

Article

Integrative Multi-Omics Analysis Unveils Candidate Genes and Functional Variants for Growth and Reproductive Traits in Duroc Pigs

Zhuofan Yan ^{1,2,3,4,5}, Xiyue Li ⁶, Wenbo Yang ^{1,2,3,4,5}, Peng Zhou ¹ , Weiya Zhang ⁶, Xinyun Li ^{1,7}, Liangliang Fu ¹, Jingjin Li ¹ and Xiaoyong Du ^{1,2,3,4,5,*}

- ¹ Key Laboratory of Agricultural Animal Genetics, Breeding and Reproduction, Ministry of Education, College of Animal Science and Technology, Huazhong Agricultural University, Wuhan 430070, China; zfyang@webmail.hzau.edu.cn (Z.Y.); wbyang_9527@webmail.hzau.edu.cn (W.Y.); zhou-peng@webmail.hzau.edu.cn (P.Z.); xyli@mail.hzau.edu.cn (X.L.); fuliangliang2017@mail.hzau.edu.cn (L.F.); jingjinli@webmail.hzau.edu.cn (J.L.)
- ² College of Informatics, Huazhong Agricultural University, Wuhan 430070, China
- ³ Key Laboratory of Intelligent Technology in Animal Husbandry, Ministry of Agriculture and Rural Affairs, Wuhan 430070, China
- ⁴ Engineering Research Center of Smart Agricultural Technology, Ministry of Education, Wuhan 430070, China
- ⁵ Hubei Province Research Center of Engineering Technology of Agricultural Big Data, Wuhan 430070, China
- ⁶ College of Animal Science and Technology, Hebei Agricultural University, Baoding 071001, China; lixiyue198@163.com (X.L.); shyzy@hebau.edu.cn (W.Z.)
- ⁷ Hubei Hongshan Laboratory, Frontiers Science Center for Animal Breeding and Sustainable Production, Wuhan 430070, China
- * Correspondence: duxiaoyong@mail.hzau.edu.cn

Simple Summary

In pig breeding, improving carcass leanness, muscle yield, and reproductive performance is crucial for production efficiency. In this study, we investigated the genetic basis of three economically important traits—backfat thickness, loin muscle area, and total teat number—in 1624 Duroc pigs. Using genome-wide association and multi-omics analyses, we identified 21 significant genetic loci linked to these traits. Genes associated with fat deposition showed enrichment in adipose tissue regulatory regions, whereas genes influencing teat number were primarily active in intestinal tissues. For teat number, two key candidate genes, *ABCD4* and *YLPM1*, were highlighted, with functional variants predicted to alter transcription factor binding in hormone- and cholesterol-related pathways that are essential for mammary gland development. Our findings suggest potential molecular mechanisms underlying pig growth and reproduction and offer practical genetic markers to guide precision breeding for improved meat production and reproductive efficiency.

Abstract

Pigs play a vital role in global food security as a major source of animal protein. Enhancing growth and reproductive traits is of great economic importance to the swine industry. To systematically identify genetic determinants underlying key economic traits, we performed an integrative multi-omics analysis in a population of 1624 Duroc pigs, focusing on backfat thickness (BF), loin muscle area (LMA), and total teat number (TTN). Our genome-wide association study (GWAS) identified twenty-one significant single nucleotide polymorphisms (SNPs)—fourteen for BF, three for LMA, and seven for TTN. Candidate genes located within 1 Mb of these SNPs, such as *ZC3HAV1L* and *FAM3A* for BF, *PTGR2* for LMA, and *VRTN* and *ABCD4* for TTN, were further investigated. Functional genomic annotations revealed that genetic variants near the significant SNPs were enriched in tissue-specific enhancer elements, implying regulatory potential. Transcriptome-wide association study



Academic Editor: Bo Zhou

Received: 24 October 2025

Revised: 5 December 2025

Accepted: 15 December 2025

Published: 17 December 2025

Citation: Yan, Z.; Li, X.; Yang, W.; Zhou, P.; Zhang, W.; Li, X.; Fu, L.; Li, J.; Du, X. Integrative Multi-Omics Analysis Unveils Candidate Genes and Functional Variants for Growth and Reproductive Traits in Duroc Pigs. *Animals* **2025**, *15*, 3627. <https://doi.org/10.3390/ani15243627>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

(TWAS) further supported the candidate genes such as *ABCD4* and *YLPM1* for TTN and predicted several putative functional mutations that may affect transcription factor binding sites. This study demonstrates the power of integrative genomics to prioritize candidate genes and causal variants for animal complex traits, offering valuable resources for future precision breeding in pigs.

Keywords: GWAS; pig; economic traits; TWAS; candidate gene

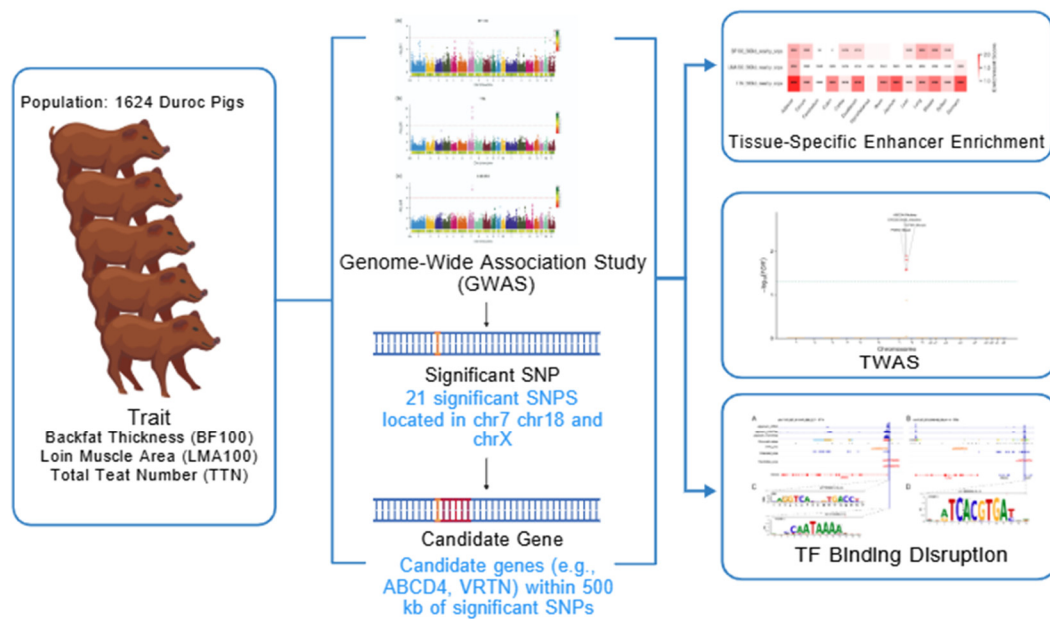
1. Introduction

Pigs are a foundation of global livestock production, and improving pork productivity remains essential for meeting increasing demand. Key economic traits—backfat thickness (BF), loin muscle area (LMA), and total teat number (TTN)—directly affect meat quality, carcass leanness, and sow productivity, thereby determining the efficiency and sustainability of swine production systems. These traits demonstrate moderate heritability (e.g., BF: ~0.37 [1,2]; LMA: 0.35–0.47 [3,4]) and are key breeding objectives in swine production, while TTN has been shown to play a critical role in regulating lactation capacity and piglet survival [5–7].

