Integrating Genome-Wide Association Study with RNA-Sequencing Reveals HDAC9 as a Candidate Gene Influencing Loin Muscle Area in Beijing Black Pigs

Renda Hou 1,†, Li Chen 2,†, Xiance Liu 3, Hai Liu 3, Guohua Shi 3, Xinhua Hou 1, Run Zhang 1, Man Yang 1, Naiqi Niu 1, Lixian Wang 1 and Longchao Zhang 1,*

1 Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China
2 Chongqing Academy of Animal Science, Chongqing 402460, China
3 Beijing Heiliu Animal Husbandry Technology Co., Ltd., Beijing 102211, China
* Correspondence: zhlchias@163.com
† These authors contributed equally to this work.

Simple Summary: In this study, the genomic variants associated with LMA in Beijing Black pigs revealed a total of 329 differentially expressed genes relating to LMA. Integrating a GWAS with RNA sequencing revealed HDAC9 as a crucial candidate gene for this trait. KEGG analysis using RNA-seq data indicated that the DEGs were mainly enriched in the JAK/STAT and oxytocin signaling pathways. We believe that our study makes a significant contribution to the literature because, although many candidate genes related to LMA have been reported, the molecular markers and candidate genes for LMA in Beijing Black pigs have not been identified.

Abstract: Loin muscle area (LMA) is an important meat production trait and plays a key role in determining carcass leanness. Genome-wide association study (GWAS) and RNA sequencing (RNA-seq) analysis were used to identify candidate LMA genes in Beijing Black pigs, a popular breed among consumers in northern China. Ten single nucleotide polymorphisms (SNPs) in sus scrofa chromosome (SSC) 9 were significantly associated with LMA. These SNPs were mapped to a 2.90 Mb (84.94–87.84 Mb) region. A total of 11 annotated genes were mapped on this region, namely MEOX2, CRPPA, SOSTDC1, LRRC72, ANKMY2, BZW2, TSPAN13, AGR2, AHR, SNX13, and HDAC9. In addition, RNA-seq analysis was performed between the high- and low-LMA groups, and 329 differentially expressed genes (DEGs) were identified. Further, Kyoto Encyclopedia of Genes and Genomes analysis based on DEGs revealed that the JAK/STAT signaling pathway and oxytocin signaling pathway may be responsible for LMA. Both GWAS and RNA-seq analysis identified the HDAC9 gene, indicating that it may be an important candidate gene affecting LMA in Beijing Black pigs. The findings provide valuable molecular insights into the mechanisms that influence LMA content in pigs, which can be utilized in targeted approaches to enhance meat quality and commercial profitability.

Keywords: Beijing Black pig; HDAC9; genome-wide association study; loin muscle area; transcriptome

1. Introduction

The loin muscle is one of the most economically valuable parts of pork. The loin muscle area (LMA) is an important factor in estimating lean meat percentage [1], which is a key indicator in the evaluation of carcass quality grading [2]. It is widely believed that genetics play the largest role in the development of the loin muscle. The heritability of LMA ranges from 0.35 to 0.47 [3,4], indicating that this trait can be improved by genetic methods. Given the importance of LMA in production, there have been several molecular studies on LMA traits in recent years, and a large number of candidate genes and single nucleotide polymorphisms (SNPs) have been identified. To date, 415 quantitative trait loci...
(QTLs) associated with LMA have been reported in the pig QTL database (http://www.animalgenome.org/cgi-bin/QTLdb/SS/index, accessed on 24 August 2022). Previous studies have identified a few molecular markers and candidate genes for pork LMA (e.g., FGF21 [5], FUBP3 [6], and DRGA0016148 [7]) using genome-wide association studies (GWAS). The above markers and candidate genes have played important roles in the genetic improvement of LMA.

Developed in the 1960s from Berkshire, Large White, and Chinese indigenous pig breeds (Dingxian, Shenzhou, Zhourxian, respectively) [8], the Beijing Black pig has become an important pig breed in northern China due to the good flavor of its meat. Although many candidate genes related to loin muscle area have been reported, the molecular markers and candidate genes for LMA in Beijing Black pigs have yet to be identified. Therefore, the aim of this study is to identify putative genetic loci or candidate genes for LMA in Beijing Black pigs using GWAS and RNA sequencing (RNA-seq) joint analysis.

2. Materials and Methods

2.1. Animals and LMA Measurement

In this study, a total of 650 Beijing Black pigs were raised by Beijing Hei6 Husbandry Technology Co., Ltd., Jintun, China, and slaughtered at 210 ± 40 days of age. The loin muscle area at the last rib was recorded using sulfuric acid paper and scanned using a Lenovo M700 scanner at 300 pixels. The actual loin muscle area was subsequently calculated by dividing the pixel value obtained by selecting the loin muscle area contour with PS2020 by the standard pixel value.

