JAST (Journal of Animal Science and Technology) TITLE PAGE

Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research article
Article Title (within 20 words without abbreviations)	**Weighted single-step genome-wide association study (WssGWAS) to reveal new candidate genes for productive traits of Landrace pig in Korea**
Running Title (within 10 words)	WssGWAS to reveal new candidate genes for the growth
Author	Jun Park1, Chong-Sam Na1
Affiliation	1 Department of Animal Biotechnology, Jeonbuk National University, Jeonju 54896, Korea
ORCID (for more information, please visit https://orcid.org)	Jun Park(https://orcid.org/0000-0003-2682-5177) Chong Sam Na(https://orcid.org/0000-0002-8979-5633)
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	Not applicable.
Acknowledgements	Not applicable.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Park J, Na CS. Data curation: Park J. Formal analysis: Park J Methodology: Park J, Na CS. Software: Park J. Validation: Na CS. Investigation: Na CS. Writing - original draft: Park J. Writing - review & editing: Park J, Na CS.
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because there are no human and animal participants.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Chong Sam Na
Email address – this is where your proofs will be sent	csna@jbnu.ac.kr
Secondary Email address	
Address	Department of Animal Biotechnology, Jeonbuk National University, Jeonju 54896, Korea
Cell phone number	+82-10-6238-0720
Office phone number	+82-63-270-2607
Fax number	+82-63-270-2614

Abstract

The objective of this study was to identify genomic regions and candidate genes associated with productive traits using a total of 37,099 productive records and 6,683 SNP data obtained from five Great-Grand-Parents (GGP) farms in Landrace. The estimated of heritabilities for days to 105kg (AGE), average daily gain (ADG), backfat thickness (BF), and eye muscle area (EMA) were 0.49, 0.49, 0.56, and 0.23, respectively. We identified a genetic window that explained 2.05-2.34% for each trait of the total genetic variance. We observed a clear partitioning of the four traits into two groups, and the most significant genomic region for AGE and ADG were located on the SSC 1, while BF and EMA were located on SSC 2. We conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG), which revealed results in three biological processes, four cellular component, three molecular function, and six KEGG pathway. Significant SNPs can be used as markers for quantitative trait loci (QTL) investigation and genomic selection (GS) for productive traits in Landrace pig.

Keywords: Gene ontology; Kyoto Encyclopedia of Genes and Genomes; Landrace pigs; Productive traits; Weighted single-step GWAS

Introduction

Pig breeding for economic traits has undergone continuous improvement over time, with ongoing research in this field. Productive traits such as ADG, AGE, and BF have moderate to high heritability. ADG and AGE directly influence pig growth [1, 2]. BF is a trait linked to reproductive performance of Landrace and Yorkshire sows [3], making it crucial for enhancing and maintaining mothering ability of the dam.

According to the Korean Swine Performance Recording Standards (KSPRS) established by the Ministry of Agriculture, Food and Rural Affairs (MAFRA), performance testing is conducted within a weight range of 70-110 kg, with the endpoint set at 90kg. Days to reach 90 kg and BF are adjusted to assess growth trait performance. However, the endpoint weight of 90 kg has remained unchanged since its establishment in 1984, reflecting the market weight of finishing pigs at that time. With current trend of market weights surpassing 110 kg, there is a growing consensus that the endpoint weight for performance testing should be increased. Consequently, there is a need to develop a new adjustment formula for performance testing, resulting in the creation of a 105 kg-based adjustment formula by the National Institute of Animal Science (NIAS).

Genome-wide association studies (GWAS) have been widely applied in various fields, including the identification of economic traits. Multiple candidate genes and significant markers have been reported for the same trait, with associations between multiple traits observed at the same locus. These results are inherent to quantitative traits, single-marker GWAS analyses might have limited power for detecting quantitative trait loci (QTLs) and mapping accuracy [4]. The cost of analyzing SNP panels and the imbalance between individuals with genomic data and those without genomic data present additional limitations.

The WssGWAS method has emerged as a powerful approach that leverages GEBVs derived from genotypes, phenotypes, and pedigree information to estimate the effects of single-nucleotide polymorphisms (SNPs) [5]. This method effectively addresses the issue of unequal variances among SNPs, leading to more accurate estimation of SNP effects [6]. WssGWAS is more effective than GWAS in analyzing traits that are influenced by QTLs with significant effects or when there is insufficient phenotype and genotype data available. Recent studies have successfully used this approach to identified various economic traits in livestock species [7-9].

We investigate the genetic regions and candidate genes associated with productive traits (adjusted to 105 kg body weight) in Landrace pig using WssGWAS. Also, we conducted GO and KEGG enrichment analyses to gain deeper insights into the underlying biological processes and functional terms associated with the identified candidate genes for productive traits.

Materials and Methods

Ethical approval

This article does not require IRB/IACUC approval because there are no human and animal participants.

Animals and phenotypes

We obtained the total 37,099 productive records (9,818 males and 27,281 females) born from 2015 to 2021 at five GGP farms (Table S1). We adjusted to evaluate for productive traits (AGE, ADG, BF, and EMA to 105kg) with

method outlined by the NIAS in Korea (https://www.nias.go.kr/images/promote/result/file/2021_2_5.pdf), and the equations used are as follows:

$$Adjusted\ AGE = Measure\ age - \frac{(105 - Measure\ weight) \times (Measure\ Age - \alpha)}{Measure\ weight}$$

Where α is the correction factor used to adjust AGE to 105kg as follows:

$$\alpha : Sire = 63.3\ ;\ Dam = 47.3$$

ADG adjusted to 105kg is calculated using the following equation:

$$adjusted\ ADG = \frac{105\ kg}{adjusted\ AGE}$$

BF adjusted to 105kg is calculated using the following equation:

$$adjusted\ BF = Measure\ BF \times \frac{(105 - Measure\ weight) \times (Measure\ BF - \beta)}{Measure\ weight}$$

Where β is the correction factor used to adjust BF to 105kg as follows:

$$\beta : Sire = 2.6\ ;\ Dam = 3.7$$

EMA adjusted to 105kg is calculated using the following equation:

$$adjusted\ EMA = Measure\ EMA \times \frac{(105 - Measure\ weight) \times (Measure\ EMA - \gamma)}{Measure\ weight}$$

Where γ is the correction factor used to adjust EMA to 105kg as follows:

$$\gamma : Sire = 29.1\ ;\ Dam = 33.0$$

SNP data and quality control (QC)

Illumina Porcine 60K V1 and V2 were used and V2 was selected as a reference panel for imputation. Prior to imputation, phasing was performed using Shapeit4 [10], a fast and accurate method for haplotype estimation using a PBWT-based approach to select informative conditioning haplotypes. Imputation was then conducted using Impute5 [11], assuming phased samples having no missing alleles. After imputation, quality control (QC) was performed by PLINK v1.09 [12] to exclude SNPs with low call rates (< 90%), low minor allele frequencies (< 0.01), or deviation from Hardy-Weinberg equilibrium (10^{-6}). After QC, we used the number of animals and SNPs were 6,683 and 35,420, respectively.