Genome-wide association studies (GWAS) have become a powerful tool for investigating such complex traits [8–11]. In pigs, GWAS has successfully identified numerous single nucleotide polymorphisms (SNPs) and candidate genes associated with growth and reproduction, such as *CACNA1E* and *ACBD6* for BF [12], and *VRTN* for TTN [13]. However, a major limitation persists: the majority of significant SNPs are located in non-coding regions, making it challenging to interpret their regulatory functions, identify causal mutations, and confirm effector genes [14–16]. The emergence of multi-omics resources offers a promising path forward. Functional genomic annotations, such as tissue-specific enhancer maps, can uncover the biological context in which non-coding variants function [17]. Simultaneously, transcriptome-wide association studies (TWAS) utilize expression quantitative trait loci (eQTL) to identify genes associated with traits, thereby establishing a direct link between genetic variation and gene function [18,19]. Platforms such as the FarmGTEx TWAS-Server have enabled such analyses across multiple major tissue types in pigs [20]. Although previous studies have provided valuable genetic insights into these traits [1,21,22], a systematic understanding of the key genes and their regulatory networks underlying BF, LMA, and TTN remains incomplete. This gap highlights the need to extend beyond GWAS through integrated multi-omics approaches.

Therefore, to address this need, we conducted an integrative multi-omics analysis in a large population of Duroc pigs. We first performed GWAS for backfat thickness adjusted to 100 kg body weight (BF100), loin muscle area adjusted to 100 kg body weight (LMA100), and TTN, followed by tissue-specific enhancer enrichment analysis to explore the regulatory context of the associated variants. We then integrated TWAS to identify functional genes influencing these traits and further combined this with motif analysis to predict potential causal regulatory variants (Figure 1). This study aims to identify novel genetic variants, key candidate genes, and their regulatory mechanisms, thereby providing a scientific foundation for precision breeding in pigs.

Integrative Genomics Reveals Mechanisms of Swine Economic Traits



Providing genetic resources and theoretical foundation for swine improvement

Figure 1. Workflow of the integrative multi-omics analysis. The schematic outlines the analytical pipeline. A population of 1624 Duroc pigs was phenotyped for BF100 (backfat thickness adjusted to 100 kg body weight), LMA100 (loin muscle area adjusted to 100 kg body weight) and TTN (total teat number). GWAS identified significant SNPs, from which candidate genes were selected. Convergent evidence from three complementary analyses—Tissue-Specific Enhancer Enrichment, TWAS, and TF Binding effect—was then integrated to prioritize high-confidence candidate genes and functional variants. This figure was created with BioGDP.com [23].

2. Materials and Methods

2.1. Data Collection and Phenotyping

2.1.1. Animal Population and Phenotypic Measurements

The study comprised a mixed cohort of 1624 Duroc pigs (964 males and 660 females) sourced from two commercial breeding companies in Guangxi and Hubei Provinces, China. All animals were raised under standardized conditions from 30 to 100 kg body weight, with uniform feed and management protocols to minimize environmental variability [24]. Backfat thickness (BF) and loin muscle area (LMA) at the 10th–11th rib were measured using an Aloka 500V SSD B ultrasound system (Corometrics Medical Systems, Wallingford, CT, USA) equipped with a linear transducer [25]. LMA values were quantified using proprietary image analysis software [26]. The measured BF and LMA were then adjusted to a standardized body weight of 100 kg (resulting in the traits BF100 and LMA100, respectively) using the correction formulas prescribed by the Chinese agricultural standard NY/T 2894-2016 [27].

2.1.2. Genotyping and Data Processing

Genomic DNA was extracted from ear tissue using the phenol-chloroform method and genotyped using the CC1 Porcine SNP50 BeadChip [28]. Since the reference genome for the chip data in this study was *Sus scrofa* version 10.2, we utilized the UCSC LiftOver tool (<https://genome.ucsc.edu/cgi-bin/hgLiftOver> (accessed on 1 September 2024)) to convert genomic coordinates from version 10.2 to 11.1 and checked allele consistency using bedtools [29]. The data were then converted to PLINK (v2.0) BED format for subsequent analyses.

2.2. Genome-Wide Association Study (GWAS)

2.2.1. Genotype Data Quality Control and Population Structure Analysis

To ensure data quality, we performed SNP quality control using PLINK (v2.0), applying parameters `--maf 0.01` and `--mind 0.1` to exclude low-quality SNPs. After filtering, 1624 individuals and 43,800 SNPs were retained. The Genomic Relationship Matrix (GRM) was constructed using GCTA v1.93.2beta software [30], followed by principal component analysis (PCA) to explore population structure.

2.2.2. Association Analysis

GWAS was performed using a mixed linear model (MLM) implemented in the R package MVP v1.4.5 [31]. To minimize potential confounding effects, several fixed covariates were included in the model, including sex, company, year, and month, which account for differences in biological characteristics and production management conditions. Population structure was controlled by incorporating the top five principal components (PCs) derived from PCA. The genome-wide significance threshold was set at $p < 1.14 \times 10^{-6}$ (Bonferroni correction: $0.05/43,800$).

2.3. SNP Enrichment Analysis

To explore the tissues and biological functions potentially involved with genetic loci associated with the target traits, we conducted an enrichment analysis using two publicly available pig regulatory genomics resources. First, given the relatively low density of the SNP-chip data used in our GWAS and the need to account for linkage disequilibrium (LD) effects, we adopted a candidate-variant extension strategy using the whole-genome sequencing (WGS) variant dataset reported by Li et al. [32]. For each genome-wide significant SNP, a ± 500 kb window upstream and downstream was defined, and all variants from the Li et al. [32] dataset falling within these windows were extracted to form an “extended candidate variant set.” Next, we evaluated whether these extended variants were enriched for regulatory annotations from Pan et al. [17], which mapped active enhancers across 14 major pig tissues. Enrichment analysis was performed using the Genomic Association Tester (GAT; <https://gat.readthedocs.io/>) (accessed on 20 October 2025)) [33].

2.4. Candidate Gene Identification and Functional Annotation

Significant SNPs ($p < 1.14 \times 10^{-6}$) were annotated using the Ensembl database (Release 110). Genes located within 500 kb upstream or downstream of significant SNPs were prioritized, while additional candidate genes related to pig growth and reproduction were supplemented through literature mining. Gene Ontology Biological Processes (GO BP) and KEGG pathway enrichment analyses for BF100, LMA100, and total teat number (TTN) were conducted using Metascape (<http://metascape.org>) (accessed on 20 October 2025)) with parameters: Min Overlap = 3, p -value cutoff = 0.01, and Min Enrichment = 1.5.

2.5. TWAS Analysis

To provide transcriptomic references for candidate genes and identify the primary tissues in which they act, cross-tissue transcriptome-wide association analysis (TWAS) was performed using the FarmGTEX platform (<https://www.farmgtex.org/>) (accessed on 10 September 2025)) [34,35]. GWAS summary statistics for BF100, LMA100, and TTN were used as input for imputation via the MetaXcan algorithm implemented on the FarmGTEX TWAS-Server. Expression reference models from 34 pig tissues (including adipose, liver, muscle, ovary, testis, brain, heart, and others) were integrated for cross-tissue analysis. Bonferroni-corrected significance thresholds were calculated as $-\log_{10}(0.05/N)$, where N is the total number of tested genes.

