many of the significant SNPs were in high linkage disequilibrium (LD) with each other, suggesting that they are inherited as a group (Figure 3, Supplementary Figure S3-10).

**DISCUSSION**

By applying PCA to the correlation-coefficient matrix of related phenotypes, we were able to extract essential and non-redundant information. Multiple trait analysis is a more useful method of locating candidate variants than multiple single trait analysis in the context of systems genetics, since a summary of several correlated traits provides a comprehensive view of the genetics of meat quality. Thus, we found the locus which influences multiple correlated traits by applying PCA on multiple traits. We approximated
Figure 2. Manhattan plots of GWAS results for the traits and QTL mapping. GWAS for the integrated phenotypes using illumina Porcine 60K SNP BeadChip of 181 Duroc samples. Each panel A), B) and C) show GWAS results against residual of PC1, PC2, and PC3. The x-axis of the Manhattan plot shows the genomic position, the y-axis represents the log10 base transformed p-values, LD-adjusted Bonferroni significance levels (7.2E-06) was applied. Meat-quality-related QTL regions located in significant SNPs are shown.

Figure 3. Example result of linkage disequilibrium around the significant SNPs of eigenvector 1 PCA results on chromosome 2. Linkage disequilibrium (LD) was determined using Haploview, and LD blocks were defined using the Gabriel rule. Darker boxes represent higher LD percentages. Pairwise D’ values (’100) are indicated; blue blocks indicate D’ = 0.
the effective number of independent SNPs across the genome, and applied a Bonferroni LD-adjusted significance (7.2E-06). We corrected our figure to produce an accurate number of tests to determine the regions of the genome that should be studied further. The inflation factors seemed to be high, but we corrected for the population stratification in this analysis using the GRAMMAR approach. Yang et al. 2011 showed that the magnitude of genomic inflation factors depends on sample size, heritability, linkage disequilibrium structure and the number of causal variants under polygenic inheritance even when there is no population structure.

The 26 SNPs associated with integrated meat-quality phenotypes and the loci identified are located in or near 23 genes (Table 1). Among them, ANK1, BMP6, SHH, PIP4K2A, and FOXN2 are known to be strong candidates for meat quality. It has been suggested that SHH and BMP6 contribute to both juiciness and tenderness of the meat (Fonseca et al., 2003). MARC0096007 (SSC18: 2334006) and ASGA0030822, (SSC7: 5278629) are located near SHH gene and BMP6 gene, respectively. The SHH and BMP6 are members of the hedgehog signaling pathway. The hedgehog family of proteins and, SHH in particular, may be important in slow oxidative (SO) fiber clustering in pigs. Recent evidence suggests that fiber-type proportions play a role in the quality of meat, and in pigs it has been suggested that the SO fibers contribute to juiciness and tenderness (Fonseca et al., 2003).

FOXXN2, ANK1, and PIP4K2A are related to meat quality. INRA0052583 (SSC17: 12436823) and ASGA0015307 (SSC3:98152224) are located near ANK1 gene and FOXXN2 gene, respectively. ANK1 gene was first discovered in erythrocytes, but has since also been found in brain and muscle tissues. ANK1 is known to be significantly associated with many meat-quality-related traits (water-binding capacity, drip loss, heat loss, pH, IMF, and shear force) (Wimmers et al., 2007). ANK1 and FOXXN2 are related to type 2 diabetes (Wolfrum et al., 2004; Tabassum et al., 2008; Imamura et al., 2012). SNPs (ASGA0048292 and ASGA0048295, SSC10: 57954471, 57985436) are located within an intron of PIP4K2A, which is an important gene for determining whole-body insulin responsiveness, and mutation of the PIP4K2A gene could increase insulin sensitivity of skeletal muscle without any significant change in adipose tissue (Xu et al., 2012). This leads indirectly to a decrease in transport of glucose into white adipose tissues and reduced adiposity, because the increased insulin-stimulated glucose transport in the muscle leads to a lower basal insulin-secretion requirement. PDE3B, located near SNP of ASGA0010696, is a major regulator of an adrenal steroidogenesis (Tsai et al., 2011) and modulates basal corticosterone synthesis (Tsai and Beavo, 2011). Corticosterone can affect meat quality by changing pre-and post-slaughter muscle metabolism (Gao et al., 2008).

One limitation of this study was the small sample size, which resulted in low power, and we found a small number of significant SNPs. Despite the limitations, this study does have its strengths. Even though this data has low power, significant SNPs were detected and they are related with meat quality. Meat quality is a complex trait that is affected by many factors. Therefore, consideration of the overall phenotype is a useful approach to evaluate meat quality. This study also used many phenotypes, and extracted non-redundant and integrated information from correlated phenotypes. Other studies have analyzed a single phenotype, or even if several phenotypes were considered, they were analyzed separately then considered together. Unlike these other studies (Luo et al., 2012; Ma et al., 2013), we found some SNPs possibly associated with all phenotypes when analyzed together using PCs. The significant SNPs in or near genes seemed to be related to several meat-quality phenotypes.