2.2. Genotyping and Quality Control

Using the salting-out procedure, DNA for genomic analysis was isolated from all samples [9]. All the animals were genotyped using the Illumina porcine 50 K BeadChip (Illumina, San Diego, CA, USA) containing 50,697 SNPs. In compliance with the quality control requirements, SNPs and samples with call-rates lower than 90 percent or SNPs with a minor allele frequency (MAF) lower than 0.05 across all samples genotyped with each chip were excluded. PLINK version 1.90 was used to perform these quality controls.

2.3. Genome-Wide Association Study

In order to conduct a GWAS, a single-locus, mixed linear model GWAS was employed (EMMAX: Efficient Mixed-Model Association eXpedited). It has been demonstrated that the EMMAX statistic test approach performs better than the principal component analysis and genomic control when sample structure (stratification and relatedness) is taken into account [10]. In our analysis, the sex and slaughter batch were used as fixed effects, and the slaughter age was used as a covariate. The genotype data were used in EMMAX to construct the n × n genetic matrix of identity between individuals. The model used can be expressed as

\[ y = X\beta + Z\mu + e \]

where \( y \) is n × 1 the vector of observed phenotypic values of the animals, \( X \) is n × f matrix of fixed effects, \( \beta \) is q × 1 vector representing coefficients of the fixed effects, \( Z \) is n × t matrix relating the instances of the random effect to LMA, \( \mu \) is random animal effects, and \( e \) is the residual effect.

2.4. Total RNA Isolation, Purification, and Quantification

The high (H) and low (L) LMA groups were made up of eight animals with a high LMA (42.04 cm²) and eight animals with a low LMA (24.20 cm²) (Table S1), and the mean age at slaughter for the H- and L-LMA groups was 210 ± 10 days. Following the manufacturer’s instructions, the TRIlzol reagent (Invitrogen, Carlsbad, CA, USA) was used to extract total RNA from Beijing Black pig longissimus dorsi muscles. An ND-1000 spectrophotometer (NanoDrop, Wilmington, DE, USA) was used to measure RNA concentration and quality. RNA integrity was assessed using an Agilent 21000 with RIN number >7.0.
2.5. cDNA Library Construction and Sequencing

Total RNA was purified using two rounds of purification. The RNA was purified using poly T oligo-attached magnetic beads, lysed into small divalent fragments under high-temperature conditions, reverse-transcribed to cDNA, and then used to synthesize U-labeled double stranded DNA with Escherichia coli DNA polymerase I, Rnase H, and dUTP. The A-base was added to the blunt ends of each strand and then ligated to Illumina multiplex barcode adapters, which included custom unique molecular identifiers for minimizing sequence-dependent bias and amplification noise [11]. U-labeled secondary strands were treated with heat-labeled UDG enzymes and amplified using polymerase chain reaction (PCR). The average insert size of the final cDNA library was 300 bp (±50 bp). Two-terminal sequencing was performed on an Illumina HiSeq 4000 (LC Bio, Hangzhou, China), according to the protocol recommended by the supplier.

2.6. Data Quality Control and Differential Gene Expression Data Analysis

The Illumina sequencing platform was used for this project. Prior to assembly, reads that were determined to be of poor quality (reads that included sequencing adaptors, sequencing primers, or a quality score of <20) were eliminated. The filtered reads were mapped to the porcine genome using HISAT2 v2.1.0 [12] (Sus scrofa 11.1). In order to convert the aligned file to BAM format, SAMtools v1.6.0 [13] was used. The raw counts of genes and transcript levels were calculated using StringTie [14]. DESeq2 was used for differential expression analysis [15]. Using the Benjamini–Hochberg method [16] with adjusted \( p < 0.05 \), genes with an expression fold change (FC) > 2 were considered to be differentially expressed genes (DEGs). Principal component analysis (PCA) and heatmaps were performed on the normalized gene expression data to visualize the expression intensity values of the resulting muscle area transcripts.

2.7. Gene Ontology (GO) Annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Analysis of DEGs

All DEGs were mapped to GO terms in the GO database [17] (http://www.geneontology.org, accessed on 22 July 2022). Based on the hypergeometric test, GO terms in DEGs that were significantly enriched were found, with \( p < 0.05 \) set as the norm. Only DEGs with a \( p < 0.05 \) in DESeq2 were used for further analysis in this study. Pathways in the KEGG database were found using the Kobas functional annotation tool (v3.0).