2.6. Identification of Candidate Regulatory Variants

To pinpoint putative regulatory variants that may drive GWAS associations, we implemented a stepwise screening strategy. For each significant SNP, we first determined whether any TWAS significant gene resided within a ± 500 kb flanking region. For each SNP meeting this criterion, we extracted all variants within a ± 500 kb window centered on the GWAS SNP from the WGS variant dataset reported by Li et al. [32]. From this set of variants, we retained only those that directly overlapped with the promoter or enhancer elements defined in the porcine functional genomic atlas of Pan et al. [17]. These overlapping variants were subjected to subsequent transcription factor (TF) motif analysis.

2.7. TF Motif Analysis

To evaluate the potential regulatory impact of the variants identified in Section 2.6, we performed TF motif analysis using the R package motifbreakR [36]. The variants obtained from the overlap-based filtering in Section 2.6 were used as input. For each variant, motifbreakR predicted its potential effect on TF binding. A p -value threshold of 1×10^{-4} was applied to identify statistically significant effects, and only those classified as having a “strong” impact were retained for further interpretation.

3. Results

3.1. Phenotypic and Genotypic Data Summary

Summary statistics for backfat thickness adjusted to 100 kg body weight (BF100), loin muscle area adjusted to 100 kg body weight (LMA100), and total teat number (TTN) are presented in Table 1. A total of 1546 individuals were phenotyped for BF100 (mean = 9.74 cm) and LMA100 (mean = 49.49 cm²), and 1235 individuals were phenotyped for TTN (mean = 13.64). Moderate-to-high coefficients of variation (CV) were observed for BF100 (19.11%) and LMA100 (21.43%), reflecting substantial phenotypic variability conducive to genetic analyses. In contrast, TTN exhibited limited variation (CV = 7.74%), suggesting relatively stable expression of this reproductive trait within the population.

Table 1. Descriptive statistics of phenotypic traits in the Duroc population.

Trait	Effective Record Count	Mean	Standard Deviation	Maximum	Minimum	Coefficient of Variation (%)
BF100 (cm)	1546	9.74	1.86	20.30	5.38	19.11
LMA100 (cm ²)	1546	49.49	10.61	71.64	25.76	21.43
TTN	1235	13.64	1.06	16.00	10.00	7.74

Genotype data were obtained from 1624 individuals (1546 for BF100 and LMA100, 1235 for TTN), with a total of 43,800 SNPs after quality control. SNP density plots for each trait are shown in Supplementary Figure S1. Principal component analysis (PCA) revealed clear clustering patterns among individuals (Supplementary Figure S2). For the BF100 and LMA100 datasets ($n = 1546$), samples separated into three distinct genetic clusters. In contrast, the TTN dataset ($n = 1235$) formed two primary clusters. These clusters corresponded closely to the two commercial origins of the pigs. This population structure was effectively controlled for in all subsequent association models.

3.2. Genome-Wide Association Analysis (GWAS) of BF100, LMA100, and TTN

GWAS using a mixed linear model identified 21 significant SNPs across the three traits after Bonferroni correction (Figure 2). The genomic inflation factors (λ) were 1.1125 for BF100, 1.0387 for LMA100, and 0.9643 for TTN, indicating adequate control of population structure.

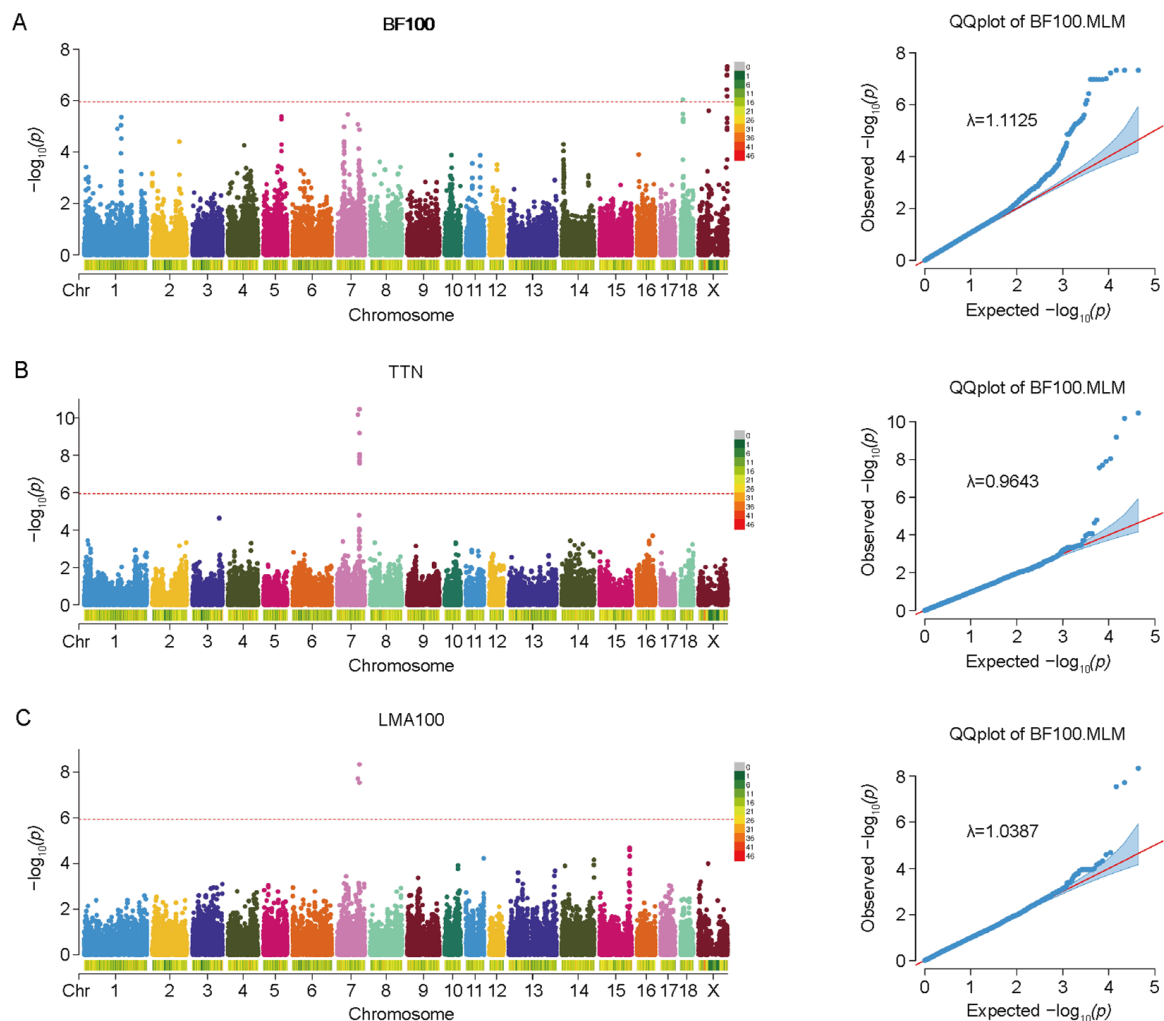


Figure 2. Manhattan plots and Q-Q plots of SNP dominance effects for three traits: (A) BF100 (backfat thickness adjusted to 100 kg body weight), (B) TTN (total teat number), and (C) LMA100 (loin muscle area adjusted to 100 kg body weight). The red dotted line indicates the potential significance threshold.