Statistical analysis

We estimated the genetic parameters for AGE, ADG, BF, and EMA with average information restricted maximum likelihood (AIREML) method. We considered two approaches: pedigree-based BLUP (PBLUP) and ssGBLUP. Each trait was estimated with a single-trait animal model, and the equation as follows:

$$y = Xb + Za + e$$

where y is the vector of observations; b is the vector of fixed effects (herd-birth year-season, sex); a is the vector of additive genetic effects; e is the vector of residuals; and X and Z are the incidence matrices for b, a, and e. Heritability was estimated as $h^2 = \frac{\sigma_a^2}{\sigma_a^2+\sigma_e^2}$, where σ_a^2 and σ_e^2 were additive genetic and residual variances, respectively.

Furthermore, GEBVs calculated using ssGBLUP approach, and marker effects were derived from these GEBVs. In contrast to the conventional BLUP approach, ssGBLUP substituted the inverse of the pedigree relationship matrix

(A^{-1}) with the inverse of the combined matrix H^{-1}, which incorporated both the pedigree and genomic relationship matrices [13]. The H^{-1} can be represented as follows:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

where A_{22}^{-1} is the inverse of numerator relationship matrix for pigs with genotyped, and G refers to the genomic relationship matrix [14]. G is presented below:

$$G = \frac{ZDZ'}{\sum_{i=1}^{M} 2p_i(1-p_i)}$$

where Z is a matrix of gene content adjusted for allele frequencies (*0*, *1* or *2* for AA, Aa and aa, respectively), D is a diagonal matrix of weights for SNP variances (initially $D = I$), M is the number of SNPs, and p_i is the minor allele frequency of i^{th} SNP. Estimates of SNP effects and weights for WssGWAS were obtained according to following steps [5]:

1. First step (t = 1): $D = I$; $G_{(t)} = D_{(t)}Z'\lambda$, where $\lambda = \frac{1}{\sum_{i=1}^{M} 2p_i(1-p_i)}$ [5];
2. Calculate GEBVs;
3. Convert GEBVs to SNP effects ($\hat{u}$): $\hat{u} = \lambda D_{(t)}Z'G_{(t)}^{-1}\hat{a}_g$, where $\hat{a}_g$ was the GEBV of animal that was also genotyped;
4. Calculate the weight for each SNP: $d_{i(t+1)} = \hat{u}_{i(t)}^2 2p_i(1-p_i)$, where i was the i^{th} SNP;
5. Normalize SNP weights to keep the total genetic variance constant:

$$D_{(t+1)} = \frac{tr(D_{(1)})}{tr(D_{(t+1)})} D_{(t+1)}$$

6. $G_{(t+1)} = ZD_{(t+1)}Z'\lambda$ was calculated;
7. $t = t + 1$ and loop to step 2.

The procedure was iteratively performed for a total of three cycles, taking into account the achieved accuracies of GEBV [15, 16]. During each iteration, the weights of single nucleotide polymorphisms (SNPs) were updated (steps 4 and 5), and utilized to construct G matrices (step 6), update GEBV (step 2), and estimate SNP effects (step 3). Subsequently, the proportion of genetic variance explained by each consecutive set of SNPs, referred to as i^{th} SNP windows, was calculated [16]. In a previous study, values for a_i were determined based on LD decay distance analysis of the population, considering the distance where r^2 drops below 0.2 [17]. In this study, LD decay distance was not calculated separately, and to facilitate comparison with the previous study's findings [17], the same value of 0.8 Mb was adopted. SNPs were positioned within a 0.8 Mb region, and the percentage of genetic variance explained by each 0.8 Mb window was determined as follows:

$$\frac{Var(a_i)}{\sigma_a^2} \times 100 = \frac{Var(\sum_{j=1}^{x} Z_j\hat{u}_j)}{\sigma_a^2}$$

where a_i is the genetic value of the i^{th} SNP window that consisted of a region of consecutive SNPs located within 0.8 Mb, σ_a^2 was the total additive genetic variance, Z_j was the vector of gene content of the j^{th} SNP for all individuals, and $\hat{u}_j$ was the effect of the j^{th} SNP within the i^{th} window. To visualize the distribution of these SNP windows, Manhattan plots were generated using the R software and CMplot package [18, 19]. The procedures described above were implemented iteratively using the software suite of BLUPF90 programs [20].

Identification of candidate genes and functional enrichment analysis

We conducted to identify specific genomic regions associated with productive traits by examining QTL using genomic windows that accounted for more than 1.0% of the total genetic variance.

These genomic windows, previously employed in similar studies [17], represent regions of the genome that contribute significantly to the genetic variation underlying productive traits.

Our focus on these candidate QTL regions aimed to uncover genetic markers or regions that play a pivotal role in influencing growth-related characteristics. Notably, we observed a significant deviation from the expected average genetic variance explained by the 0.8 Mb window, which accounted for 0.0495% of the genetic variance on average (dividing 100% by the number of 2022 genomic regions). The 1% threshold exceeded the anticipated average genetic variance explained by the 0.8 Mb window by more than 20-times. To identify genes within the identified QTL regions, particularly within the significant windows, we utilized the ensemble *Sus scrofa* 11.1 database (https://www.ensembl.org/biomart). Furthermore, to gain deeper insights into the biological processes associated with these regions, we performed gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses using the Database for Annotation, Visualization, and Integrated Discovery (DAVID v6.8, https://david.ncifcrf.gov/). GO terms and KEGG pathways showing significant enrichment were determined based on a *p*-value threshold of < 0.05. Through these analyses, we gained valuable knowledge regarding crucial molecular pathways and biological functions associated with the observed genetic variations.

