

RESEARCH

Open Access



Identification of key LncRNAs and mRNAs associated with intramuscular fat in pig via WGCNA

Wenqiang Li¹, Suozhou Yang¹, Huixin Liu¹, Zhi Cao¹, Fei Xu¹, Chao Ning¹, Qin Zhang¹, Dan Wang^{1*} and Hui Tang^{1*}

Abstract

Background Intramuscular fat (IMF) not only directly affects the tenderness, juiciness, and overall flavour of meat but also plays a significant role in influencing consumer preferences for pork. Therefore, exploring key biomarkers that influence IMF deposition is highly important for breeding high-quality pork. IMF is a typical quantitative trait that is regulated by the interaction of multiple coding and noncoding RNAs. Traditional differential analysis methods typically focus on individual genes, making it difficult to identify key genes and their underlying mechanisms accurately. Weighted gene coexpression network analysis (WGCNA) is an efficient and accurate method for identifying and characterizing key pathways and genes associated with complex traits. Therefore, the aim of this study was to construct an mRNA–lncRNA coexpression network related to IMF using WGCNA to explore and identify potential candidate genes that influence IMF in pigs.

Results Full-length transcriptome sequencing was performed on 31 220-day-old Jiangquan black pigs raised in the same environment, and a gene expression matrix comprising 25,609 genes was constructed. Nine coexpression modules were identified through WGCNA, with the number of genes in these modules ranging from 33 to 3648. The magenta module ($\text{corr} = 0.7, P < 0.01$) and the turquoise module ($\text{corr} = -0.77, P < 0.01$) were significantly associated with IMF deposition. Hub genes in each module were identified on the basis of the screening criteria of $\text{GS} > 0.4$ and $\text{MM} > 0.8$. Combined with enrichment analysis and protein–protein interaction (PPI) analysis, 18 key mRNAs potentially related to IMF were selected: *CRKL*, *CBL*, *PDGFRB*, *DOCK1*, *YWHAH*, *HSP90AB1*, *LOC100524873*, *NDUFA1*, *NDUFA11*, *NDUFA12*, *NDUFA2*, *NDUFAB1*, *NDUFB10*, *NDUFB3*, *NDUFB7*, *NDUFS5*, *NDUFS6*, and *UQCRL10*. To explore the regulatory role of lncRNAs in the process of IMF deposition, we constructed an lncRNA–mRNA–pathway network on the basis of the relationships between lncRNAs and key mRNAs, as well as the results of Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. This network includes four key lncRNAs (*TGOLN2*, *LOC100521518*, *LOC100524915*, and *LOC100622481*) and predicts the potential mechanisms by which lncRNAs regulate IMF deposition.

Conclusions Through WGCNA, enrichment analysis, and PPI analysis, 18 mRNAs and four lncRNAs potentially involved in IMF deposition were identified, and the lncRNA regulatory pathways were preliminarily explored. Our

*Correspondence:

Dan Wang
wangd_18@163.com
Hui Tang
tanghui@sdau.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

findings provide new insights into the regulatory mechanisms of pig IMF deposition and lay the foundation for further exploration of the molecular mechanisms underlying pig fat deposition.

Keywords RNA-seq, Weighted gene coexpression network analysis (WGCNA), Intramuscular fat (IMF), Pigs, lncRNA, Protein–protein interaction (PPI)

Background

Pork is a major source of protein in the human diet, and as living standards improve, increasing attention is given to the quality of pork [1]. Intramuscular fat (IMF) is composed of fat tissue in the form of white spots or streaks within muscle fibre bundles. The content of IMF is closely related to the flavour and tenderness of pork, making it an important indicator for assessing pork quality [2]. However, in recent decades, the IMF levels in modern pig breeds have decreased because modern breeding technologies focus too much on increasing lean meat yield while neglecting improvements in the IMF content. This not only affects eating quality but also has a negative effect on the economic benefits of the pork industry [3]. Therefore, to better meet the demands of consumers and the market, studying the molecular mechanisms of IMF deposition in pigs has always been a focus of scientific research [4–6].