The 21 genes among our findings are also located in meat quality QTL regions. Most of the significant SNPs are located in LD block. The SNPs tag the block regions which are related with phenotypes and locate within the previously reported QTL regions, and thus the SNPs may contribute these complex traits. A QTL region is generally large, therefore fine mapping is needed. In this context, our findings narrow down the meat quality QTL regions. There are limitations that the traits are highly correlated so that they were possibly influenced by the same gene. Also, the LD blocks are large containing several genes, so we cannot exclude closely linked genes that affect the different traits. However, this innovative approach using integrated phenotypes can be applied on other GWAS of complex-trait including meat-quality.

Traits determining meat-quality are difficult to improve by traditional selection and breeding methods, because heritability is quite low (Davoli and Braglia, 2007), the methods are expensive and some are only possible after the animal has been slaughtered. However, knowledge of the number of genes involved in the qualitative characteristics of meat would provide new information that could be used in breeding strategies intended to improve meat quality.

Meat-quality is determined by various factors, thus, consideration of the various phenotypes with simultaneously restricted QTL regions is a useful way to evaluate meat quality. This may lead to an innovative approach of the GWAS of complex-trait including meat-quality. If the significant SNPs of this study and additional SNPs from other studies using the same method as this study are used in marker-assisted selection, it will contribute to improving the meat and carcass quality in pig in a more efficient method than the traditional pig selection.
ACKNOWLEDGEMENTS

The work was supported by a grant from the Next-Generation BioGreen 21 Program (No. PJ008116, PJ009032), Rural Development Administration, Republic of Korea.

REFERENCES

Andersson, L. 2001. Genetic dissection of phenotypic diversity in farm animals. Nat. Rev. Genet. 2:130-138.

Andersson, L. and M. Georges. 2004. Domestic-animal genomics: deciphering the genetics of complex traits. Nat. Rev. Genet. 5:202-212.

Aulchenko, Y. S., Andersson, L. and Minzhel, I. 2005. Genealogy and LD in a large pedigree-based sample from the Finnish breed: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-265.

Barrett, J., B. Fry, J. Maller, and M. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:262-265.

Davoli, R. and Braglia, S. 2007. Molecular approaches in pig breeding to improve meat quality. Brief. Funct. Genomic. Proteomic. 6:313-321.

De Vries, A., L. Faucitano, A. Sosnicki, and G. Plastow. 2000. The use of gene technology for optimal development of pork meat quality. Food. Chem. 69:397-405.

Duggal, P., E. M. Gillanders, T. N. Holmes, and J. E. Bailey-Wilson. 2008. Establishing an adjusted p-value threshold to control the family-wise type I error in genome wide association studies. BMC Genomics 9:516.

Fonseca, S., I. Wilson, G. Horgan, and C. Maltin. 2003. Slow fiber cluster pattern in pig longissimus thoracis muscle: implications for myogenesis. J. Anim. Sci. 81:973-983.

Friendly, M. 2002. Carcass and meat quality traits in pigs: a review. Meat Sci. 64:219-237.

Tabassum, R., S. Chavali, O. P. Dwivedi, N. Tandon, and D. Bharadwaj. 2008. Genetic variants of FOXA2: risk of type 2 diabetes and effect on metabolic traits in North Indians. J. Hum. Genet. 53:957-965.

Tsai, L.-C. L. and J. A. Beavo. 2011. The roles of cyclic nucleotide phosphodiesterases (PDEs) in steroidogenesis. Curr. Opin. Pharmacol. 11:670-675.

Tsai, L.-C. L., M. Shimizu-Albertine, and J. A. Beavo. 2011. The high-affinity cAMP-specific phosphodiesterase 8B controls steroidogenesis in the mouse adrenal gland. Mol. Pharmacol. 79:639-648.

Wimmers, K., E. Murani, M. Te Pas, K. Chang, R. Davoli, J. Merks, H. Henne, M. Murania, N. Da Costa, and B. Harlizius. 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.

Wolf, C. E., G. Asilmaz, E. Luca, J. M. Friedman, and M. Stoffel. 2004. Foxxa2 regulates lipid metabolism and ketogenesis in the liver during fasting and in diabetes. Nature 432:1027-1032.

Xu, T., W. Huang, X. Zhang, B. Ye, H. Zhou, and S. Hou. 2012. Identification and characterization of genes related to the development of breast muscles in Pekin duck. Mol. Biol. Rep. 39:7647-7655.

Yang, J., W. N. Weedon, S. Purcell, G. Lettre, K. Estrada, C. J. Willer, A. V. Smith, E. Ingelsson, J. R. O’Connell, and M. Mangino. 2011. Genomic inflation factors under polygenic inheritance. Eur. J. Hum. Genet. 19:807-812.