

Identification of genes related to intramuscular fat content of pigs using genome-wide association study

Sohyoung Won¹, Jaehoon Jung¹, Eungwoo Park², and Heebal Kim^{1,3,*}

* **Corresponding Author:** Heebal Kim
Tel: +82-2-880-4803, **Fax:** +82-2-876-8827,
E-mail: heebal@snu.ac.kr

¹ Department of Agricultural Biotechnology and
Research Institute of Population Genomics, Seoul
National University, Seoul 151-742, Korea

² Animal Genomics & Bioinformatics Division, National
Institute of Animal Science, Rural Development
Administration, Wanju 55365, Korea

³ CHO&KIM genomics, Seoul 05836, Korea

ORCID

Sohyoung Won
<https://orcid.org/0000-0003-2039-6017>
Heebal Kim
<https://orcid.org/0000-0003-3064-1303>

Submitted Mar 21, 2017; Revised Apr 18, 2017;
Accepted Jun 24, 2017

Objective: The aim of this study is to identify single nucleotide polymorphisms (SNPs) and genes related to pig IMF and estimate the heritability of intramuscular fat content (IMF).

Methods: Genome-wide association study (GWAS) on 704 inbred Berkshires was performed for IMF. To consider the inbreeding among samples, associations of the SNPs with IMF were tested as random effects in a mixed linear model using the genetic relationship matrix by GEMMA. Significant genes were compared with reported pig IMF quantitative trait loci (QTL) regions and functional classification of the identified genes were also performed. Heritability of IMF was estimated by GCTA tool.

Results: Total 365 SNPs were found to be significant from a cutoff of p-value <0.01 and the 365 significant SNPs were annotated across 120 genes. Twenty five genes were on pig IMF QTL regions. Bone morphogenetic protein-binding endothelial cell precursor-derived regulator, forkhead box protein O1, ectodysplasin A receptor, ring finger protein 149, cluster of differentiation, tyrosine-protein phosphatase non-receptor type 1, SRY (sex determining region Y)-box 9 (SOX9), MYC proto-oncogene, and macrophage migration inhibitory factor were related to mitogen-activated protein kinase pathway, which regulates the differentiation to adipocytes. These genes and the genes mapped on QTLs could be the candidate genes affecting IMF. Heritability of IMF was estimated as 0.52, which was relatively high, suggesting that a considerable portion of the total variance of IMF is explained by the SNP information.

Conclusion: Our results can contribute to breeding pigs with better IMF and therefore, producing pork with better sensory qualities.

Keywords: Intramuscular Fat Content; Genome-wide Association Study; Heritability; Single Nucleotide Polymorphisms; Pig

INTRODUCTION

Intramuscular fat content (IMF), which stands for the amount of fat located throughout skeletal muscles, is a major quality trait of meat affecting sensory attributes such as flavor and texture. IMF is decided by the number and size of intramuscular adipocytes, and is directly related to the juiciness and tenderness of meat [1]. Pork with higher IMF tends to have better flavor, juiciness and tenderness, resulting in higher overall acceptability [2]. Therefore, breeding pigs to have higher IMF can produce pork that is more palatable.

Genome-wide association study (GWAS) enables determining the impact of genetic variants on various traits of animals affecting productivity. GWAS and genotyped single nucleotide polymorphism (SNP) data can identify genes associated to a certain economic trait of animals. Previous GWAS studies about IMF of pigs have found that heart-type fatty acid binding protein and long-chain-fatty-acid-CoA ligase 4 polymorphisms have an association with IMF of different pig populations [3]. Also, splicing factor, arginine/serine-rich 18 gene is reported to be related to the regulation of intramuscular fat deposition in pigs [4]. Polymorphic microsatellite loci CSSM34

and ETH10 were associated with marbling scores, which show the IMF in the Angus, Shorthorn and Wagyu cattle [1].

Many previous studies using GWAS to determine the association of genomic data with meat quality traits such as IMF focused on finding quantitative trait loci (QTL) [5-7]. Out bred line-cross model analysis suggested QTLs on chromosomes 2, 4, and 6, and the half-sib model analysis suggested linkage for chromosomes 4 and 7 [5]. The data of QTLs discovered from previous studies developed into a QTL database. The QTL database shows where QTL regions are located throughout chromosomes for each economic trait and animal. Using the QTL database, we can check whether a gene associated with a specific trait is within the known QTL region of the trait or not.