2.8. Quantitative Real-Time PCR (qRT-PCR) Validation

Quantitative real-time PCR was performed on 10 randomly chosen genes to ensure that the sequencing findings were accurate. Total RNA was reverse-transcribed according to the manufacturer’s instructions using a PrimeScript RT kit (TaKaRa, Kusatsu, Japan). The qRT-PCR was used to normalize the transcript levels of the examined genes to those of GAPDH and to compute the \( 2^{-\Delta\Delta C_t} \) values [18,19]. QuantStudio 7 Flex (ABI, USA) was used to perform qRT-PCR; the reactions were conducted in triplicate. Amplification conditions included 40 cycles of 95 °C for 5 s, 60 °C for 34 s, 95 °C for 15 s, 60 °C for 1 s, and 95 °C for 15 s. Three replications of each experiment were carried out. Primer 5.0 was used to create primers based on sequence data from the NCBI database (Table S2).

3. Results

3.1. Genome-Wide Association SNPs for LMA

The mean and standard deviations of LMA of the 650-strong Beijing black pig population were 31.5 cm\(^2\) and 4.69 cm\(^2\), respectively. A total of 35,401 SNPs were in the final dataset used for the GWAS after quality control (Table S3). For GWAS, a threshold of \( 1.04 \times 10^{-4} \) (3.69/35,401) was used. Figure 1 shows the Manhattan plot and the quantile-quantile (Q-Q) plot from this GWAS. When performing a GWAS, dividing the population into groups may yield false-positive results. The Q–Q plot demonstrates that there was no
clear systematic bias in the traits as a whole. From the GWAS on SSC9, 10 genome-wide important SNPs that affect LMA were identified (Table 1).

![Manhattan plot for LMA](image)

**Figure 1.** GWAS for LMA. (A) Manhattan plots for LMA. The red horizontal line indicates the FDR threshold (1.04 × 10⁻⁴). (B) Quantile–quantile plot of SNPs after quality control in genome-wide association analysis for LMA.

| Marker          | Position   | Rs          | p-Value       | Var (%) c |
|-----------------|------------|-------------|---------------|-----------|
| WU_10.2_9_93503815 | 84,937,219 | rs336943465 | 5.19 × 10⁻⁶   | 2.08      |
| ALGA0054101    | 85,233,732 | rs81413966  | 1.10 × 10⁻⁵   | 1.94      |
| ASGA0043938    | 85,259,596 | rs81413967  | 6.12 × 10⁻⁶   | 2.05      |
| ASGA0043959    | 86,791,976 | rs81414045  | 7.94 × 10⁻⁵   | 1.56      |
| MARC0073290    | 86,843,385 | rs81259452  | 5.91 × 10⁻⁵   | 2.06      |
| MARC0046912    | 86,880,793 | rs81238869  | 2.02 × 10⁻⁵   | 1.82      |
| MARC0065750    | 86,920,421 | rs81253220  | 9.91 × 10⁻⁶   | 1.96      |
| MARC0055652    | 87,156,018 | rs81245907  | 3.52 × 10⁻⁵   | 1.71      |
| ASGA0043969    | 87,237,648 | rs81414101  | 4.76 × 10⁻⁵   | 1.66      |
| MARC0069139    | 87,837,573 | rs81256749  | 1.01 × 10⁻⁴   | 1.51      |

*Data from Sus scrofa Build 11.1; Rs, reference SNP; Var (%), phenotypic variation explained by the SNP.

On SSC9, 10 SNPs displayed significant genome-wide associations with LMA. These SNPs were mapped to a 2.90 Mb (84.94–87.84 Mb) region. The strongest association of these SNPs was with MARC0069139 at 87.84 Mb. A total of 11 annotated genes were mapped on this region. They were mesenchyme homeobox 2 (MEOX2), CDP-L-ribitol pyrophosphorylase A (CRPPA), sclerostin domain containing 1 (SOSTDC1), leucine rich repeat containing 72 (LRRC72), ankyrin repeat and MYND domain containing 2 (ANKMY2), basic leucine zipper and W2 domains 2 (BZW2), tetraspanin 13 (TSPAN13), anterior gradient 2 (AGR2), aryl hydrocarbon receptor (AHR), sorting nexin 13 (SNX13), and histone deacetylase 9 (HDAC9). For MARC0069139, individuals with the AA genotype had a higher LMA than those with the GG genotype (p < 0.01, Figure 2A). The frequency of the A allele was higher than that of the G allele (Figure 2B).
**basic leucine zipper and W2 domains 2 (**BZW2), tetraspanin 13 (**TSPAN13), anterior gradient 2 (**AGR2), aryl hydrocarbon receptor (**AHR), sorting nexin 13 (**SNX13), and histone deacetylase 9 (**HDAC9). For MARC0069139, individuals with the AA genotype had a higher LMA than those with the GG genotype (**p < 0.01, Figure 2A). The frequency of the A allele was higher than that of the G allele (Figure 2B).