Results and Discussion

Variance component and heritability

The estimates of the heritabilities for AGE, ADG, BF, and EMA were 0.49, 0.49, 0.56, and 0.23, respectively (Table 1). Results showed that the heritability of ssGBLUP was higher than that of PBLUP, which only used pedigree information. The ssGBLUP method, which incorporates both pedigree and genetic information, theoretically provides more accurate estimates of genetic parameters [7].

Genome-wide association study (GWAS)

In most cases, major economic traits of livestock are quantitative traits except for some traits. These quantitative traits are characterized by a complex genetic structure. Exploration of candidate genes for such traits has always been an important goal of animal breeding programs. In this study, the genetic variance explained by a 0.8 Mb window for each growth trait was estimated using WssGWAS (Fig 1). Specifically, we explained 2.05%, 3.23%, 9.27%, and 9.96% of the total genetic variation for AGE, ADG, BF, and EMA, respectively, with the most significant window explaining approximately 2.05-2.34% of the total genetic variation (Table 2). Furthermore, within the identified window regions of this study, we presented the SNP markers, their corresponding chromosome (Chr), positions, and the associated genetic variance values explained by each marker (Table S2-S5).

Previous GWAS studies have reported significance regions on SSC 1, 3, 6, 8, and 13 for ADG and on SSC 1, 3, 6, 8, and 10 for AGE, explaining a total of 8.09% and 4.08% of the genetic variance, respectively [21]. Moreover, candidate QTL regions on SSC 4 and 14 for AGE, on SSC 4 and 2 for ADG, and on SSC 2, 3, and 10 for BF explain

a total of 6.48%, 5.96%, and 6.76% of genetic variance, respectively [4]. The utilization of the WssGWAS, which incorporates SNP windows for genetic variance estimation, offers improved capabilities in identifying previously unknown QTLs compared to conventional GWAS methods. This approach mitigates the risk of overestimating the number of detected QTLs and false positives resulting from linkage disequilibrium [22, 23]. Furthermore, the iterative weighting of SNPs enhances the detection of QTLs with larger effects [16]. In this study, a total of 10 iterations were conducted, and the genomic accuracy for each trait was presented (Table S6). As the number of iterations increased, there was a corresponding increase in genetic accuracy, consistent with previous study [5]. The highest increase was observed at the 3rd iteration, followed by a gradual decrease. Unlike the study that reported a decrease in weights at certain iterations [5], our study showed an increase in accuracy up to 0.02 to 0.04 over 10 iterations, as compared to the first iteration where all SNP weights were set to I. While the optimal number of iterations for each trait was not conclusively determined in our study, we chose to use the results from the 3rd iteration, which exhibited the highest genetic accuracy, for the GWAS analysis.

Candidate gene for AGE and ADG

We have successfully identified three significant regions (SSC 1, 7, and 14) that are associated with AGE. These regions explain 1.03-2.03% of the total genetic variance for AGE. Additionally, we conducted gene annotation and identified five genes with potential as candidate genes. Similarly, ADG is discovered five relevant QTL regions (SSC 1, 2, 7, and 14) that account for 1.01-2.14% of the total genetic variance. Within these regions, we have annotated seven genes. Notably, although three QTL regions associated with AGE are also found to be associated with ADG, the proportions of genetic variance explained differ between the two traits.

When considering complex quantitative traits, it is important to acknowledge that linear gene effects may not consistently align with average trait values. Instead, a nonlinear assumption is often more appropriate [21], as gene contributions can exhibit nonlinearity and pleiotropic effects between traits may manifest [4]. Pleiotropic quantitative trait loci (QTLs) are prevalent in the porcine genome, as exemplified by the presence of QTLs associated with vertebral number, body length, and nipple number on SSC 7 [24]. Considering the overlap in the identified genomic regions and the substantial genetic correlation observed between ADG and AGE, it is reasonable to infer that the genes associated with these traits are shared.

Within the identified genomic regions, we observed the presence of *RELCH* in close proximity to *MC4R* on SSC 1. *RELCH* has been previously recognized as one of the seven potential candidate genes associated with pig fatness traits [25] and has demonstrated an association with pig fat depth [26]. Functionally, *RELCH* is involved in regulating intracellular cholesterol distribution, specifically from recycling endosomes to the trans-Golgi network. Gene ontology analysis further revealed enrichment in biological processes related to neuroactive ligand-receptor interaction [27]. These findings provide valuable insights into the potential regulatory mechanisms underlying fatness traits in pigs and highlight the role of *RELCH* in cholesterol metabolism and neuroactive signaling pathways.

RNF152 emerges as a promising candidate gene associated with pig fatness and body composition traits [25, 26], specifically backfat thickness in Duroc pigs as revealed by ssGWAS analysis [28]. This gene acts as a negative regulator of the mTOR signaling pathway [29, 30], a key pathway governing cellular metabolism, survival, and proliferation through the regulation of anabolic processes such as protein, lipid, and nucleotide synthesis. The pivotal role of the mTOR pathway in cellular function has been extensively documented [31-34]. Our study highlights the

potential pleiotropic effects within the SSC 1 region, which exhibited remarkable significance for both AGE and ADG traits. These findings provide valuable insights into the genetic architecture underlying productive traits and the interplay of key molecular pathways in pigs.

CDH20 has been identified as a candidate gene for pig fatness traits and days to reach 100 kg in previous studies [25, 35]. *CDH20* encodes a type 2 classical cadherin, which is a calcium-dependent cell-cell adhesion glycoprotein and a potential candidate for tumor suppression [36]. Additionally, *CDH20* is involved in the cell adhesion pathway. This study is the first to report its association with porcine growth and fatness traits [37].

TMEM132C has been identified as a potential candidate gene for growth and fatness-related traits in Bamaxiang pigs using a customized 1.4 million SNP array [38]. It has also been implicated as one of the candidate genes for average backfat at 100 kg [39].

NDUFV1 is located in the SSC 2 region and plays a critical role in energy metabolism [40]. Previous investigations have consistently demonstrated a significant downregulation of *NDUFV1* expression in placental tissues, particularly when compared to the control group representing normal pregnancies. Notably, *NDUFV1* plays a crucial role in facilitating energy production within the mitochondrial matrix and membrane, thereby influencing essential metabolic processes [41].

Candidate gene for BF and EMA

BF had the highest explained genetic variance and identified the highest number of candidate genes. Specifically, six relevant regions located on SSC 2, 5, 14, and 18 were identified, explaining 1.27-2.34% of the total genetic variance, and 21 genes were annotated. EMA had lower heritability other traits such as AGE, ADG, and BF, but it was moderate heritability. Moreover, the significant genetic regions identified for EMA did not coincide with those found for BF, although SSC 2 and 14 exhibited similar levels of variance explained. Similar to BF, the region with the highest significant genetic variance explained was SSC 2 with 2.07% for EMA, while SSC 6, SSC 7, SC14, and SSC 15 were also identified as regions associated with EMA.