In this study, we analyzed the SNP data and IMF of pigs using GWAS to identify SNPs associated with IMF. To adjust the effect of inbreeding, a genetic relationship matrix was constructed and used during GWAS. Significant SNPs were matched to the nearest genes within 100 kb. We compared the identified genes with the QTL database of pig IMF and classified the function of the identified genes. We also estimated the heritability of IMF using the data. This study aims to search genes associated to IMF of pig and furthermore, to provide knowledge for breeding pigs having better IMF consequently, better sensory qualities.

MATERIALS AND METHODS

Ethics statement

The study protocol and the standard operating procedures (No. 2009-077, C-grade) of Berkshire pigs were reviewed and approved by National Institute of Animal Science's Institutional Animal Care and Use Committee.

Animals and phenotype records

An inbred Berkshire population was used for analysis, and IMF of the Berkshire sample was measured. Seven hundred and four samples were examined. Among them, 367 samples were male, 204 samples were female and the sex of 133 samples was unknown. The IMF of each pig was measured by chemical fat extraction procedures.

Genotyping and quality control

The genomic DNAs of pigs were genotyped on the Illumina Porcine 60 K SNP Beadchip. A total of 62,163 SNPs were genotyped. We discarded the markers with low minor allele frequency (<0.05), significant deviation from Hardy-Weinberg equilibrium ($p < 10^{-3}$), and low genotype call rate ($<95\%$). Among 62,163 SNPs, 40,191 SNPs passed quality control. The Hardy-Weinberg test failed 3,651 SNPs, 3,304 SNPs failed the genotype missingness test, and 19,829 SNPs failed the minor allele frequency test.

Genome-wide association analysis

The phenotype (IMF) was standardized to z-scores by subtract-

ing the mean and then dividing by the standard deviation, in each sex group (male, female, unknown) separately. Single trait, univariate linear mixed model was used for the analysis assuming additive effect of SNPs. SNP effects were treated as random effects and sex was added as a covariate. Software GEMMA was used to calculate the genetic relationship matrix of individuals and to test the effects of SNPs by likelihood ratio test [8]. The cutoff for statistical significance of genes was $p < 0.01$.

Gene annotation and functional classification

Gene annotation of significant SNPs was based on the Ensembl Genes 89 database of *Sus scrofa* genes (Sscrofa 10.2). Significant SNPs were annotated to the nearest genes within a distance of 100 kb. Functional classification of genes was performed on DAVID, an online functional annotation database. *Sus scrofa* was selected as both species and background option. The cutoff of gene ontology was $p < 0.05$.

Heritability estimation

The GCTA tool [9] was used to calculate heritability for IMF. We calculated the genetic relationship matrix between all pairs of samples using all the autosomal SNPs. We then estimated the variance of genetic component by restricted maximum likelihood analysis, and heritability by dividing the estimated genetic variance by the total variance measured.

RESULTS

Identification of significant single nucleotide polymorphisms

A total of 365 SNPs from all 19 chromosomes were identified as significant SNPs as the result of GWAS. Chromosome 14 contained 53 significant SNPs which was the largest number among all chromosomes. There were 40 and 35 significant SNPs on chromosome 7 and 11 respectively, which contained second and third number of significant SNPs. The statistical significance values of the association between each SNP and IMF calculated as $-\log_{10}$ (p-value) across 18 autosomal chromosomes and chromosome X was plotted in the form of a Manhattan plot (Figure 1).

The 365 significant SNPs found from GWAS were annotated to the nearest genes within 100 kb. Among the 365 significantly identified SNPs, 153 were annotated across 120 genes. There were some SNPs annotated to same genes and none of the significant SNPs on chromosome 8 and 15 had genes within 100 kb distance. Full information of significant SNPs, their chromosome number, position, closest gene, distance from the closest gene, raw p-value is on Supplementary Table S1.