**Figure 2.** Identification of MARC0069139 using GWAS of loin muscle area. (A) The difference analysis among three genotypes of MARC0069139 in Beijing Black pigs. * **p < 0.05, ** **p < 0.01. (B) Allele frequencies of MARC0069139 in Beijing Black pigs.

### 3.2. RNA-Seq Data

Sixteen cDNA libraries from the H and L groups were sequenced, yielding 712,114,178 clean reads. An average of 44,507,136 reads were obtained for each sample (ranging from 37,346,508 to 52,970,650 reads). The quality values of Q20 and Q30 were 99.97% and 98.10 %, respectively. Total RNA-seq resulted in an average of 97.01% of reads that passed the quality control filters. An average of 97.03% of the reads were mapped to the reference genome (**Sus scrofa** 11.1). These results indicated that the sequencing data in this study were reliable and could be used for subsequent bioinformatics analysis (Table 2).

**Table 2.** RNA sequencing data analysis.

| Sample_Name | Raw_Reads   | Clean_Reads   | Q30 (%) | GC_Content (%) | Mapped (%) |
|-------------|-------------|---------------|---------|----------------|------------|
| H1          | 42,921,002  | 41,411,838    | 98.03   | 52.00          | 96.48      |
| H2          | 51,521,550  | 50,088,198    | 98.55   | 51.50          | 97.22      |
| H3          | 49,671,858  | 47,485,598    | 97.96   | 52.00          | 95.60      |
| H4          | 45,497,868  | 44,432,658    | 97.49   | 52.00          | 97.66      |
| H5          | 40,988,148  | 39,848,532    | 97.98   | 52.00          | 97.22      |
| H6          | 39,841,846  | 38,699,270    | 97.97   | 52.00          | 97.13      |
| H7          | 47,118,756  | 45,666,934    | 98.21   | 51.00          | 96.92      |
| H8          | 50,646,242  | 49,107,270    | 97.97   | 52.00          | 96.96      |
| L1          | 54,959,674  | 52,970,650    | 98.22   | 52.00          | 96.38      |
| L2          | 52,534,364  | 51,224,866    | 98.06   | 52.00          | 97.51      |
| L3          | 45,241,998  | 43,696,868    | 98.32   | 51.00          | 96.58      |
| L4          | 42,172,534  | 41,077,200    | 98.25   | 51.50          | 97.40      |
| L5          | 41,125,930  | 39,957,600    | 98.17   | 52.00          | 97.16      |
| L6          | 38,604,910  | 37,465,804    | 98.13   | 52.00          | 97.05      |
| L7          | 52,999,242  | 51,634,384    | 98.00   | 53.00          | 97.42      |
| L8          | 38,195,034  | 37,346,508    | 98.25   | 52.00          | 97.78      |
3.3. DEGs between Higher and Lower LMA Groups

Gene expression was detected using counts. We identified 329 DEGs associated with LMA in group H and L ($p < 0.05$, fold change $\geq 2$). Compared to the L group, 161 DEGs were upregulated and 168 DEGs were downregulated in group H (Figure 3, Table S4). Heat maps (Figure 4A) were generated and PCA (Figure 4B) was performed by normalizing the gene expression data. The samples in groups H and L were clearly clustered into two branches of high and low groups, and genes with similar expression patterns were clustered together, indicating significant intersample differences and good reproducibility.

![Advanced Volcano Plot](image)

**Figure 3.** Volcano Plot of DEGs in longissimus dorsi muscle comparing H (high) group and L (low) group. The x-axis shows the values of log2 (fold change), while the average expression values of log10 ($p$-Value) are displayed by the y-axis. The red and blue dots represent the significantly differentially expressed transcripts ($p$-Value < 0.05) comparing the L group, with red for upregulated genes and blue for downregulated genes. The gray dots indicate the transcripts with expression levels which are not statistically significant ($p$-Value > 0.05).
3.4. GO Enrichment and KEGG Pathway Analysis

To determine the functions of the DEGs, GO and KEGG enrichment analyses were carried out. The results of the GO analysis showed that the DEGs were mainly involved in the regulation of the defense response to bacteria, toll-like receptor signaling pathway, leukocyte mediated immunity, cell adhesion mediator activity, and chemokine receptor binding (Figure 5). In addition, some DEGs enriched in GO terms were correlated with myogenesis, such as negative regulation of myoblast fusion, muscle atrophy, and leptin receptor activity (Table S5).