ANO1, also known as *TMEM16A*, is a $Ca2^{+}$-activated chloride channel that plays a vital role in various physiological functions [42]. This channel is critical for maintaining the *STT* of urinary tract muscles in female mice and women. Sex differences in this context are likely influenced by *ANO1* expression in SMCs of the urethra, and this gene is also involved in smooth muscle contraction [43, 44].

PSMD13, also referred to as *S11, Rpn9, p40.5, or HSPC027*, is a 376 amino acid protein belonging to the proteasome subunit S11 family. It is located in the SSC 2 region and has been identified as being associated with loin depth in previous studies [45]. *COX8H* is a candidate gene situated in the SSC 2 region. It has been reported to explain 3.51% and 5.87% of the total genetic variation for BF and lean percent, respectively, in Yorkshire pigs [46]. Additionally, it has been identified as one of the highly expressed genes in intramuscular adipose tissues of Erhualian pigs [47]. *MAP3K11* belongs to the serine/threonine kinase family and plays a crucial role in the FGFR signaling pathway, which regulates cartilage and bone formation [48]. Furthermore, a previous study has suggested a potential association between *MAP3K11* and body weight in sheep [49].

AKAP3, located in the SSC5 region, is a member of the AKAP family. It interacts with the regulatory subunit of PKA [50]. While it has been predominantly studied in sperm and cancer, previous research has shown the expression of *AKAP3* in the longissimus dorsi muscle of pigs [51]. The expression of *AKAP3* in skeletal muscle and its binding

to PKA's regulatory subunit have the potential to affect glycogen content in the muscle, thereby impacting meat quality after post-mortem modifications [51]. *FGF6* is a key regulator of skeletal muscle development that influences muscle fiber diameter and intramuscular fat content [52, 53]. Additionally, *FGF6* has been employed in gene delivery systems for skeletal muscle repair [54].

ZYX is located in the SSC18 region and is closely associated with multiple QTLs related to tissue and texture characteristics [55]. *ZYX* is a protein present in focal adhesions depending on active fibers and interacts with the actin-crosslinking protein alpha-actinin. *ZYX* is involved in cellular organization, signal transduction, cellular response to mechanical stress, and cell adhesion [56-59]. Structurally, *ZYX* consists of an N-terminal domain that interacts with proteins involved in signal transduction and a C-terminal LIM domain that plays a crucial role in regulating cell proliferation, differentiation, and protein-protein and/or protein-DNA interactions [60].

MED9, located in the SSC 2 region, is an essential gene for the maintenance of white adipose tissues and adipogenesis in *Piscirickettsia salmonis* [61]. *MED9* also interacts with PPARs, which are important for inflammatory processes [62]. Polymorphism in the *SERPING1* gene has been found to be significantly associated with tenderness and pH24 in both dominant and co-dominant models. Furthermore, this gene can influence the postmortem pH of muscle by regulating glycolysis [63].

GO terms and KEGG pathway enrichment analysis

Enrichment analyses uncovered significant associations between multiple terms and productive traits. Specifically, we observed enrichment in three biological processes, four cellular components, three molecular functions, and six KEGG pathways (Table 3). Notably, the most significant GO term was GO:0004190, which pertains to chromatin. Furthermore, the GO:0005509 category, encompassing calcium ion binding, exhibited enrichment for nine candidate genes, constituting the majority of the candidates.

The process of actin filament bundle assembly (GO:0051017) involves the construction of actin filament bundles with varying degrees of tightness and orientation. It represents a vital aspect of cellular structure and function. Notably, the selective sweep gene *AIF1L* emerged as a significant molecule, playing an essential role in cell survival and contributing to proinflammatory activities of immune cells, including monocytes/macrophages and activated T lymphocytes [64, 65].

Chloride transmembrane transport (GO:1902476) refers to the movement of chloride across a membrane. Previous studies have implicated *ANO9* as a gene associated with marbling depth in both purebred and crossbred pigs. The genetic region containing this gene accounts for 3.34% of the total genetic variance for loin depth [45]. Additionally, the *CLCN1* gene participates in the transmission of nerve impulses, a crucial cellular communication process involved in the interaction between adipocytes and myogenic cells [66]. The interplay between these cell types is significant for various aspects of growth and development, including the regulation of myogenesis rate and extent, muscle growth, adipogenesis, lipogenesis/lipolysis, and energy substrate utilization [67].

Calcium ion binding (GO:0005509) denotes the process of binding to a calcium ion ($Ca2^{+}$). Prior research has identified EHD1 as a candidate gene that likely possesses functional relevance to meat quality in Beijing black pigs [68]. Additionally, a GWAS study revealed a significant association between *EHD1* and the meat-to-fat ratio (MFR) [69]. Furthermore, using *EHD1* knockout mice, researchers demonstrated the regulatory role of *EHD1* in cholesterol homeostasis and lipid droplet storage [70].

In conclusion, this study offers novel insights into the genetic basis of productive traits in pigs. The identified biological processes, pathways, and candidate genes serve as valuable resources for future investigations for genetic improvement with these traits. Significant SNPs can be used as markers for quantitative trait loci (QTL) investigation and genomic selection (GS) for productive traits in Landrace pig.

Tables and Figures

Tables

Table 1. Variance components and heritabilities for productive traits

Traits	Method	σ_a^{2*}	σ_e^{2*}	σ_P^{2*}	$h^2\ (SE)^*$
AGE (days)	PBLUP	47.66	58.18	105.84	0.45 (0.01)
	ssGBLUP	54.130	56.73	110.86	0.49 (0.01)
ADG (g)	PBLUP	766.41	923.76	1690.20	0.45 (0.01)
	ssGBLUP	889.23	890.09	1779.30	0.49 (0.01)
BF (mm)	PBLUP	3.69	3.40	7.109	0.52 (0.01)
	ssGBLUP	4.18	3.27	7.46	0.56 (0.01)
EMA (cm^2)	PLBUP	1.89	6.46	8.34	0.22 (0.01)
	ssGBLUP	1.96	6.50	8.45	0.23 (0.01)

$*\sigma_a^2$: additive genetic, σ_e^2: residual, σ_P^2: phenotypic variances, $h^2\ (SE)$: heritability and standard error.