Mapping on quantitative trait loci database

Identified genes were compared with IMF QTL regions base on the Pig QTLdb. Total 25 genes from the 120 significant genes, which is 20.8%, were included in reported pig IMF QTL regions. Seven genes on chromosome 9, 6 genes on both chromosome 2

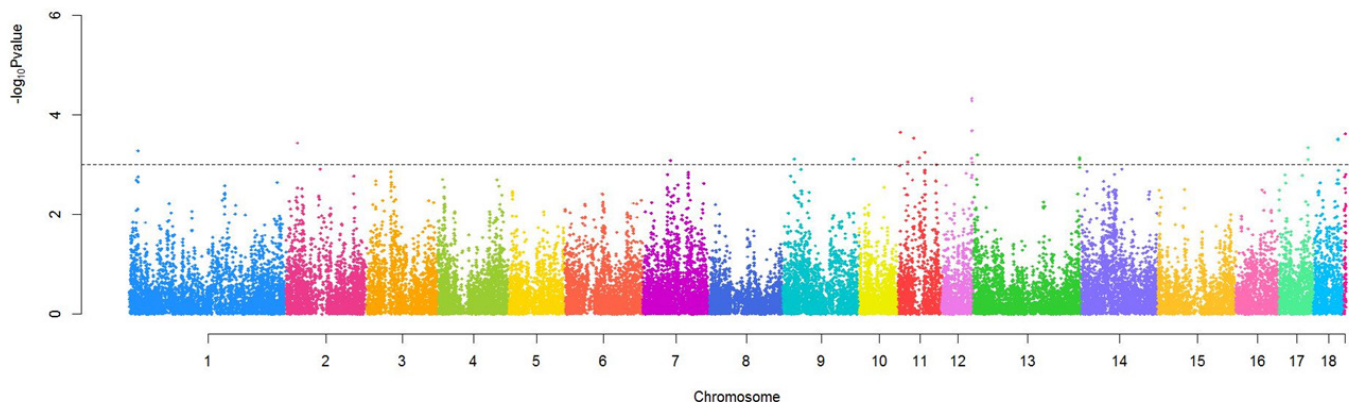


Figure 1. The statistical significance values of the association of single nucleotide polymorphisms (SNPs) across 18 autosomal chromosomes and the X chromosome with intramuscular fat content (IMF) plotted as $-\log_{10} p$ values. The horizontal dotted line indicates the cutoff $p = 0.01$.

and 6, and 2 genes on chromosome 4, 7, and 17 each were mapped on QTLs (Figure 2). This suggests a considerable proportion of the genes identified from this study was consistent with the previous QTL studies, and those genes can be considered as genes that are located on the section of the genome having high correlation with IMF of pigs.

Functional classification

Identified genes were classified by their biological function and ontology. Regulation of mitogen-activated protein kinase (MAPK) cascade was the most significant gene ontologies from GOTERM_BP_5. The full result of functionally annotated genes are listed on Supplementary Table S2. Especially, bone morphogenetic protein (BMP)-binding endothelial cell precursor-derived regulator (BMPER), forkhead box protein O1 (FOXO1), ectodysplasin A

receptor, ring finger protein 149 (RNF149), cluster of differentiation (CD40), tyrosine-protein phosphatase non-receptor type 1 (PTPN1), SRY (sex determining region Y)-box 9 (SOX9), MYC proto-oncogene (MYC), and macrophage migration inhibitory factor were categorized as genes related to both MAPK cascade and the regulation of MAPK cascade. In addition, growth/differentiation factor 9 and BMP6 were related to the regulation of MAPK pathway. FOXO1, RNF149, PTPN1, MYC were additionally annotated to negative regulation of MAPK cascade and regulation of stress-activated MAPK cascade.

Estimated heritability

Heritability of IMF was estimated by GCTA. The total variance of the sample was 1.020818 and the genetic variance was 0.526911. The genetic variance was estimated by the variance of genome-