The KEGG tool was used to systematically analyze cellular metabolic pathways and gene product functions to study the complex functions of genes. The 20 most enriched pathways \((p < 0.05)\) are shown in Figure 6. KEGG analysis showed that the DEGs were enriched in pathways related to myogenesis, including glycolipid metabolism, and the JAK/STAT, PI3K-Akt, adipocytokine, Rap1, and oxytocin signaling pathways (Table S6).

3.5. Validation of Gene Expression Using qRT-PCR

Ten DEGs (PVALB, GNG4, S100A1, ARRD3C3, IER3, DNAJB1, CITED2, HSPA6, CXCL10, and ADIPOQ) were used to determine the difference in gene expression between the H and L groups using qRT-PCR to validate the accuracy of transcriptome sequencing (Figure 7B). The qRT-PCR and RNA-seq data revealed a consistent pattern of gene expression (Figure 7A).
Figure 5. Most highly enriched GO terms of DEGs between H group and L group. The y-axis displays the number of DEGs, and the x-axis represents the GO terms. The bar colors correspond to different GO categories, with purple for molecular function, green for biological process, and orange for cellular component.
The column diagrams for Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of DEGs. The x-axis represents the numbers of DEGs. The y-axis represents the functions of pathways. The color depth represents p-value.

Figure 6. The column diagrams for Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of DEGs. The x-axis represents the numbers of DEGs. The y-axis represents the functions of pathways. The color depth represents p-value.

Figure 7. Comparison of gene expression. (A) Comparison of the fold change between RNA-seq and qRT-PCR. (B) Gene expression of H and L group by qPCR. * p < 0.05, ** p < 0.01.
4. Discussion

There have been several reports of quantitative trait loci (QTL) for LMA. For instance, Edwards et al. [20] used 510 Duroc × Pietran F2 animals genotyped for 124 microsatellite markers that were uniformly spaced across the genome and detected a significant QTL that affects LMA and the related backfat trait. Furthermore, on SSC12 the SRY-related HMG-box 15 (SOX15) locus was displayed as a QTL for LMA in GWAS using the Landrace × KNP F2 intercross populations [21]. Using the Meishan × Yorkshire F2 population, QTLs for LMA were mapped on SSC 1, 2, 4, 11, 12, and 14 [22]. Ponsuksili et al. identified an LMA-related QTL (SSC12, 90–97 cM) in the Duroc × Berlin Miniature Pig F2 population [23]. In the current study, QTL signals of LMA were mapped using GWAS on SSC9 of Beijing Black pigs. These results suggest that LMA-regulated genes exhibit significant population specificity.

On SSC9, previous studies have mapped QTLs for LMA in the 45.3–109.6 Mb region. Our study identified 10 genome-wide significant SNPs across 84.94–87.84 Mb, with regions of overlap with those reported in previous reports. This region contains 11 annotated genes. MEOX2 is a key gene in the regulatory framework controlling limb muscle development in vertebrate embryos [24]. The loss of MEOX2 results in smaller limb muscles that harbor reduced numbers of myofibers [25]. Mutations in the CRPPA gene can produce a deficit of functioning α-dystroglycan and damaged muscle fibers, which impairs skeletal muscle development, structure, and function [26]. In addition to generating dystroglycanopathies, CRPPA mutations have also been linked to limb girdle muscular dystrophy, Walker–Warburg syndrome, and muscle–eye–brain disease [27,28]. When SOSTDC1 is knocked down, it promotes the migration of linear morphea fibroblasts [29]. A previous study suggested that ANKMY2 plays a key role in the regulation of Shh signaling in vivo [30]. BZW2 has been shown to promote fibrosarcoma tumor growth [31]. Downregulation of BZW2 results in apoptosis and a reduction in cell cycle progression and proliferation [32]. TSPAN13 reduces proliferation and invasion and enhances apoptosis of breast cancer cells in vitro and in vivo [33]. AGR2 is a promoter of cancer cell proliferation, invasion and survival, chemotherapy resistance, metastasis, and tumor growth [34–38]. AHR enables cells to adapt to changing conditions by sensing compounds from the environment, diet, microbiome, and cellular metabolism, which plays an important role in development of and immunity to cancer [39–42]. SNX13 has been shown to be a good candidate for milk production traits in cattle [43]. HDAC9 is highly expressed in skeletal muscle [44].