Across the three traits, a total of 21 genome-wide significant SNPs were identified—14 for BF100, 3 for LMA100, and 7 for TTN. For BF100, the 14 significant SNPs were predominantly located on the X chromosome, with the strongest signal observed at rs327767193 within the *FAM3A* gene. Three SNPs on chromosome 7 (Chr7:91308348:G:T, rs330032123, and rs711580469) were found to be associated with both LMA100 and TTN.

The strongest of these overlapping signals, rs711580469, mapped to the *VRTN* gene, a key regulator of vertebral development conserved across species [37,38]. Additionally, seven TTN-associated SNPs clustered on chromosome 7, including the lead SNP rs330032123 within *ABCD4*, flanked by *VSX2* and *VRTN*.

Candidate genes located within ± 0.5 Mb of these significant SNPs are summarized in Table 2.

Table 2. Summary of significant SNPs and candidate genes associated with BF100, LMA100, and TTN in GWAS.

Traits	SNP ID	Chrom	Position	p-Value	Candidate Genes
BF100	rs81236473	18	10555467	9.34×10^{-7}	ZC3HAV1L
	rs81333323	X	123188729	3.72×10^{-7}	CNGA2\GABRE
	rs332246267	X	123225920	1.06×10^{-7}	CNGA2\GABRE
	rs329206246	X	123251807	1.06×10^{-7}	CNGA2\GABRE
	rs336224043	X	123302262	1.06×10^{-7}	CNGA2\GABRE
	rs334203110	X	123655172	1.06×10^{-7}	GABRA3
	rs320267534	X	123827689	1.06×10^{-7}	BGN/SLC6A8
	rs80918182	X	123849200	1.06×10^{-7}	BGN/SLC6A8
	rs330196368	X	123955387	9.97×10^{-8}	BGN/SLC6A8
	rs326373823	X	124344831	6.02×10^{-8}	BGN/SLC6A8
	rs81245332	X	124482822	6.78×10^{-7}	BGN/SLC6A8
	rs330863063	X	124702511	4.72×10^{-8}	SLC6A8/FAM3A
	rs327767193	X	125021305	4.72×10^{-8}	FAM3A
	rs344761734	X	125135139	4.72×10^{-8}	FAM3A
LMA100/TTN	/	7	91308348	1.93×10^{-8}	PIGH
	rs330032123	7	97584287	2.91×10^{-8}	PTGR2/FAM161B/LIN52/ABCD4/VRTN/SYNDIG1L/LTBP2/AREL1/YLPM1/PROX2
	rs711580469	7	97622770	4.54×10^{-9}	PTGR2/FAM161B/LIN52/ABCD4/VRTN/SYNDIG1L/LTBP2/AREL1/YLPM1/PROX2/DLST
TTN	rs81396056	7	97732109	1.95×10^{-8}	ABCD4/NPC2/VRTN/SYNDIG1L/LTBP2/AREL1/YLPM1/PROX2/DLST
	rs80891836	7	98022168	1.25×10^{-8}	ABCD4/NPC2/VRTN/SYNDIG1L/LTBP2/AREL1/YLPM1/PROX2/DLST
	rs80836267	7	98089286	8.95×10^{-9}	VRTN/SYNDIG1L/LTBP2/AREL1/YLPM1/PROX2/DLST
	rs336641062	7	98116120	2.68×10^{-8}	VRTN/RPS6KL1/SYNDIG1L/LTBP2/AREL1/YLPM1/PROX2/DLST

3.3. Enrichment Analysis of Trait-Associated SNPs

To investigate the tissue-specific regulatory architecture underlying the target traits, we performed enrichment analyses using the extended candidate variant sets derived from our GWAS results. Each set was constructed by extending genome-wide significant SNPs by ± 500 kb and incorporating all variants from the whole-genome sequencing (WGS) variant dataset reported by Li et al. [32], thereby accounting for LD effects and capturing nearby regulatory variants. The functional relevance of these extended variants was then assessed by testing their overlap with enhancer annotations across 14 major pig tissues, as defined in the pig functional genome atlas reported by Pan et al. [17]. Enrichment was evaluated using the Genomic Association Tester (GAT; <https://gat.readthedocs.io/>) [33].

Distinct tissue enrichment patterns were observed for the three economic traits (Figure 3). Adipose tissue showed significant enrichment for all three traits, suggesting that regulatory variants influencing BF100, LMA100, and TTN are all partially mediated through adipose-related regulatory mechanisms.

The complete statistical details, including *p*-values and false discovery rates (FDR), are provided in Supplementary Table S1.

3.4. TWAS Reveals Functional Genes and Underlying Regulatory Variants for TTN

To identify genes whose expression is associated with TTN, we performed a transcriptome-wide association study (TWAS) using the FarmGTE_x TWAS-Server [20]. The analysis integrated expression prediction models across 34 tissues, including metabolic, reproductive, neural, and systemic tissue types (see Methods). The cross-tissue TWAS Manhattan plot for TTN is shown in Figure 4, highlighting significant genes and their pri-

many associated tissues. The analysis identified four genes significantly associated with the TTN trait: *ABCD4*, *ERG28*, *PSEN1*, and *YLPM1*. The association for *ABCD4* was detected in pituitary tissue, a central endocrine gland that regulates mammary development via hormones such as prolactin [39], suggesting that genetic variation influencing *ABCD4* expression may affect teat development via this endocrine pathway. Similarly, the association of *YLPM1* with morula—an early embryonic stage—suggests that its genetic regulation may be established during early development and exert lasting effects on processes relevant to later mammary morphogenesis. These TWAS signals were used to prioritize candidate regulatory variants for further functional investigation (Figure 5).

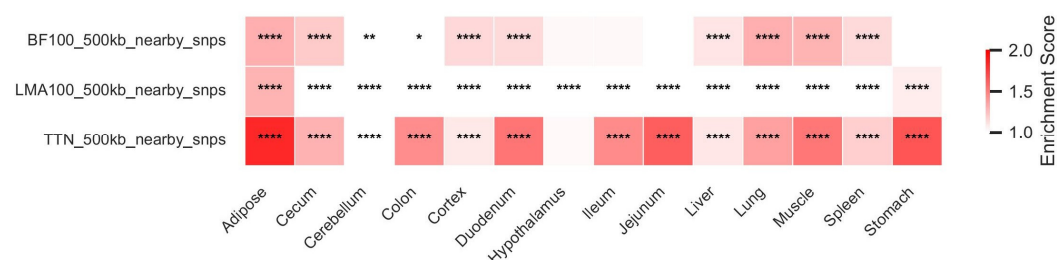


Figure 3. Enrichment of genetic variants for pig economic traits in tissue-specific enhancers. The heatmap displays the enrichment scores of candidate regions (within ± 500 kb of significant SNPs) associated with backfat thickness (BF100), loin muscle area (LMA100), and total teat number (TTN) across enhancers from 17 distinct tissues. Enrichment analysis was performed using the Genomic Association Tester (GAT). The significance markers in the cells are categorized based on p -values: **** (p -values ≤ 0.0001), ** (p -values ≤ 0.01), * (p -values ≤ 0.05), with no marker indicating no significant enrichment. The full set of enrichment statistics is available in Supplementary Table S1.