295 **Table 2.** Significance regions and candidate genes for productive traits

Traits	SSC[1]	Position (Mb)	gVar (%)[2]	nSNP	Candidate Genes
AGE (days)	1	159.24-159.88	2.05	9	*RELCH, PIGN, RNF152, CDH20*
ADG (g)	1	159.24-159.88	2.22	9	*RELCH, PIGN, RNF152, CDH20*
	2	4.97-5.71	1.01	10	*NDUFV1, CABP4, CORO1B, PTPRCAP*
BF (mm)	2	2.46-3.26	2.34	15	*ACTE1, SHANK2, CTTN, ANO1*
	5	0.07-0.42	1.46	5	*PSMD13, COX8H*
		6.64-7.42	1.25	24	*MAP3K11*
		65.61-66.36	1.68	18	*NDUFA9, AKAP3, DYRK4, RAD51AP1, FGF6, C12orf4, TIGAR*
	14	19.67-20.42	1.27	10	*AADAT, MFAP3L, CLCN3, NEK1, SH3RF1*
	18	6.88-7.67	1.27	28	*ZYX, FAM131B*
EMA (cm^2)	2	13.04-13.46	2.07	21	*CTNND1, BTBD18, TMX2, MED19, SERPING1*
	6	10.19-10.99	1.26	23	*DDB1, VWCE, PPAG3*
		129.64-130.41	1.62	21	*TTLL7, ADGRL2*
		102.18-102.96	1.24	13	*AKAIN1, DLGAP1*
	7	109.35-110.14	1.42	24	*ENSSSCG00000052115, ENSSSCG00000037928*
	14	26.65-27.30	1.34	13	*TMEM132C, ENSSSCG00000042937*
	15	121.01-121.81	1.01	15	*CRYBA2, CFAP65, IHH*

296 [1]*Sus scrofa* chromosome; [2]represents the proportion of genetic variance explained by 0.8 Mb.
297
298

299 **Table 3.** Significant gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG)
300 pathways associated with productive traits of Landrace pigs ($p < 0.05$)

Gene ontology and KEGG pathway	nGenes	*p*-Value	Gene
GO:0051017-actin filament bundle assembly	2	0.02	*CORO1B, RHOD*
GO:1902476-chloride transmembrane transport	3	0.03	*ANO1, ANO9, CLCN1*
GO:0006303-double-strand break repair via nonhomologous end joining	3	0.01	*KDM2A, NHEJ1, PRPF19*
GO:0005886-plasma membrane	7	0.02	*CDH20, CORO1B, PIGN, PTPRCAP, RHOD, SPTBN2, SYT12*
GO:0035861-site of double-strand break	3	0.03	*DDB1, NHEJ1, PRPF19*
GO:0000785-chromatin	5	0.04	*RAD51AP1, CDCA5, CCND2, DPF2, MEN1*
GO:0008076-voltage-gated potassium channel complex	3	0.04	*CTTN, KCNA1, KCNA6*
GO:0004190-aspartic-type endopeptidase activity	6	0.00	*PGA5, pregnancy-associated glycoprotein 2-like, PPAG3, PIP*
GO:0005247-voltage-gated chloride channel activity	2	0.04	*CLCN1, CLCN3*
GO:0005509-calcium ion binding	9	0.04	*EHD1, IHH, NAALADL1, CDH20, CABP4, CAPN1, LTBP3, SYT12, VWCE*
ssc05012: Parkinson disease	6	0.03	*COX8H, NDUFV1, NDUFA9, PSMD13, PRKACB, UBE2L6*
ssc00982: Drug metabolism - cytochrome P450	3	0.04	*GSTK1*
ssc04340: Hedgehog signaling pathway	3	0.04	*IHH, CCND2, PRKACB*
ssc00480: Glutathione metabolism	2	0.04	*glutathione S-transferase P-like*
ssc00980: Metabolism of xenobiotics by cytochrome P450	2	0.04	*glutathione S-transferase P-like*
ssc05204: Chemical carcinogenesis - DNA adducts	2	0.04	*glutathione S-transferase P-like*