In the current study, we identified a total of 329 DEGs relating to LMA. Several DEGs, such as adiponectin, C1Q and collagen domain containing (ADIPOQ), Interferon Alpha Inducible Protein 6 (IFI6), and leptin receptor (LEPR), were previously reported to affect muscle fiber growth and development in animals. One of the biologically active adipokines, ADIPOQ, is essential for mediating antidiabetic and anti-atherogenic actions [45]. ADIPOQ may also play a significant role in the regulation of muscle fiber type, as evidenced by the fact that it is expressed by skeletal muscle fibers and affects the types of muscle fiber [46]. When compared to other tissues, muscle and bone have higher levels of IFI6 expression, suggesting that this gene is most active in these tissues. Kayan et al. demonstrated that IFI6 was significantly associated with LMA using the F2 population of Duroc × Pietran pigs [47]. In previous studies, LEPR has been used as a candidate gene for LMA [48] and plays an important role in muscle atrophy [49]. According to the DEGs, the JAK/STAT and oxytocin signaling pathways (known to be associated with myofiber growth and development) were enriched. Studies have shown that the JAK/STAT pathway controls the myogenic development of adult satellite cells (MuSCs), a cell type that is essential for postnatal skeletal muscle growth and damage repair [50]. Inhibiting oxytocin signaling in young animals inhibits muscle regeneration, whereas systemic administration of oxytocin promotes muscle regeneration swiftly by increasing aged muscle stem cell activation/proliferation via the MAPK/ERK signaling pathway [51].

Of the 329 DEGs, HDAC9 was shown to be a candidate gene for LMA in both GWAS and RNA-seq analyses. To date, 18 histone deacetylases have been identified in mammals
and are divided into four classes. Among them, Class II HDACs (HDAC4, HDAC5, HDAC7, and HDAC9) are highly expressed in skeletal muscle and directly bind MEF2, repressing expression of MEF2-dependent genes [44]. The ability of MRFs to transform fibroblasts into myoblasts is one of their most notable and distinguishing features. When MEF2 is cotransfected with MRFs, this ability is considerably improved [52]. Previous studies have shown that in vitro and in vivo the HDAC9 gene is a direct transcriptional target of MEF2. HDAC9 proteins can bind to MEF2 proteins and inhibit their transcriptional activity. The transcriptional repressor, HDAC9, thereby generates a negative feedback loop in the muscle development transcriptional circuitry [53]. The transcriptome results indicated a lower expression of HDAC9 in the high LMA group compared to the low LMA group, consistent with findings of previous reports. HDAC9 must therefore play an important role in LMA in Beijing Black pigs.

5. Conclusions

This study describes the genomic variants associated with LMA in Beijing Black pigs. Our results suggest that MEOX2, CRPPA, SOSTDC1, LRRC72, ANKMY2, BZW2, TSPAN13, AGR2, AHR, SNX13, and HDAC9 on SSC9 may play important roles in affecting LMA. Furthermore, KEGG analysis using RNA-seq data indicated that the DEGs were mainly enriched in the JAK/STAT and oxytocin signaling pathways. Integrating GWAS with RNA-seq revealed that HDAC9 is a crucial candidate gene for this trait. The identified variations provide valuable molecular information that may be harnessed to increase LMA in pork. This study provides new insights into the molecular mechanisms that regulate LMA in pigs.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/biology11111635/s1: Table S1. Descriptive statistics of loin muscle area from Beijing Black Pig population for RNA sequencing; Table S2. The details of primer design for qPCR of mRNA; Table S3. Distribution of SNPs after quality control and average distances on each chromosome; Table S4. Statistics of differentially expressed genes between H group and L group by RNA-seq; Table S5. Functional annotation clustering of differentially expressed genes in the longissimus dorsi muscle from Beijing Black pigs; Table S6. Pathway analysis for differentially expressed genes.

Author Contributions: Conceptualization, R.H., L.C. and L.Z.; formal analysis, R.H., L.C., X.L., H.L. and L.Z.; investigation, R.H., L.C., G.S., X.H., R.Z., M.Y., N.N. and L.W.; methodology, L.C., X.L., H.L., G.S., X.H. and R.Z.; resources: L.W. and L.Z.; supervision: R.H., L.C., L.W. and L.Z.; writing—original draft: R.H. and L.Z.; writing—review and editing: R.H. and L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Agricultural Science and Technology Innovation Program (CAAS-ZDRW202006-04), National Key R&D Program of China (2021YFD1301101), National Swine Industry Technology System (CARS-35).

Institutional Review Board Statement: In accordance with criteria issued by the Council of China, all animals were slaughtered via electrical stunning and exsanguination at the Beijing Fifth Meat Processing Factory. The Chinese Academy of Agricultural Sciences (CAAS) (Beijing) Science Research Department approved the experiments with animals (No. IAS2022-38).

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2021) in the National Genomics Data Center (Nucleic Acids Res 2022), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA008225) which is publicly accessible at https://ngdc.cnbc.ac.cn/gsa.

Acknowledgments: We thank the members of Zhengkui Zhou (Institute of Animal Science, CAAS) for helpful discussions.