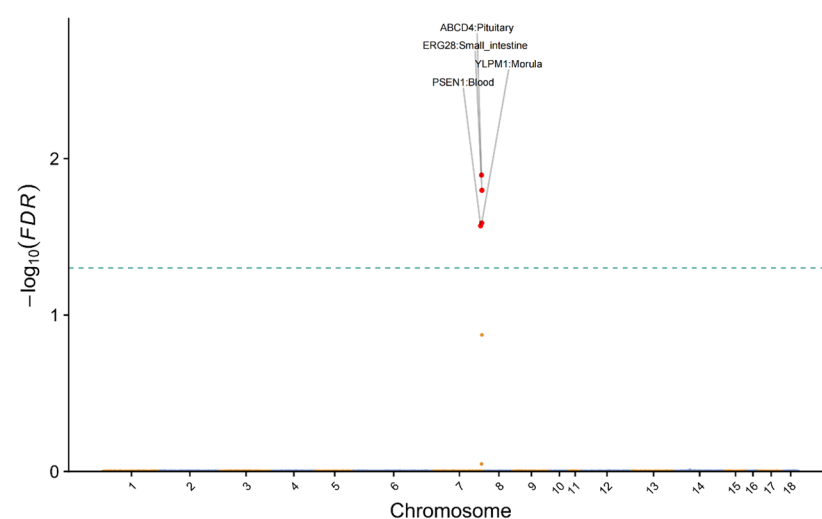


Figure 4. Manhattan plot of transcriptome-wide association study (TWAS) for total teat number (TTN). The y -axis indicates the association significance ($-\log_{10}(\text{FDR})$). The dashed line indicates the multiple testing-corrected significance threshold ($\text{FDR} < 0.05$).

For the *ABCD4* gene, which is involved in vitamin B12 metabolism [40]—a pathway with implications for lipid metabolism [41]—we identified a candidate regulatory variant, rs3472164889, within a promoter region marked by active epigenomic modifications in the jejunum (Figure 5A). Motif analysis predicted that this variant may affect the binding of transcription factor ESR1 (Figure 5C). Experimental studies have shown that reduced ESR1 expression in the mammary gland correlates with alterations in glandular morphology [42].

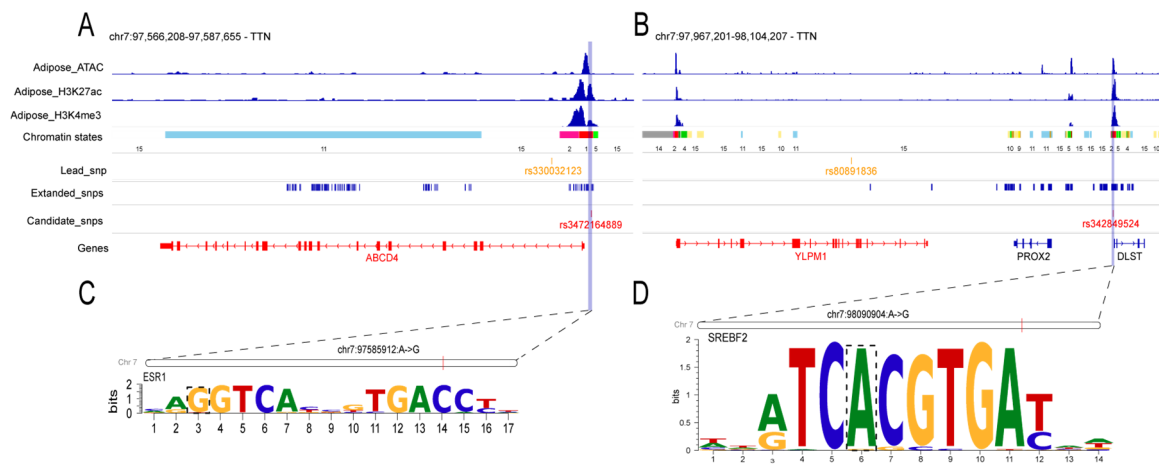


Figure 5. Predicted regulatory mechanisms and candidate functional mutations for TTN-associated genes. (A,B) Genomic views of two loci on chromosome 7 show adipose epigenomic profiles and GWAS SNPs, identifying *ABCD4* and *YLPM1* as candidate genes. The top four tracks display epigenomic profiles and chromatin state annotations (ChromHMM 15-state model) from Pan et al. [17], where numeric states (1–15) correspond to active promoters (TssA, TssAHet), transcriptional regions (TxFlnk, TxFlnkWk, TxFlnkHet), enhancers (EnhA, EnhAMe, EnhAWk, EnhAHet, EnhPois), open chromatin (ATAC_Is), bivalent or repressed promoters (TssBiv, Repr, ReprWk), and quiescent regions (Qui). (C,D) Transcription factor (TF) motif analysis predicts candidate SNPs alter binding of key transcription factors: ESR1 at the *ABCD4* promoter (C) and SREBF2 downstream of *YLPM1* (D).

For the *YLPM1* gene, our analysis highlighted a candidate functional mutation, rs342849524, located 48.62 kb downstream of the gene within a promoter region (Figure 5B). *YLPM1* plays a key role in cellular differentiation by repressing telomerase activity through binding to the TERT promoter [43]. Motif analysis predicted that this mutation may affect transcription factor SREBF2, which has been confirmed as a sterol-regulatory element-binding transcription factor with significant effects on lipid synthesis [44,45] (Figure 5D).

3.5. Functional Enrichment Analysis of Candidate Target Genes

To characterize the biological pathways underlying the identified candidate genes, Gene Ontology (GO) Biological Process and KEGG pathway enrichment analyses were performed for BF100-, LMA100-, and TTN-associated genes using Metascape (Figure 6). Genes associated with BF100 were significantly enriched in processes related to neurodevelopment, such as synapse assembly and axonogenesis, and in the MAPK signaling pathway, which has known roles in adipogenesis regulation. Genes associated with LMA100 were enriched in pathways involved in cellular metabolism and structural organization, including organic acid catabolic process, macroautophagy, and cilium organization. In contrast, TTN-associated genes were primarily enriched in cofactor biosynthesis and amino acid metabolic processes.

Collectively, these results reveal distinct biological processes and molecular pathways contributing to the regulation of growth and reproductive traits in pigs.

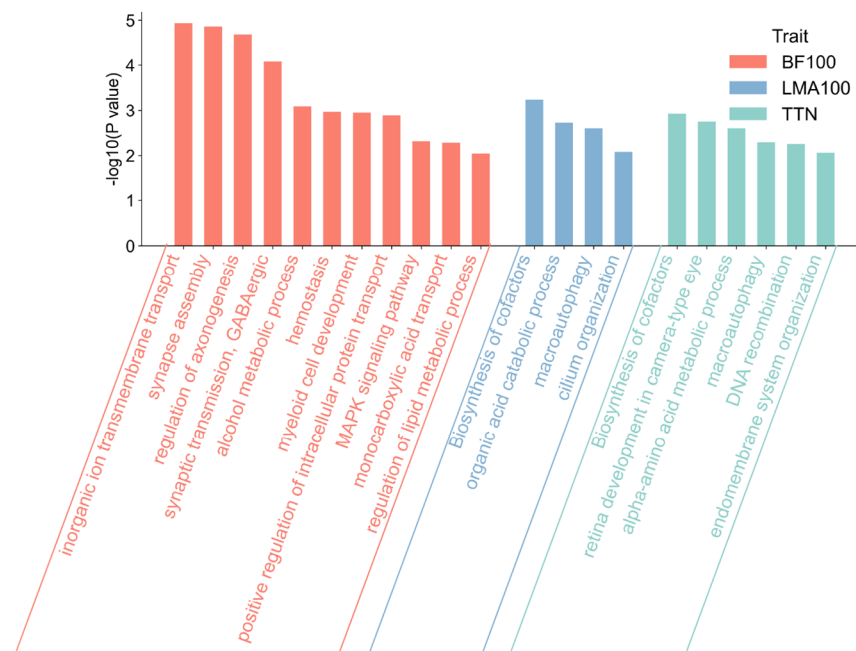


Figure 6. GO and KEGG enrichment analysis of candidate genes associated with BF100, LMA100, and TTN traits.