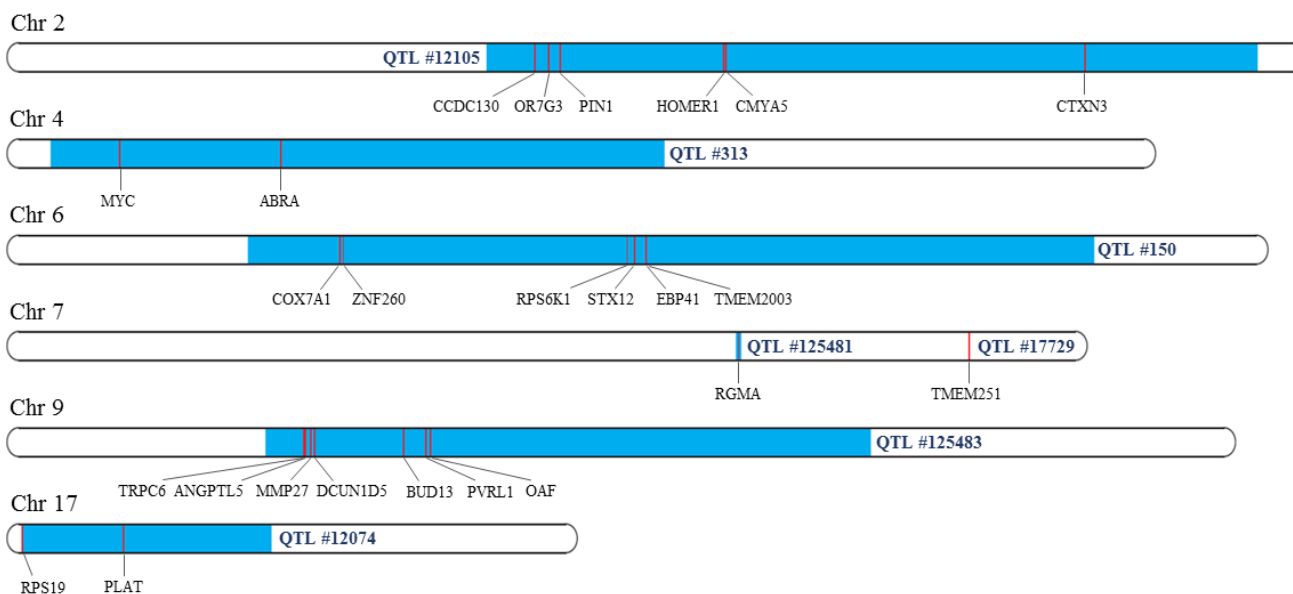


Figure 2. Location of significant genes mapped on quantitative trait loci (QTLs). The light gray areas indicate where genes are located and the dark gray areas indicate QTL regions.

wide SNPs. The estimated heritability was 0.516166, approximately 0.52, and the standard deviation of the estimated heritability was 0.061655.

In previous studies, heritability of IMF was estimated at 0.39 [10], 0.44 [11], 0.52 [12], 0.65 [13]. Estimated heritability of IMF in the referred studies had values between 0.39 and 0.65. The heritability estimated from the SNP and phenotype data in this study, 0.52, was in the range of reported estimations and was according with the previous studies.

DISCUSSION

Genes related to MAPK cascade and adipocyte differentiation

MAPK cascade was the most significant gene ontology (GO) term from the functional annotation results of significant genes and other GO terms related to MAPK cascade appeared multiple times as well. MAPK pathway regulates various cell functions such as proliferation, differentiation and mitosis [14]. Moreover, MAPK pathway is closely related to the differentiation of preadipocytes to adipocytes [15,16]. Some of the proteins involved in MAPK pathway also regulate adipocyte differentiation. For example, MAPK phosphatase-1 (MKP-1) downregulates the expression of p42/p44 MAPK and plays an important role in adipocyte differentiation [15]. In addition, inhibition of p38MAPK decreases adipocyte differentiation in humans and therefore p38MAPK activation can be seen as a requirement for primary human adipocyte differentiation [16]. Since IMF is determined by the amount of adipocytes, genes related to MAPK pathway could affect IMF by regulating the amount of adipocyte differentiation.

Some of the significant genes related to MAPK cascade or the regulation of MAPK cascade (BMPER, FOXO1, SOX9, PTPN1, CD40) are previously reported to have influence on adipocyte differentiation. BMPER directly interacts with BMPs [17]. Some BMPs activate p38MAPK pathway through the MAPK kinase (MAPKKK) cascade [18] and BMPER could be needed for adipocyte differentiation to activate p38MAPK. Furthermore, BMP4 has an effect on lipid accumulation as well as expression of adipocyte markers [19]. Furthermore, BMP2 and BMP7 induces adipocyte differentiation at low concentrate in C3H10T1/2 cell line [20]. FOXO1 is expressed in the early stages of adipocyte differentiation and acts as a preadipocyte differentiation preventing substituent [21]. Epidermal growth factor (EGF) repeat containing transmembrane protein (pref1) activates MAPK and upregulates SOX9 resulting in inhibition of adipocyte differentiation [22]. CD40 is related to the activation of MAPK [23] and PTPN1 is a negative regulator of CD40 [24]. PTPN1 polymorphisms are reported to be associated with adipocyte related measures such as body fat percentage [25].