RNA sequencing (RNA-seq) is a next-generation sequencing (NGS)-based technique that provides low background noise and a high dynamic range, making it ideal for the quantitative analysis of gene expression [7]. In recent years, researchers have used RNA-seq technology to study the transcriptome of pork fat-related traits, identifying a large number of genes associated with fat deposition and metabolism [8, 9]. Li et al. [10], Zhao et al. [11] and Xing et al. [12] conducted differential gene expression analysis using transcriptomic data from different pig breeds, identifying a large number of genes associated with the variation in pork quality traits and changes in their regulation. For example, *FABP3* regulates lipid-responsive transcription factors and is involved in lipid-mediated transcriptional programs [10], and *FASN*, a key enzyme involved in fat deposition and fatty acid synthesis, is highly expressed in pigs with greater fat deposition ability [11]. As regulatory factors of gene expression, lncRNAs play an important regulatory role in fat deposition in pigs [13]. Wang et al. analysed the gene profiles of Laiwu pigs with different IMF contents. They identified 17 candidate lncRNAs, among which seven target genes, including *ERLEC1*, *RBM47*, and *MYOF*, were confirmed to be associated with the IMF content [14]. Zhao et al. established an lncRNA–miRNA–mRNA regulatory network for IMF by analysing the differential gene expression between Songliao black pigs and Landrace pigs [15]. They identified several potential target relationships, such as *MSTRG.19948.1*, which targets *LPIN1* and *HSP90AA1*, and *MSTRG.13120.1*, which targets *LPIN1*,

which may influence IMF deposition. However, IMF deposition in pigs is regulated by a complex network of interrelated genes and their products. Therefore, when studying the regulatory mechanisms of IMF deposition, it is essential to consider the interactions between multiple mRNAs as well as the relationships between mRNAs and lncRNAs. Traditional single-gene differential expression analyses often overlook valuable information, as they focus on the effects of individual genes rather than the broader gene networks [16]. Weighted gene coexpression network analysis (WGCNA) is an effective method for studying systems biology and plays a crucial role in revealing the characteristics of gene networks associated with complex traits [17]. It can reveal gene interactions at a systemic level and assist researchers in identifying regulatory hubs of coexpressed genes by revealing modules that are strongly correlated with target traits [18, 19]. WGCNA has been successfully applied to identify genes and biological processes involved in IMF deposition in pigs. Zappaterra et al. used WGCNA to analyse the gene expression profiles of Italian Landrace pigs with varying IMF contents [20]. They identified hub genes (*FGF2*, *ZNF518A*, and *U2SURP*) associated with IMF deposition, which are involved primarily in the regulation of DNA transcription, cell differentiation, primary cilium morphogenesis, and intracellular signalling cascades. However, research on constructing coexpression networks using both mRNA and lncRNA data is still relatively limited. In addition, constructing coexpression networks within the same breed helps better eliminate interbreed differences. Therefore, applying the WGCNA method to construct an mRNA–lncRNA coexpression network in the Jiangquan black pig population to explore the characteristic genes and molecular regulatory mechanisms of IMF deposition is of significant research value.

In this study, full-length transcriptome data from 31 Jiangquan black pigs were used to construct lncRNA–mRNA coexpression modules and identify key modules and hub genes related to IMF deposition. Through PPI analysis and enrichment analysis, an lncRNA–mRNA–pathway network associated with IMF deposition was constructed. This research provides a foundation for further exploration of the molecular regulatory mechanisms underlying IMF deposition and may offer theoretical support for screening feature genes related to IMF deposition and for the breeding of superior pig breeds.