4. Discussion

Duroc pigs are a key genetic resource in global swine breeding, with traits such as backfat thickness (BF), loin muscle area (LMA), and total teat number (TTN) directly affecting production efficiency and economic returns. However, the functional regulatory mechanisms underlying these traits remain incompletely understood. In this study, we investigated their genetic basis by integrating genome-wide association studies (GWAS), SNP enrichment analysis, transcriptome-wide association studies (TWAS), and transcription factor (TF) motif analysis. Through this multi-level genomic approach, we aimed to identify key loci and functional regulatory elements underlying phenotypic variation in these traits, thereby providing a deeper understanding of their underlying genetic mechanisms.

Applying this approach, we investigated the functional context of significant associations. For backfat thickness adjusted to 100 kg body weight (BF100), GWAS signals near adipogenesis-related genes (*ZC3HAV1L*, *FAM3A*) [46,47] were further supported by their enrichment in adipose-specific enhancers. This suggests that the genetic influence on BF100 is mediated, at least in part, by variants that modulate regulatory activity specifically in adipose tissue. This moves beyond mere statistical association to implicate specific tissue-level regulatory mechanisms. Notably, genetic variants associated with BF100, loin muscle area adjusted to 100 kg body weight (LMA100), and TTN all showed enrichment in adipose tissue enhancers, suggesting adipose-derived signals may serve as a common regulatory node influencing these distinct traits.

Our TWAS and subsequent regulatory analyses nominated *ABCD4* and *YLPM1* as prioritized candidate genes for TTN. For *ABCD4*, the promoter variant rs3472164889 is predicted to affect the binding of ESR1, a key transcriptional regulator in mammary biology. Reduced ESR1 expression has been experimentally linked to altered mammary gland morphology [42]. Independently, functional studies indicate that *ABCD4* can modulate mammary epithelial cell proliferation and apoptosis, and upregulate the prolactin receptor (*PRLR*) [48]. For *YLPM1*, the downstream variant rs342849524 is predicted to influence the binding of SREBF2, a master transcriptional regulator of cholesterol and fatty acid synthesis [44,45]. Given that cholesterol is the essential precursor for all steroid hormones, genetic variation at this locus could modulate the metabolic substrate pool available for

hormone biosynthesis, thereby potentially influencing mammary gland development. Furthermore, *YLPM1* itself has been reported to regulate cellular differentiation, including through the repression of telomerase activity [43]. This suggests that *YLPM1* may also affect mammary development by modulating the proliferation or differentiation state of mammary epithelial cells, complementing its potential role in lipid-mediated hormonal regulation. Notably, the TWAS signal for *YLPM1* originated from morula, indicating that its genetic regulation is established early and may exert lasting influence on cell fate decisions pertinent to mammary morphogenesis.

Several limitations should be considered when interpreting our findings. First, the moderate density of the SNP chip limited fine-mapping resolution, and the absence of genotype imputation may have reduced the detection of trait-associated signals. Second, the lack of a population-matched, multi-tissue expression quantitative trait loci (eQTL) reference precluded formal colocalization analysis (e.g., using COLOC), which would be necessary to strengthen causal inference for TWAS-identified genes such as *ABCD4* and *YLPM1*. Future studies should prioritize: (1) high-depth whole-genome sequencing in expanded Duroc populations to improve mapping precision; (2) development of breed-specific multi-tissue eQTL resources; and (3) experimental validation, both in vitro and in vivo, of the functional roles of candidate genes identified in this study.

5. Conclusions

This study employed an integrative multi-omics approach to dissect the genetic architecture of key economic traits in Duroc pigs. A total of 21 significant SNPs and several candidate genes were identified, such as *FAM3A* and *ZC3HAV1L* for backfat thickness. Furthermore, by integrating transcriptome-wide association and motif analyses, *ABCD4* and *YLPM1* were implicated as candidate genes of teat number. Collectively, these findings provide a set of supported candidate genes and variants, establishing a foundation for future functional studies and molecular breeding applications in pigs.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ani15243627/s1>, Figure S1: SNP Density Plots for BF100, TTN, and LMA100 Traits; Figure S2: Principal Component Analysis (PCA) of Genomic Variation in Duroc Pigs. Table S1: Complete statistical results of the tissue-specific enhancer enrichment analysis for BF100, LMA100, and TTN.

Author Contributions: Conceptualization, project administration and funding acquisition, X.D.; conceptualization and project administration X.L. (Xinyun Li); methodology, validation, software and data curation, Z.Y., X.L. (Xiyue Li) and W.Y.; formal analysis and visualization, Z.Y. and W.Z.; investigation and writing—original draft preparation, Z.Y. and P.Z.; writing—review and editing, resources, funding acquisition and supervision, L.F. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Biological Breeding-National Science and Technology Major Project (2024ZD04077), the National Natural Science Foundation of China (32302704), the National Key R&D Program of China (No. 2018YFD0502000), the National Key R&D Program of China (2021YFD1300800) and IAEA CRP project (CRPD31030).

Institutional Review Board Statement: The care and management of the experimental animals were conducted in compliance with the Hubei Province regulations on Experimental Animals, and approved in 2025 by the Ethics Committee (HZAUSW-2025-0031) of Huazhong Agricultural University, Wuhan City, Hubei Province, China.