301

Figure

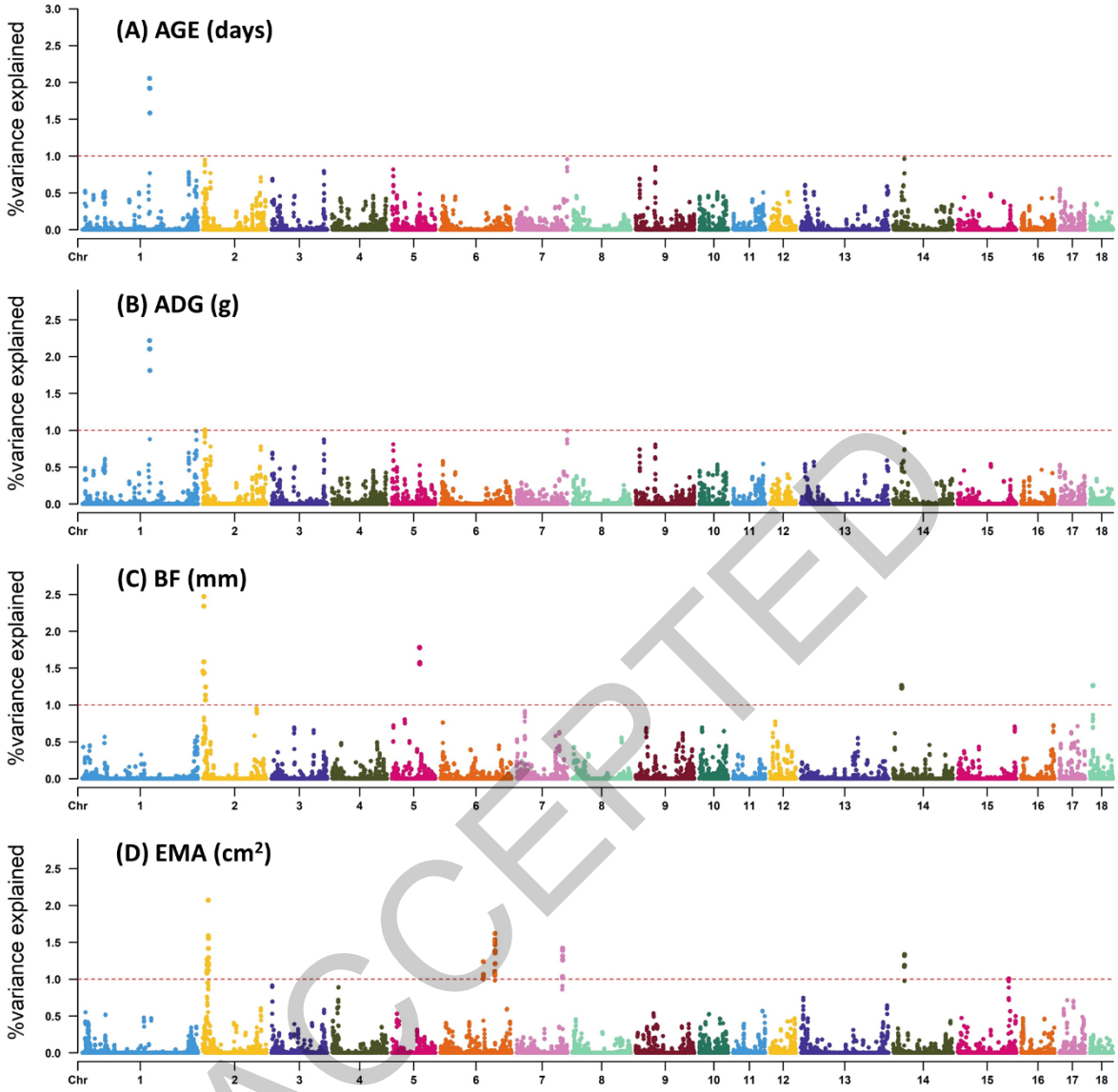

Figure 1. Proportion of genetic variances of productive traits explained by 0.8 Mb windows

References

1. Fontanesi L, Schiavo G, Galimberti G, Calo DG, Russo V. A genomewide association study for average daily gain in Italian Large White pigs. J Anim Sci. 2014;92(4):1385-94. 10.2527/jas.2013-7059

2. Ding R, Yang M, Wang X, Quan J, Zhuang Z, Zhou S, et al. Genetic Architecture of Feeding Behavior and Feed Efficiency in a Duroc Pig Population. Front Genet. 2018;9:220. 10.3389/fgene.2018.00220

3. Vargovic L, Bunter K, Hermesch S, editors. Economic benefit of additional recording for welfare traits in maternal breeding objectives for pigs. Proc Assoc Advmt Anim Breed Genet; 2021.

4. Ruan D, Zhuang Z, Ding R, Qiu Y, Zhou S, Wu J, et al. Weighted Single-Step GWAS Identified Candidate Genes Associated with Growth Traits in a Duroc Pig Population. Genes (Basel). 2021;12(1):117. 10.3390/genes12010117

5. Wang H, Misztal I, Aguilar I, Legarra A, Muir WM. Genome-wide association mapping including phenotypes from relatives without genotypes. Genet Res. 2012;94(2):73-83. 10.1017/S0016672312000274

6. Zhang X, Lourenco D, Aguilar I, Legarra A, Misztal I. Weighting Strategies for Single-Step Genomic BLUP: An Iterative Approach for Accurate Calculation of GEBV and GWAS. Front Genet. 2016;7:151. 10.3389/fgene.2016.00151

7. Marques DBD, Bastiaansen JWM, Broekhuijse M, Lopes MS, Knol EF, Harlizius B, et al. Weighted single-step GWAS and gene network analysis reveal new candidate genes for semen traits in pigs. Genet Sel Evol. 2018;50(1):40. 10.1186/s12711-018-0412-z

8. Luo H, Hu L, Brito LF, Dou J, Sammad A, Chang Y, et al. Weighted single-step GWAS and RNA sequencing reveals key candidate genes associated with physiological indicators of heat stress in Holstein cattle. J Anim Sci Biotechnol. 2022;13(1):108. 10.1186/s40104-022-00748-6

9. Brunes LC, Baldi F, Lopes FB, Lobo RB, Espigolan R, Costa MFO, et al. Weighted single-step genome-wide association study and pathway analyses for feed efficiency traits in Nellore cattle. J Anim Breed Genet. 2021;138(1):23-44. 10.1111/jbg.12496

10. Delaneau O, Zagury JF, Robinson MR, Marchini JL, Dermitzakis ET. Accurate, scalable and integrative haplotype estimation. Nat Commun. 2019;10(1):5436. 10.1038/s41467-019-13225-y

11. Rubinacci S, Delaneau O, Marchini J. Genotype imputation using the Positional Burrows Wheeler Transform. PLoS Genet. 2020;16(11):e1009049. 10.1371/journal.pgen.1009049

12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-75. 10.1086/519795

13. Aguilar I, Misztal I, Legarra A, Tsuruta S. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. J Anim Breed Genet. 2011;128(6):422-8. 10.1111/j.1439-0388.2010.00912.x

14. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91(11):4414-23. 10.3168/jds.2007-0980

15. Legarra A, Robert-Granie C, Manfredi E, Elsen JM. Performance of genomic selection in mice. Genetics. 2008;180(1):611-8. 10.1534/genetics.108.088575

16. Wang HY, Misztal I, Aguilar I, Legarra A, Fernando RL, Vitezica Z, et al. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. Front Genet. 2014;5:134. ARTN 134 10.3389/fgene.2014.00134

17. Zhuang Z, Ding R, Peng L, Wu J, Ye Y, Zhou S, et al. Genome-wide association analyses identify known and novel loci for teat number in Duroc pigs using single-locus and multi-locus models. BMC Genom. 2020;21(1):344. 10.1186/s12864-020-6742-6

18. R Core Team R. R: A language and environment for statistical computing. 2013.

19. Yin L. CMplot: circle manhattan plot. R package version. 2020;3(2).

20. Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee D, editors. BLUPF90 and related programs (BGF90). Proc of the 7th WCGALP; 2002: Montpellier.