Conflicts of Interest: The authors declare no conflict of interest.
27. Gencpınar, P.; Uyanik, G.; Haspolat, S.; Oygür, N.; Duman, O. Clinical and Molecular Manifestations of Congenital Muscular Alpha-Dystroglycanopathy due to an ISPD Gene Mutation. *Neurophysiology* 2019, 51, 373–378. [CrossRef]

28. Song, D.; Xu, X.; Ge, L.; Chang, X.; Wei, C.; Liu, J.; Yang, H.; Qu, S.; Bao, X.; Toda, T.; et al. A splice site mutation c.1251G>A of ISPD gene is a common cause of congenital muscular dystrophy in Chinese patients. *Clin. Genet.* 2020, 97, 789–790. [CrossRef]

29. Badshah, I.I.; Brown, S.; Weibel, L.; Rose, A.; Way, B.; Seibene, N.; Imman, G.; Harper, J.; O'Shaughnessy, R.F.L. Differential expression of secreted factors SSTSDC1 and ADAMTS8 cause profibrotic changes in linear morphoea fibroblasts. *Br. J. Dermatol.* 2019, 180, 1135–1149. [CrossRef]

30. Saita, S.; Shirane, M.; Ishitani, T.; Shimizu, N.; Nakayama, K.I. Role of the ANKMY2-FKBP38 axis in regulation of the Sonic hedgehog (Shh) signaling pathway. *J. Biol. Chem.* 2014, 289, 25639–25654. [CrossRef]

31. Gençpınar, P.; Uyanık, G.; Haspolat, S.; Oygür, N.; Duman, Ö. Clinical and Molecular Manifestations of Congenital Muscular Dystrophy (CMD) in Turkish Patients. *J. Med. Genet.* 2011, 48, 5125–5135. [CrossRef]

32. Li, G.; Lu, A.; Chen, A.; Geng, S.; Xu, Y.; Chen, X.; Yang, J. BZW2/5MP1 acts as a promising target in hepatocellular carcinoma. *J. Cancer* 2021, 12, 5125–5135. [CrossRef]

33. Wang, Y.W.; Zhao, S.; Yuan, X.Y.; Liu, Y.; Zhang, K.; Wang, J.; Zhu, J.; Ma, R. miR-4732-5p promotes breast cancer progression by targeting TSPAN13. *J. Cell. Mol. Med.* 2013, 23, 2549–2557. [CrossRef]

34. Barraclough, D.L.; Platt-Higgins, A.; de Silva Rudland, S.; Barraclough, R.; Winstanley, J.; West, C.R.; Rudland, P.S. The metastasis-associated antigen 4 (MTA4) is a novel tumour suppressor. *Cancer Res.* 2009, 69, 3164–3171. [CrossRef]

35. Liu, D.; Rudland, P.S.; Sibson, D.R.; Platt-Higgins, A.; Barraclough, R. Human homologue of cement gland protein, a novel metastasis inducer associated with breast carcinomas. *Cancer Res.* 2005, 65, 3796–3805. [CrossRef]

36. Park, K.; Chung, Y.J.; So, H.; Kim, K.; Park, J.; Oh, M.; Jo, M.; Choi, K.; Lee, E.J.; Choi, Y.L.; et al. AGR2, a mucinous ovarian cancer marker, promotes cell proliferation and migration. *Exp. Mol. Med.* 2011, 43, 91–100. [CrossRef]

37. Zhang, J.S.; Gong, A.; Cheville, J.C.; Smith, D.I.; Young, C.Y.; Smith, C.D.; Young, C.; Papadopoulo, E.; Luna, R.E.; et al. Overexpression of elF5 or its protein mimic 5MP perturbs elF2 function and induces ATF4 translation through delayed re-initiation. *Nucleic Acids Res.* 2016, 44, 8704–8713. [CrossRef]

38. Dumartin, L.; Whiteman, H.J.; Weeks, M.E.; Hariharan, D.; Dmitrovic, B.; Iacobuzio-Donahue, C.A.; Brentnall, T.A.; Bronner, M.; Feakins, R.M.; Timms, J.F.; et al. AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of pathogenesis and survival. *Cancer Res.* 2013, 71, 7091–7102. [CrossRef]

39. Gomez, A.; Bindesbøll, C.; Satheesh, S.V.; Grimaldi, G.; Hutin, D.; MacPherson, L.; Ahmed, S.; Tamblyn, L.; Cho, T.; Nebb, H.I.; et al. Characterization of TCDD-inducible poly-ADP-ribose polymerase (TIPARP/ARTD14) catalytic activity. *Biochem. J.* 2018, 475, 3827–3846. [CrossRef]