Informed Consent Statement: Informed consent was obtained from the animals' owner involved in the study.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: We acknowledge the support from the breeding pig project presided over by Chousheng Liu and Xiaotian Qiu, staff members of the National Animal Husbandry Service (NAHS).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gozalo-Marcilla, M.; Buntjer, J.; Johnsson, M.; Batista, L.; Diez, F.; Werner, C.R.; Chen, C.-Y.; Gorjanc, G.; Mellanby, R.J.; Hickey, J.M.; et al. Genetic Architecture and Major Genes for Backfat Thickness in Pig Lines of Diverse Genetic Backgrounds. *Genet. Sel. Evol.* **2021**, *53*, 76. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Johnson, Z.B.; Nugent, R.A., III. Heritability of Body Length and Measures of Body Density and Their Relationship to Backfat Thickness and Loin Muscle Area in Swine. *J. Anim. Sci.* **2003**, *81*, 1943–1949. [\[CrossRef\]](#)
3. Zhuang, Z.; Li, S.; Ding, R.; Yang, M.; Zheng, E.; Yang, H.; Gu, T.; Xu, Z.; Cai, G.; Wu, Z.; et al. Meta-Analysis of Genome-Wide Association Studies for Loin Muscle Area and Loin Muscle Depth in Two Duroc Pig Populations. *PLoS ONE* **2019**, *14*, e0218263. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Hou, R.; Chen, L.; Liu, X.; Liu, H.; Shi, G.; Hou, X.; Zhang, R.; Yang, M.; Niu, N.; Wang, L.; et al. Integrating Genome-Wide Association Study with RNA-Sequencing Reveals HDAC9 as a Candidate Gene Influencing Loin Muscle Area in Beijing Black Pigs. *Biology* **2022**, *11*, 1635. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Bovo, S.; Ballan, M.; Schiavo, G.; Ribani, A.; Tinarelli, S.; Utzeri, V.J.; Dall’Olio, S.; Gallo, M.; Fontanesi, L. Single-Marker and Haplotype-Based Genome-Wide Association Studies for the Number of Teats in Two Heavy Pig Breeds. *Anim. Genet.* **2021**, *52*, 440–450. [\[CrossRef\]](#)
6. Moscatelli, G.; Dall’Olio, S.; Bovo, S.; Schiavo, G.; Kazemi, H.; Ribani, A.; Zambonelli, P.; Tinarelli, S.; Gallo, M.; Bertolini, F.; et al. Genome-Wide Association Studies for the Number of Teats and Teat Asymmetry Patterns in Large White Pigs. *Anim. Genet.* **2020**, *51*, 595–600. [\[CrossRef\]](#)
7. Zaalberg, R.M.; Chu, T.T.; Bovbjerg, H.; Jensen, J.; Villumsen, T.M. Genetic Parameters for Early Piglet Weight, Litter Traits and Number of Functional Teats in Organic Pigs. *Animal* **2023**, *17*, 100717. [\[CrossRef\]](#)
8. Abdellaoui, A.; Yengo, L.; Verweij, K.J.H.; Visscher, P.M. 15 Years of GWAS Discovery: Realizing the Promise. *Am. J. Hum. Genet.* **2023**, *110*, 179–194. [\[CrossRef\]](#)
9. Ding, R.; Zhuang, Z.; Qiu, Y.; Wang, X.; Wu, J.; Zhou, S.; Ruan, D.; Xu, C.; Hong, L.; Gu, T.; et al. A Composite Strategy of Genome-Wide Association Study and Copy Number Variation Analysis for Carcass Traits in a Duroc Pig Population. *BMC Genom.* **2022**, *23*, 590. [\[CrossRef\]](#)
10. Liu, L.; Zhou, J.; Chen, C.J.; Zhang, J.; Wen, W.; Tian, J.; Zhang, Z.; Gu, Y. GWAS-Based Identification of New Loci for Milk Yield, Fat, and Protein in Holstein Cattle. *Animals* **2020**, *10*, 2048. [\[CrossRef\]](#)
11. McCarthy, M.I.; Abecasis, G.R.; Cardon, L.R.; Goldstein, D.B.; Little, J.; Ioannidis, J.P.A.; Hirschhorn, J.N. Genome-Wide Association Studies for Complex Traits: Consensus, Uncertainty and Challenges. *Nat. Rev. Genet.* **2008**, *9*, 356–369. [\[CrossRef\]](#)
12. Long, Y.; Ruan, G.R.; Su, Y.; Xiao, S.J.; Zhang, Z.Y.; Ren, J.; Ding, N.S.; Huang, L.S. Genome-Wide Association Study Identifies QTLs for EBV of Backfat Thickness and Average Daily Gain in Duroc Pigs. *Genetika* **2015**, *51*, 371–378. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Park, J.; Do, K.-T.; Park, K.-D.; Lee, H.-K. Genome-Wide Association Study Using a Single-Step Approach for Teat Number in Duroc, Landrace and Yorkshire Pigs in Korea. *Anim. Genet.* **2023**, *54*, 743–751. [\[CrossRef\]](#)
14. Maurano, M.T.; Humbert, R.; Rynes, E.; Thurman, R.E.; Haugen, E.; Wang, H.; Reynolds, A.P.; Sandstrom, R.; Qu, H.; Brody, J.; et al. Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science* **2012**, *337*, 1190–1195. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Li, B.; Ritchie, M.D. From GWAS to Gene: Transcriptome-Wide Association Studies and Other Methods to Functionally Understand GWAS Discoveries. *Front. Genet.* **2021**, *12*, 713230. [\[CrossRef\]](#)
16. Orozco, G.; Schoenfelder, S.; Walker, N.; Eyre, S.; Fraser, P. 3D Genome Organization Links Non-Coding Disease-Associated Variants to Genes. *Front. Cell Dev. Biol.* **2022**, *10*, 995388. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Pan, Z.; Yao, Y.; Yin, H.; Cai, Z.; Wang, Y.; Bai, L.; Kern, C.; Halstead, M.; Chanthavixay, G.; Trakooljul, N.; et al. Pig Genome Functional Annotation Enhances the Biological Interpretation of Complex Traits and Human Disease. *Nat. Commun.* **2021**, *12*, 5848. [\[CrossRef\]](#)
18. Zhang, Y.; Wang, M.; Li, Z.; Yang, X.; Li, K.; Xie, A.; Dong, F.; Wang, S.; Yan, J.; Liu, J. An Overview of Detecting Gene-Trait Associations by Integrating GWAS Summary Statistics and eQTLs. *Sci. China Life Sci.* **2024**, *67*, 1133–1154. [\[CrossRef\]](#)
19. Shao, M.; Chen, K.; Zhang, S.; Tian, M.; Shen, Y.; Cao, C.; Gu, N. Multiome-Wide Association Studies: Novel Approaches for Understanding Diseases. *Genom. Proteom. Bioinform.* **2024**, *22*, qzae077. [\[CrossRef\]](#)