Materials and methods

Samples and IMF measurement

In this study, 31 healthy Jiangquan black pigs (approximately 105 ± 5 kg, 15 males and 16 females) at 220 days of age, all of which were raised on the same pig farm in Linyi, China, were selected. None of the pigs were related, and they were allowed free access to food and raised under the same environmental conditions. The experimental pigs were slaughtered and sampled immediately after electric shock treatment at 220 days of age. Approximately 2 g of longissimus dorsi muscle (LDM) tissue was collected from the third lumbar vertebra of each pig and immediately preserved in liquid nitrogen for total RNA extraction. Additionally, approximately 200 g of LDM was collected for IMF content measurement. The IMF content was determined using the Soxhlet petroleum ether extraction method according to the “Technical Regulation for Determination of Pork Quality” (NY/T 821–2004, China). First, the meat sample was cut into small pieces, surface moisture was removed, and the sample was placed in an oven to dry at 105°C until it reached a constant weight. Next, the dried meat sample was ground into a powder using a grinder, and approximately 5–10 g of the ground sample was weighed and placed in an extraction flask. Then, a Soxhlet extractor was used to add an appropriate amount of petroleum ether, and continuous extraction was performed for approximately 8 h. After extraction, the sample was removed and dried at 80°C until a constant weight was reached, after which the extracted fat was weighed. Finally, the IMF content was calculated on the basis of the fat weight and the initial weight of the sample (Supplementary Table 1). All animals were treated in accordance with the “Guidelines for the Care and Use of Experimental Animals” of Shandong Agricultural University (No. SDUA-2021-062).

Total RNA isolation, library preparation, and sequencing

Total RNA was extracted using TRIzol reagent (Invitrogen, USA) according to the manufacturer’s protocol. The quality and concentration of the extracted RNA were assessed using an Agilent 2100 Bioanalyzer (Agilent, USA). Ribosomal RNA (rRNA) was then depleted using the Epicenter Ribo-Zero rRNA Removal Kit (Epicentre, USA) following the manufacturer’s instructions. CDNA libraries were constructed by degrading the second strand of cDNA containing uracil (U) with the USER enzyme (NEB, USA), followed by polymerase chain reaction (PCR) amplification. Clustering of the index-coded samples was performed on a cBot Cluster Generation System using the TruSeq PE Cluster Kit v3-cBot-HS (Illumina, USA) according to the manufacturer’s instructions. After cluster generation, the library preparations were sequenced on an Illumina NovaSeq 6000 platform (Illumina, USA), and 150 bp paired-end reads were generated.

Raw data quality control and transcript assembly

Clean reads were obtained by filtering the raw data using Fastp software [21]. The filtering process involved three steps: first, adapter contamination was removed; second, reads with more than 10% uncertain bases were discarded; and third, reads with more than 50% low-quality bases (Phred quality score < 20) were excluded. Indexing was performed using HISAT2 (v2.2.1) software [22] with the reference genome (Sscrofa11.1, https://www.ncbi.nlm.nih.gov/assembly/GCF_000003025.6/) obtained from the NCBI database. HISAT2 was then used to align the cleaned reads to the reference genome (`--dta --max-intronlen 5000000`), generating SAM format files that contained the read mapping information for each sample. The transcripts for each sample were assembled using StringTie (v2.2.1) [23], and the GTF files from 31 samples were merged into a nonredundant GTF file. The FPKM (reads per kilobase of the exon model per million mapped reads) values for all the genes across the samples were subsequently obtained using Cufflinks (v2.1.1) [24]. Finally, on the basis of the annotation information of the reference genome, gene expression profiles containing only mRNAs and lncRNAs were extracted for subsequent analysis.

Establishment of weighted gene coexpression networks and identification of modules

WGCNA was employed to perform correlation network analysis on large, high-dimensional datasets [25]. This method clusters highly correlated genes into the same modules and identifies complex relationships between modules and traits. Genes with low expression levels in living organisms are less likely to have biological functions [26]. Therefore, we excluded genes from the gene expression profile that had an average FPKM < 0.5 across the 31 individuals [27]. A one-step function is used to construct the network and identify consensus modules in this study. First, the gene expression matrix is converted into an adjacency matrix, and an unsupervised coexpression relationship is constructed on the basis of the Pearson correlation coefficient between gene pairs. The adjacency matrix is then weighted using a power exponent β (soft threshold), with the choice of β determined by the scale-free topology criterion. When the soft threshold is set to 3, the scale-free topology fit index starts to level off, indicating good network connectivity. Next, the adjacency matrix is transformed into a topological matrix, and the topological overlap matrix (TOM) is used to measure the similarity between gene pairs. Using 1-TOM, genes with similar expression profiles are clustered into gene modules via average linkage hierarchical clustering. The minimum number of genes in each module is set to 30, with the deepSplit parameter set to 3,

while the remaining parameters are kept at their default values.