21. Tang Z, Xu J, Yin L, Yin D, Zhu M, Yu M, et al. Genome-Wide Association Study Reveals Candidate Genes for Growth Relevant Traits in Pigs. Front Genet. 2019;10:302. 10.3389/fgene.2019.00302

22. Peters SO, Kizilkaya K, Garrick DJ, Fernando RL, Reecy JM, Weaber RL, et al. Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. J Anim Sci. 2012;90(10):3398-409. 10.2527/jas.2012-4507

23. Habier D, Fernando RL, Kizilkaya K, Garrick DJ. Extension of the bayesian alphabet for genomic selection. BMC Bioinform. 2011;12(1):186. 10.1186/1471-2105-12-186

24. Li Y, Pu L, Shi L, Gao H, Zhang P, Wang L, Zhao F. Revealing New Candidate Genes for Teat Number Relevant Traits in Duroc Pigs Using Genome-Wide Association Studies. Animals (Basel). 2021;11(3):806. 10.3390/ani11030806

25. Zeng H, Zhong Z, Xu Z, Teng J, Wei C, Chen Z, et al. Meta-analysis of genome-wide association studies uncovers shared candidate genes across breeds for pig fatness trait. BMC Genom. 2022;23(1):786. 10.1186/s12864-022-09036-z

26. Heidaritabar M, Bink M, Dervishi E, Charagu P, Huisman A, Plastow GS. Genome-wide association studies for additive and dominance effects for body composition traits in commercial crossbred Pietrain pigs. J Anim Breed Genet. 2023. 10.1111/jbg.12768

27. Sobajima T, Yoshimura SI, Maeda T, Miyata H, Miyoshi E, Harada A. The Rab11-binding protein RELCH/KIAA1468 controls intracellular cholesterol distribution. J Cell Biol. 2018;217(5):1777-96. 10.1083/jcb.201709123

28. Zhang Z, Zhang Z, Oyelami FO, Sun H, Xu Z, Ma P, et al. Identification of genes related to intramuscular fat independent of backfat thickness in Duroc pigs using single-step genome-wide association. Anim Genet. 2021;52(1):108-13. 10.1111/age.13012

29. Deng L, Jiang C, Chen L, Jin JL, Wei J, Zhao LL, et al. The Ubiquitination of RagA GTPase by RNF152 Negatively Regulates mTORC1 Activation. Mol Cell. 2015;58(5):804-18. 10.1016/j.molcel.2015.03.033

30. Kadoya M, Sasai N. Negative Regulation of mTOR Signaling Restricts Cell Proliferation in the Floor Plate. Front Neurosci. 2019;13:1022. 10.3389/fnins.2019.01022

31. Gangloff YG, Mueller M, Dann SG, Svoboda P, Sticker M, Spetz JF, et al. Disruption of the mouse mTOR gene leads to early postimplantation lethality and prohibits embryonic stem cell development. Mol Cell Biol. 2004;24(21):9508-16. 10.1128/MCB.24.21.9508-9516.2004

32. Murakami M, Ichisaka T, Maeda M, Oshiro N, Hara K, Edenhofer F, et al. mTOR is essential for growth and proliferation in early mouse embryos and embryonic stem cells. Mol Cell Biol. 2004;24(15):6710-8. 10.1128/mcb.24.15.6710-6718.2004

33. Laplante M, Sabatini DM. mTOR signaling at a glance. J Cell Sci. 2009;122(20):3589-94. 10.1242/jcs.051011

34. Kim J, Guan KL. mTOR as a central hub of nutrient signalling and cell growth. Nat Cell Biol. 2019;21(1):63-71. 10.1038/s41556-018-0205-1

35. Zhang Z, Chen ZT, Diao SQ, Ye SP, Wang JY, Gao N, et al. Identifying the complex genetic architecture of growth and fatness traits in a Duroc pig population. J Integr Agric. 2021;20(6):1607-14. 10.1016/S2095-3119(20)63264-6

36. Kools P, Van Imschoot G, van Roy F. Characterization of three novel human cadherin genes (CDH7, CDH19, and CDH20) clustered on chromosome 18q22-q23 and with high homology to chicken cadherin-7. Genomics. 2000;68(3):283-95. 10.1006/geno.2000.6305

37. Lin J, Wang C, Redies C. Restricted expression of classic cadherins in the spinal cord of the chicken embryo. Front Neuroanat. 2014;8:18. 10.3389/fnana.2014.00018

38. Gong H, Xiao S, Li W, Huang T, Huang X, Yan G, et al. Unravelling the genetic loci for growth and carcass traits in Chinese Bamaxiang pigs based on a 1.4 million SNP array. J Anim Breed Genet. 2019;136(1):3-14. 10.1111/jbg.12365

39. Wei C, Zeng H, Zhong Z, Cai X, Teng J, Liu Y, et al. Integration of non-additive genome-wide association study with a multi-tissue transcriptome analysis of growth and carcass traits in Duroc pigs. Animal. 2023:100817.

40. Che L, Yang Z, Xu M, Xu S, Che L, Lin Y, et al. Maternal nutrition modulates fetal development by inducing placental efficiency changes in gilts. BMC Genom. 2017;18(1):213. 10.1186/s12864-017-3601-1

41. Xu Z, Jin X, Cai W, Zhou M, Shao P, Yang Z, et al. Proteomics Analysis Reveals Abnormal Electron Transport and Excessive Oxidative Stress Cause Mitochondrial Dysfunction in Placental Tissues of Early-Onset Preeclampsia. Proteomics Clin Appl. 2018;12(5):e1700165. 10.1002/prca.201700165

42. Kunzelmann K, Tian YM, Martins JR, Faria D, Kongsuphol P, Ousingsawat J, et al. Airway epithelial cells-Functional links between CFTR and anoctamin dependent Cl- secretion. Int J Biochem Cell Biol. 2012;44(11):1897-900. 10.1016/j.biocel.2012.06.011

43. Feng M, Wang Z, Liu Z, Liu D, Zheng K, Lu P, et al. The RyR–ClCa–VDCC axis contributes to spontaneous tone in urethral smooth muscle. J Cell Physiol. 2019;234(12):23256-67.