20. Zhang, Z.; Chen, Z.; Teng, J.; Liu, S.; Lin, Q.; Wu, J.; Gao, Y.; Bai, Z.; FarmGTEx Consortium; Li, B.; et al. FarmGTEx TWAS-Server: An Interactive Web Server for Customized TWAS Analysis. *Genom. Proteom. Bioinform.* **2025**, *23*, qzaf006. [[CrossRef](#)]
21. Luan, M.; Ruan, D.; Qiu, Y.; Ye, Y.; Zhou, S.; Yang, J.; Sun, Y.; Ma, F.; Wu, Z.; Yang, J.; et al. Genome-Wide Association Study for Loin Muscle Area of Commercial Crossbred Pigs. *Anim. Biosci.* **2023**, *36*, 861–868. [[CrossRef](#)]
22. 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* **2021**, *11*, 806. [[CrossRef](#)] [[PubMed](#)]
23. Jiang, S.; Li, H.; Zhang, L.; Mu, W.; Zhang, Y.; Chen, T.; Wu, J.; Tang, H.; Zheng, S.; Liu, Y.; et al. Generic Diagramming Platform (GDP): A Comprehensive Database of High-Quality Biomedical Graphics. *Nucleic Acids Res.* **2025**, *53*, D1670–D1676. [[CrossRef](#)]
24. Park, J.; Na, C.-S. Weighted Single-Step Genome-Wide Association Study to Reveal New Candidate Genes for Productive Traits of Landrace Pig in Korea. *J. Anim. Sci. Technol.* **2024**, *66*, 702–716. [[CrossRef](#)]
25. Suzuki, K.; Irie, M.; Kadowaki, H.; Shibata, T.; Kumagai, M.; Nishida, A. Genetic Parameter Estimates of Meat Quality Traits in Duroc Pigs Selected for Average Daily Gain, Longissimus Muscle Area, Backfat Thickness, and Intramuscular Fat Content. *J. Anim. Sci.* **2005**, *83*, 2058–2065. [[CrossRef](#)]
26. Suzuki, K.; Kadowaki, H.; Shibata, T.; Uchida, H.; Nishida, A. Selection for Daily Gain, Loin-Eye Area, Backfat Thickness and Intramuscular Fat Based on Desired Gains over Seven Generations of Duroc Pigs. *Livest. Prod. Sci.* **2005**, *97*, 193–202. [[CrossRef](#)]
27. NY/T 2894-2016; Measurement of Backfat Depth and Loin Muscle Area on Living Pig with B-Mode Ultrasound, China. China Agriculture Press: Beijing, China, 2016.
28. Li, L.-Y.; Xiao, S.-J.; Tu, J.-M.; Zhang, Z.-K.; Zheng, H.; Huang, L.-B.; Huang, Z.-Y.; Yan, M.; Liu, X.-D.; Guo, Y.-M. A Further Survey of the Quantitative Trait Loci Affecting Swine Body Size and Carcass Traits in Five Related Pig Populations. *Anim. Genet.* **2021**, *52*, 621–632. [[CrossRef](#)]
29. Quinlan, A.R.; Hall, I.M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* **2010**, *26*, 841–842. [[CrossRef](#)]
30. Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A Tool for Genome-Wide Complex Trait Analysis. *Am. J. Hum. Genet.* **2011**, *88*, 76–82. [[CrossRef](#)]
31. Yin, L.; Zhang, H.; Tang, Z.; Xu, J.; Yin, D.; Zhang, Z.; Yuan, X.; Zhu, M.; Zhao, S.; Li, X.; et al. rMVP: A Memory-Efficient, Visualization-Enhanced, and Parallel-Accelerated Tool for Genome-Wide Association Study. *Genom. Proteom. Bioinform.* **2021**, *19*, 619–628. [[CrossRef](#)] [[PubMed](#)]
32. Li, J.; Xiang, Y.; Zhang, L.; Qi, X.; Zheng, Z.; Zhou, P.; Tang, Z.; Jin, Y.; Zhao, Q.; Fu, Y.; et al. Enhancer-Promoter Interaction Maps Provide Insights into Skeletal Muscle-Related Traits in Pig Genome. *BMC Biol.* **2022**, *20*, 136. [[CrossRef](#)] [[PubMed](#)]
33. Heger, A.; Webber, C.; Goodson, M.; Ponting, C.P.; Lunter, G. GAT: A Simulation Framework for Testing the Association of Genomic Intervals. *Bioinformatics* **2013**, *29*, 2046–2048. [[CrossRef](#)]
34. Fang, L.; Teng, J.; Lin, Q.; Bai, Z.; Liu, S.; Guan, D.; Li, B.; Gao, Y.; Hou, Y.; Gong, M.; et al. The Farm Animal Genotype-Tissue Expression (FarmGTEx) Project. *Nat. Genet.* **2025**, *57*, 786–796. [[CrossRef](#)]
35. Teng, J.; Gao, Y.; Yin, H.; Bai, Z.; Liu, S.; Zeng, H.; Bai, L.; Cai, Z.; Zhao, B.; Li, X.; et al. A Compendium of Genetic Regulatory Effects across Pig Tissues. *Nat. Genet.* **2024**, *56*, 112–123. [[CrossRef](#)]
36. Coetzee, S.G.; Coetzee, G.A.; Hazelett, D.J. motifbreakR: An R/Bioconductor Package for Predicting Variant Effects at Transcription Factor Binding Sites. *Bioinformatics* **2015**, *31*, 3847–3849. [[CrossRef](#)]
37. Duan, Y.; Zhang, H.; Zhang, Z.; Gao, J.; Yang, J.; Wu, Z.; Fan, Y.; Xing, Y.; Li, L.; Xiao, S.; et al. VRTN is Required for the Development of Thoracic Vertebrae in Mammals. *Int. J. Biol. Sci.* **2018**, *14*, 667–681. [[CrossRef](#)] [[PubMed](#)]
38. Xu, M.; Huang, J.; Wang, K.; Zhang, Y.; Zhou, H.; Wang, F.; Han, J. Functional Mutations in the VRTN Gene Influence Growth Traits and Meat Quality in Hainan Black Goats. *Vet. Sci.* **2025**, *12*, 936. [[CrossRef](#)] [[PubMed](#)]
39. Hannan, F.M.; Elajnaf, T.; Vandenberg, L.N.; Kennedy, S.H.; Thakker, R.V. Hormonal Regulation of Mammary Gland Development and Lactation. *Nat. Rev. Endocrinol.* **2023**, *19*, 46–61. [[CrossRef](#)]
40. Kitai, K.; Kawaguchi, K.; Tomohiro, T.; Morita, M.; So, T.; Imanaka, T. The Lysosomal Protein ABCD4 Can Transport Vitamin B12 across Liposomal Membranes In Vitro. *J. Biol. Chem.* **2021**, *296*, 100654. [[CrossRef](#)]
41. Boachie, J.; Adaikalakoteswari, A.; Samavat, J.; Saravanan, P. Low Vitamin B12 and Lipid Metabolism: Evidence from Pre-Clinical and Clinical Studies. *Nutrients* **2020**, *12*, 1925. [[CrossRef](#)]
42. Gomez, A.L.; Altamirano, G.A.; Leturia, J.; Bosquiaz, V.L.; Muñoz-de-Toro, M.; Kass, L. Male Mammary Gland Development and Methylation Status of Estrogen Receptor Alpha in Wistar Rats Are Modified by the Developmental Exposure to a Glyphosate-Based Herbicide. *Mol. Cell. Endocrinol.* **2019**, *481*, 14–25. [[CrossRef](#)]
43. Blalock, W.L.; Piazzini, M.; Bavelloni, A.; Raffini, M.; Faenza, I.; D’Angelo, A.; Cocco, L. Identification of the PKR Nuclear Interactome Reveals Roles in Ribosome Biogenesis, mRNA Processing and Cell Division. *J. Cell. Physiol.* **2014**, *229*, 1047–1060. [[CrossRef](#)] [[PubMed](#)]
44. Bommer, G.T.; MacDougald, O.A. Regulation of Lipid Homeostasis by the Bifunctional SREBF2-miR33a Locus. *Cell Metab.* **2011**, *13*, 241–247. [[CrossRef](#)]

45. Fan, Y.-M.; Karhunen, P.J.; Levula, M.; Ilveskoski, E.; Mikkelsen, J.; Kajander, O.A.; Järvinen, O.; Oksala, N.; Thusberg, J.; Vihinen, M.; et al. Expression of Sterol Regulatory Element-Binding Transcription Factor (SREBF) 2 and SREBF Cleavage-Activating Protein (SCAP) in Human Atheroma and the Association of Their Allelic Variants with Sudden Cardiac Death. *Thromb. J.* **2008**, *6*, 17. [[CrossRef](#)] [[PubMed](#)]
46. Kang, T.; Peng, D.; Bu, G.; Gu, H.; Zhang, F.; Zhang, R.; Zhou, Y.; Xiong, Y.; Lei, M. Transcriptional Regulation Analysis of FAM3A Gene and Its Effect on Adipocyte Differentiation. *Gene* **2016**, *595*, 92–98. [[CrossRef](#)]
47. Ibragimov, E.; Pedersen, A.Ø.; Sloth, N.M.; Fredholm, M.; Karlskov-Mortensen, P. Identification of a Novel QTL for Lean Meat Percentage Using Imputed Genotypes. *Anim. Genet.* **2024**, *55*, 658–663. [[CrossRef](#)] [[PubMed](#)]
48. Guo, X.; Zhao, C.; Yang, R.; Wang, Y.; Hu, X. ABCD4 Is Associated with Mammary Gland Development in Mammals. *BMC Genom.* **2024**, *25*, 494. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.