Identification of modules associated with traits and hub genes in the key module

Correlation analysis between coexpression modules and the IMF content was performed on the basis of the relevant literature [28]. Module–trait relationships were assessed using Pearson's correlation, with a correlation coefficient (corr) greater than 0.7 and a *p* value less than 0.05 considered indicative of a significant relationship. Gene significance (GS) and module membership (MM) were calculated for each gene within the respective modules. GS reflects the relationship between gene expression levels and the IMF content, whereas MM indicates the correlation between a gene and the module's feature genes. When both GS and MM are highly correlated for a gene, the gene is not only strongly associated with the trait of interest but also likely a hub gene within the module that is related to that trait [29]. Among the modules significantly associated with IMF, genes with a GS value greater than 0.4 and an MM value greater than 0.8 were considered hub genes [30]. Visualization and analysis of all the hub lncRNAs and the top 30 hub mRNAs in key modules were performed using Cytoscape [31].

GO function and KEGG pathway analyses

Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analyses are widely utilized tools in bioinformatics. GO and KEGG enrichment analyses of all the mRNAs within the key modules were conducted using the DAVID database. The enrichment results were visualized using the ggplot2 (v2.2.0) package in R. Hub mRNAs and hub lncRNAs associated with IMF-related KEGG pathways were screened for further identification of key mRNAs.

Protein–protein interaction network construction and identification of key genes

The interactions between hub mRNAs in key modules were analysed using the STRING database to construct protein–protein interaction (PPI) networks. The “Analyze Network” plugin for Cytoscape was employed to calculate and visualize the gene degrees within the PPI network. The top-degree hub mRNAs were identified as key regulators of IMF deposition [32].

LncRNA target mRNA analysis and identification of key LncRNAs

The cis-target genes of lncRNAs are typically located in proximity to adjacent genes [33]. To identify the cis-target genes of lncRNAs, FEELnc (v0.1.1) was used to search for all coding genes located within 100 kb upstream or downstream of the hub lncRNAs [34]. Previous studies

have suggested that the trans-target genes of lncRNAs are often associated with their coexpressed genes [32]. The cor function in R was applied to calculate the Pearson correlation coefficient (PCC) between the hub lncRNAs and hub mRNAs within critical modules. LncRNAs play a significant role in regulating mRNA expression and are considered key lncRNAs when their target genes are critical mRNAs.

Construction of the LncRNA–mRNA–pathway networks

To investigate the potential regulatory mechanisms of lncRNAs, we constructed an lncRNA–mRNA–pathway network based on lncRNA–mRNA targeting pairs and key mRNAs involved in regulating the IMF-related pathway. The lncRNA–mRNA–pathway network was then visualized and analysed using Cytoscape software.

Results

Summary of sequencing data

We collected 31 LDM muscle tissue samples from Jiangquan black pigs for full-length transcriptome sequencing. An average of 132.02 million paired-end clean reads were obtained for each sample, with a Q30 value exceeding 90.85%, and the average mapping rate was greater than 94.22% (Supplementary Table 2). The final gene expression matrix of the 31 individuals included 20,195 mRNAs and 5,414 lncRNAs.

Construction of the weighted co-expression networks

After low-expression genes (FPKM < 0.5) were removed, 8,093 mRNAs and 198 lncRNAs were used in WGCNA to construct coexpression networks. Cluster analysis indicated that all samples were grouped into a single cluster, so there was no need to remove any outliers (Supplementary Fig. 1). The soft-thresholding power was set to 3, with a correlation coefficient threshold of 0.9 (Fig. 1). Modules were merged on the basis of the criterion of dissimilarity less than 0.25, with a minimum module size of 30 genes. As a result, 9 modules were identified, each represented by a distinct colour (Fig. 2). The number of genes in these modules ranged from 33 to 3,648 (Table 1).