Heritability of IMF and selection

The estimated heritability of IMF, 0.52, was relatively high. This means that a substantial part of the total phenotypic variance of

IMF is explained by the genetic variance. Here, the variance is that of the population, and thus high heritability suggests high genetic influence in the population on the whole [26]. Heritability is an important parameter for predicting the response to selection [27]. Since the Breeder's equation is given as $R = h^2 S$, where R is the response to selection, S is the selection differential and h is the heritability [28], higher heritability can result in stronger response of selection and effective selection. Therefore, the phenotype information of IMF can be useful information for selecting pigs to breed pigs having high level of IMF.

Pork containing more than 3% IMF tends to have higher sensory qualities including juiciness, tenderness and taste [29]. As IMF of pork increased from a range of 2.01% to 3.0% to higher than 3%, juiciness, tenderness, and both the intensity and desirability of taste increased. Since the current average of IMF measured from the Berkshire sample was 2.82%, if we increase IMF up to 3% by selection and breeding, we would be able to produce pork with improved juiciness, tenderness and taste.

Limitations of results

The tool used for association analysis in this study, GEMMA, adjusts the effect of sex by using sex as a covariate and uses genetic information from the X chromosome in the same way as those from autosomal chromosomes while computing the genetic relationship matrix [8]. However, since females carry two copies of X chromosomes while males carry a single copy, different methods should be used to estimate the genetic relationship for female-female pairs, male-male pairs and female-male pairs respectively in GWAS analyses [9]. Furthermore, among the 704 samples used in this study, the sex of 133, 19% of the sample, was unknown. Also, to balance the allele dosage between sexes, one of the female X chromosome is silenced by random X chromosome inactivation [30]. Therefore, additional information coding which allele was inactivated is needed to adjust GWAS analyses. In this study, information about which allele was inactivated was not provided, and this might together cause inaccuracy in the results from the X chromosome [31]. However, the proportion of significant SNPs on the X chromosome was 2.47% (9 out of 365) which was relatively low. Thus, some part of inaccuracy in the results from the X chromosome may have not affect the overall results of the study.

Owing to the small sample size of animals, the overall significance of the study was low. Small sample size lowers the estimated effect size and consequently lowers the power of the study. The estimated power of the study was only 0.21 [32]. To detect significantly associated SNPs in a study with low power, we had to use liberal statistics and a liberal cutoff (raw p-value and $p < 0.01$). This might cause some significant SNPs to be false positives, but still the SNPs detected in this study can be suggested as candidates for SNPs related to IMF of pig. Besides, we could pick out some SNPs more likely to related to IMF of pig by comparing them with known QTLs or searching their biological pathways. The genes mapped on QTLs or related to MAPK cascade may be stronger

candidates for genes that are associated with IMF of pig than others.

CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

ACKNOWLEDGMENTS

This work was supported by a grant from the Next-Generation BioGreen 21 Program (Project No. PJ01111501), Rural Development Administration, Republic of Korea.