40. MacPherson, L.; Tamblyn, L.; Rajendra, S.; Bralha, F.; McPherson, J.P.; Matthews, J. 2,3,7,8-Tetrachlorodibenzo-p-dioxin poly(ADP-ribose) polymerase (TiPARP, ARTD14) is a mono-ADP-ribosyltransferase and repressor of aryl hydrocarbon receptor transactivation. *Nucleic Acids Res.* 2013, 41, 1604–1621. [CrossRef]

41. Ema, M.; Ohé, N.; Suzuki, M.; Mimura, J.; Sogawa, K.; Ikawa, S.; Fuji-i-Kuriyama, Y. Dioxin binding activities of polymorphic forms of mouse and human aryl hydrocarbon receptors. *J. Biol. Chem.* 1994, 269, 27337–27343. [CrossRef]

42. Sadik, A.; Somarribas Patterson, L.F.; Öztürk, S.; Mohapatra, S.R.; Panitz, V.; Secker, P.F.; Pflander, P.; Loth, S.; Salem, H.; Pretzell, M.T.; et al. IL4I1 Is a Metabolic Immune Checkpoint that Activates the AHR and Promotes Tumor Progression. *Cancer Res.* 2020, 80, 1822, 1252–1270.e34. [CrossRef] [PubMed]

43. Rincón, G.; Islas-Trejo, A.; Casellas, J.; Ronin, Y.; Soller, M.; Lipkin, E.; Medrano, J.F. Fine mapping and association analysis of a quantitative trait locus for milk production traits on Bos taurus autosome 4. *J. Dairy Sci.* 2009, 92, 758–764. [CrossRef] [PubMed]

44. Basset-Duby, R.; Olson, E.N. Signaling pathways in skeletal muscle remodeling. *Annu. Rev. Biochem.* 2006, 75, 19–37. [CrossRef] [PubMed]

45. Yamauchi, T.; Hari, K.; Kubota, N.; Terauchi, Y.; Tobe, K.; Froguel, P.; Nagai, R.; Kadowaki, T. Dual roles of adiponecin/Acrp30 in vivo as an anti-diabetic and anti-atherogenic adipokine. *Curr. Drug Targets Immune Endocr. Metabol. Disord.* 2003, 3, 243–254. [CrossRef]

46. Krause, M.P.; Liu, Y.; Yu, V.; Chan, L.; Xu, A.; Ridell, M.C.; Sweeney, G.; Hawke, T.J. Adiponecin is expressed by skeletal muscle fibers and influences muscle phenotype and function. *Am. J. Physiol. Cell Physiol.* 2008, 295, C203–C212. [CrossRef]

47. Kayan, A.; Uddin, M.J.; Cinar, M.U.; Grosse-Brihnsen, C.; Phatsara, C.; Wimmers, K.; Ponsuksili, S.; Tesfaye, D.; Looft, C.; Juengst, H.; et al. Investigation on interferon alpha-inducible protein 6 (IFI6) gene as a candidate for meat and carcass quality in pig. *Meat Sci.* 2011, 88, 755–760. [CrossRef]

48. Hirose, K.; Ito, T.; Fukawa, K.; Arakawa, A.; Mikawa, S.; Hayashi, Y.; Tanaka, K. Evaluation of effects of multiple candidate genes (LEP, LEPR, MC4R, PIK3C3, and VRTN) on production traits in Duroc pigs. *Anim. Sci. J.* 2014, 85, 198–206. [CrossRef]

49. Yang, Q.; Yan, C.; Wang, X.; Gong, Z. Leptin induces muscle wasting in a zebrafish kras-driven hepatocellular carcinoma (HCC) model. *Dis. Model. Mech.* 2019, 12, dmm038240. [CrossRef]

50. Moresi, V.; Adamo, S.; Berghella, L. The JAK/STAT Pathway in Skeletal Muscle Pathophysiology. *Front. Physiol.* 2019, 10, 500. [CrossRef]
51. Elabd, C.; Cousin, W.; Upadhyayula, P.; Chen, R.Y.; Chooljian, M.S.; Li, J.; Kung, S.; Jiang, K.P.; Conboy, I.M. Oxytocin is an age-specific circulating hormone that is necessary for muscle maintenance and regeneration. *Nat. Commun.* **2014**, *5*, 4082. [CrossRef]

52. Molkentin, J.D.; Black, B.L.; Martin, J.F.; Olson, E.N. Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* **1995**, *83*, 1125–1136. [CrossRef]

53. Haberland, M.; Arnold, M.A.; McAnally, J.; Phan, D.; Kim, Y.; Olson, E.N. Regulation of HDAC9 gene expression by MEF2 establishes a negative-feedback loop in the transcriptional circuitry of muscle differentiation. *Mol. Cell. Biol.* **2007**, *27*, 516–525. [CrossRef] [PubMed]