44. Chen D, Meng W, Shu L, Liu S, Gu Y, Wang X, Feng M. ANO1 in urethral SMCs contributes to sex differences in urethral spontaneous tone. Am J Physiol Renal Physiol. 2020;319(3):F394-F402. 10.1152/ajprenal.00174.2020

45. Bergamaschi M, Maltecca C, Fix J, Schwab C, Tiezzi F. Genome-wide association study for carcass quality traits and growth in purebred and crossbred pigs. J Anim Sci. 2020;98(1):skz360. ARTN skz360 10.1093/jas/skz360

46. Lee J, Kang JH, Kim JM. Bayes Factor-Based Regulatory Gene Network Analysis of Genome-Wide Association Study of Economic Traits in a Purebred Swine Population. Genes (Basel). 2019;10(4):293. 10.3390/genes10040293

47. Sun WX, Wang HH, Jiang BC, Zhao YY, Xie ZR, Xiong K, Chen J. Global comparison of gene expression between subcutaneous and intramuscular adipose tissue of mature Erhualian pig. Genet Mol Res. 2013;12(4):5085-101. 10.4238/2013.October.29.3

48. Montero A, Okada Y, Tomita M, Ito M, Tsurukami H, Nakamura T, et al. Disruption of the fibroblast growth factor-2 gene results in decreased bone mass and bone formation. J Clin Invest. 2000;105(8):1085-93. 10.1172/JCI8641

49. Wang Z, Guo J, Guo Y, Yang Y, Teng T, Yu Q, et al. Genome-Wide Detection of CNVs and Association With Body Weight in Sheep Based on 600K SNP Arrays. Front Genet. 2020;11:558. 10.3389/fgene.2020.00558

50. Wong W, Scott JD. AKAP signalling complexes: focal points in space and time. Nat Rev Mol Cell Biol. 2004;5(12):959-70. 10.1038/nrm1527

51. Casiro S, Velez-Irizarry D, Ernst CW, Raney NE, Bates RO, Charles MG, Steibel JP. Genome-wide association study in an F2 Duroc x Pietrain resource population for economically important meat quality and carcass traits. J Anim Sci. 2017;95(2):545-58. 10.2527/jas.2016.1003

52. Zofkie W, Southard SM, Braun T, Lepper C. Fibroblast growth factor 6 regulates sizing of the muscle stem cell pool. Stem Cell Rep. 2021;16(12):2913-27. 10.1016/j.stemcr.2021.10.006

53. Armand AS, Laziz I, Chanoine C. FGF6 in myogenesis. Biochim Biophys Acta. 2006;1763(8):773-8. 10.1016/j.bbamcr.2006.06.005

458 54. Doukas J, Blease K, Craig D, Ma C, Chandler LA, Sosnowski BA, Pierce GF. Delivery of
459 FGF genes to wound repair cells enhances arteriogenesis and myogenesis in skeletal muscle.
460 Mol Ther. 2002;5(5 Pt 1):517-27. 10.1006/mthe.2002.0579

461 55. Srikanchai T, Murani E, Phatsara C, Schwerin M, Schellander K, Ponsuksili S, Wimmers K.
462 Association of ZYX polymorphisms with carcass and meat quality traits in commercial pigs.
463 Meat Sci. 2010;84(1):159-64. 10.1016/j.meatsci.2009.08.042

464 56. Macalma T, Otte J, Hensler ME, Bockholt SM, Louis HA, Kalff-Suske M, et al. Molecular
465 characterization of human zyxin. J Biol Chem. 1996;271(49):31470-8.
466 10.1074/jbc.271.49.31470

467 57. Nix DA, Fradelizi J, Bockholt S, Menichi B, Louvard D, Friederich E, Beckerle MC.
468 Targeting of zyxin to sites of actin membrane interaction and to the nucleus. J Biol Chem.
469 2001;276(37):34759-67. 10.1074/jbc.M102820200

470 58. Yoshigi M, Hoffman LM, Jensen CC, Yost HJ, Beckerle MC. Mechanical force mobilizes
471 zyxin from focal adhesions to actin filaments and regulates cytoskeletal reinforcement. J Cell
472 Biol. 2005;171(2):209-15. 10.1083/jcb.200505018

473 59. Hansen MD, Beckerle MC. Opposing roles of zyxin/LPP ACTA repeats and the LIM domain
474 region in cell-cell adhesion. J Biol Chem. 2006;281(23):16178-88. 10.1074/jbc.M512771200

475 60. Hoffman LM, Nix DA, Benson B, Boot-Hanford R, Gustafsson E, Jamora C, et al. Targeted
476 disruption of the murine zyxin gene. Mol Cell Biol. 2003;23(1):70-9. 10.1128/MCB.23.1.70-
477 79.2003

478 61. Sánchez-Roncancio C, García B, Gallardo-Hidalgo J, Yáñez JM. GWAS on Imputed Whole-
479 Genome Sequence Variants Reveal Genes Associated with Resistance to Piscirickettsia
480 salmonis in Rainbow Trout (Oncorhynchus mykiss). Genes. 2023;14(1):114.

481 62. Dean JM, He A, Tan M, Wang J, Lu D, Razani B, Lodhi IJ. MED19 regulates adipogenesis
482 and maintenance of white adipose tissue mass by mediating PPARγ-dependent gene
483 expression. Cell reports. 2020;33(1):108228.

484 63. Hwang JH, An SM, Kwon SG, Park DH, Kim TW, Kang DG, et al. Associations of the
485 polymorphisms in DHRS4, SERPING1, and APOR genes with postmortem pH in Berkshire
486 pigs. Animal biotechnology. 2017;28(4):288-93.

64. Zhao YY, Lin YQ, Xu YO. Functional Identification of Allograft Inflammatory Factor 1-Like Gene in Luning Chicken. Anim Biotechnol. 2018;29(3):234-40. 10.1080/10495398.2017.1369096

65. Rakita A, Nikolic N, Mildner M, Matiasek J, Elbe-Burger A. Re-epithelialization and immune cell behaviour in an ex vivo human skin model. Sci Rep. 2020;10(1):1. 10.1038/s41598-019-56847-4

66. Neustaeter A, Grossi D, Jafarikia M, Sargolzaei M, Schenkel F, editors. Genome-wide association study for loin marbling score in Canadian Duroc pigs. WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 10th; 2014.

67. Kokta T, Dodson M, Gertler A, Hill R. Intercellular signaling between adipose tissue and muscle tissue. Domestic animal endocrinology. 2004;27(4):303-31.

68. Yang W, Liu Z, Zhao Q, Du H, Yu J, Wang H, et al. Population Genetic Structure and Selection Signature Analysis of Beijing Black Pig. Front Genet. 2022;13:860669. 10.3389/fgene.2022.860669

69. Falker-Gieske C, Blaj I, Preuss S, Bennewitz J, Thaller G, Tetens J. GWAS for Meat and Carcass Traits Using Imputed Sequence Level Genotypes in Pooled F2-Designs in Pigs. G3 (Bethesda). 2019;9(9):2823-34. 10.1534/g3.119.400452

70. Naslavsky N, Rahajeng J, Rapaport D, Horowitz M, Caplan S. EHD1 regulates cholesterol homeostasis and lipid droplet storage. Biochem Biophys Res Commun. 2007;357(3):792-9. 10.1016/j.bbrc.2007.04.022