REFERENCES

- Hocquette JF, Gondret F, Baéza E, et al. IMF in meat-producing animals: development, genetic and nutritional control, and identification of putative markers. *Animal* 2010;4:303-19.
- Fernandez X, Monin G, Talmant A, et al. Influence of intramuscular fat content on the quality of pig meat—1. Composition of the lipid fraction and sensory characteristics of *m. longissimus lumborum*. *Meat Sci* 1999;53:59-65.
- Chen JN, Jiang YZ, Cen WM, et al. Distribution of *H-FABP* and *ACSL4* gene polymorphisms and their associations with IMF and backfat thickness in different pig populations. *Genet Mol Res* 2014;13:6759-72.
- Wang X, Xue C, Wang X, et al. Differential display of expressed genes reveals a novel function of *SFRS18* in regulation of intramuscular fat deposition. *Int J Biol Sci* 2009;5:28-33.
- de Koning DJ, Janss LL, Rattink AP, et al. Detection of quantitative trait loci for backfat thickness and Intramuscular Fat Content in pigs (*Sus scrofa*). *Genetics* 1999;152:1679-90.
- Paszek AA, Wilkie PJ, Flickinger GH, et al. Interval mapping of carcass and meat quality traits in a divergent swine cross. *Anim Biotechnol* 2009;12:155-65.
- Ovilo C, Pérez-Enciso M, Barragán C, et al. A QTL for intramuscular fat and backfat thickness is located on porcine chromosome 6. *Mamm Genome* 2000;11:344-6.
- Xiang Z, Matthew, S. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* 2012;44:821-4.
- Yang J, Lee SH, Goddard ME, et al. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011;88:76-82.
- Suzuki K, Irie M, Kadowaki H, et al. Genetic parameter estimates of meat quality traits in Duroc pigs selected for average daily gain, longissimus muscle area, backfat thickness, and IMF. *J Anim Sci* 2005;83:2058-65.
- Larzul C, Lefaucheur L, Ecolan P, et al. Phenotypic and genetic parameters for longissimus muscle fiber characteristics in relation to growth, carcass, and meat quality traits in large white pigs. *J Anim Sci* 1997;75:3126-37.
- Lo LL, McLaren DG, McKeith FK, et al. Genetic analyses of growth, real-time ultrasound, carcass, and pork quality traits in Duroc and Landrace pigs: II. Heritabilities and correlations. *J Anim Sci* 1992;70:2387-96.
- Newcom DW, Baas TJ, Schwab CR, et al. Relationship between Backfat depth and its individual layers and intramuscular fat percentage in swine [Internet]. c2004 [cited 2017 Mar 10]. *Animal Industry Report*. Available from: http://lib.dr.iastate.edu/ans_air/vol650/iss1/103
- Pearson G, Robinson F, Gibson TB, et al. Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions. *Endocr Rev* 2001;22:153-83.
- Sakaue H, Ogawa W, Nakamura T, et al. Role of MAPK phosphatase-1 (MKP-1) in adipocyte differentiation. *J Biol Chem* 2004;279:39951-7.
- Aouadi M, Jager J, Laurent K, et al. p38MAP Kinase activity is required for human primary adipocyte differentiation. *FEBS Lett* 2007;581:5591-6.
- Moser M, Binder O, Wu Y, et al. BMPER, a novel endothelial cell precursor-derived protein, antagonizes bone morphogenetic protein signaling and endothelial cell differentiation. *Mol Cell Biol* 2003;23:5664-79.
- Tseng Y, Tong-Chuan H. Bone morphogenetic proteins and adipocyte differentiation. *Cellscience Rev* 2007;3:342-60.
- Bachner D, Ahrens M, Schroder D, et al. Bmp-2 downstream targets in mesenchymal development identified by subtractive cloning from recombinant mesenchymal progenitors (C3H10T1/2). *Dev Dyn* 1998;213:398-411.
- Wang EA, Israel DI, Kelly S, et al. Bone morphogenetic protein-2 causes commitment and differentiation in C3H10T1/2 and 3T3 cells. *Growth Factors* 1993;9:57-71.
- Nakae J, Kitamura T, Kitamura Y, et al. The forkhead transcription factor *Foxo1* regulates adipocyte differentiation. *Dev Cell* 2003;4:119-29.
- Wang Y, Sul HS. Pref-1 regulates mesenchymal cell commitment and differentiation through *Sox9*. *Cell Metab* 2009;9:287-302.
- Shirakata Y, Ishii K, Yagita H, et al. Distinct subcellular localization and substrate specificity of extracellular signal-regulated kinase in B cells upon stimulation with IgM and CD40. *J Immunol* 1999;163:6589-97.
- Medgyesi D, Hobeika E, Biesen R, et al. The protein tyrosine phosphatase *PTP1B* is a negative regulator of CD40 and BAFF-R signaling and controls B cell autoimmunity. *J Exp Med* 2014;211:427-40.
- Ukkola O, Rankinen T, Lakka T, et al. Protein tyrosine phosphatase 1B variant associated with fat distribution and insulin metabolism. *Obes Res* 2005;13:829-34.
- Griffiths, Anthony JF. An introduction to genetic analysis. New York, America: Macmillan; 2005.
- Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet* 2008;9:255-66.
- Falconer DS. Introduction to quantitative genetics. New York, America: Pearson Education India; 1975.
- Daszkiewicz T, Bąk T, Denaburski, J. Quality of pork with a different intramuscular fat (IMF) content. *Pol J Food Nutr Sci* 2005;14:31-5.
- Ahn J, Lee J. X chromosome: X inactivation. *Nat Educ* 2008;1:24.

31. Tukiainen T, Pirinen M, Sarin AP, et al. Chromosome X-wide association study identifies Loci for fasting insulin and height and evidence for incomplete dosage compensation. *PLoS Genet* 2014;10:E1004127.
32. Visscher PM, Hemani G, Vinkhuyzen AA, et al. Statistical power to detect genetic (co) variance of complex traits using SNP data in unrelated samples. *PLoS Genet* 2014;10:e1004269.