Identification of key modules and hub genes

A Pearson correlation analysis was conducted to examine the relationships between module eigengenes and IMF traits (Fig. 3). The magenta module presented a significant positive correlation with the IMF content (corr = 0.7, *P* < 0.001), whereas the turquoise module presented a significant negative correlation (corr = -0.77, *P* < 0.001). On the basis of these findings, the magenta and turquoise modules were selected as key modules for subsequent analysis. For each gene within these modules, we performed a correlation analysis between MM and GS. As shown in the scatter plot (Fig. 4), MM was strongly

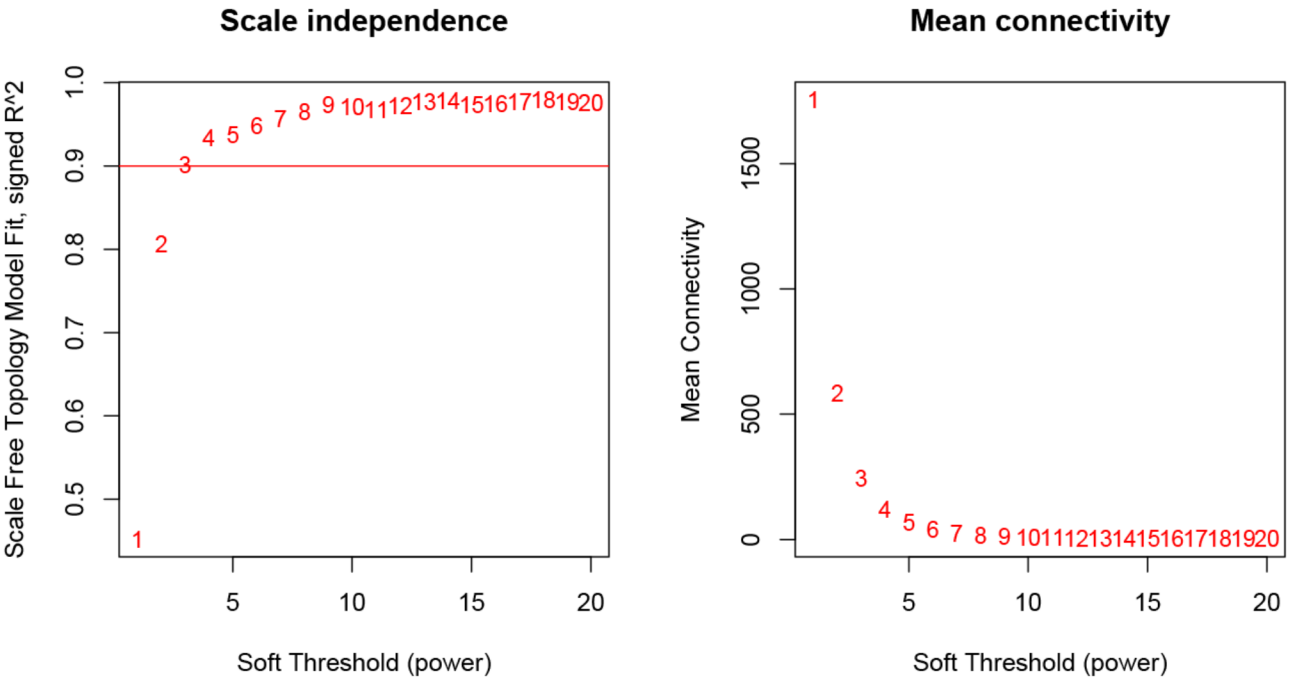


Fig. 1 Scale independence and mean connectivity of different soft thresholds (β). Scale-free fit index analysis for different soft_threshold powers (left) and mean connectivity analysis for different soft_threshold powers (right)

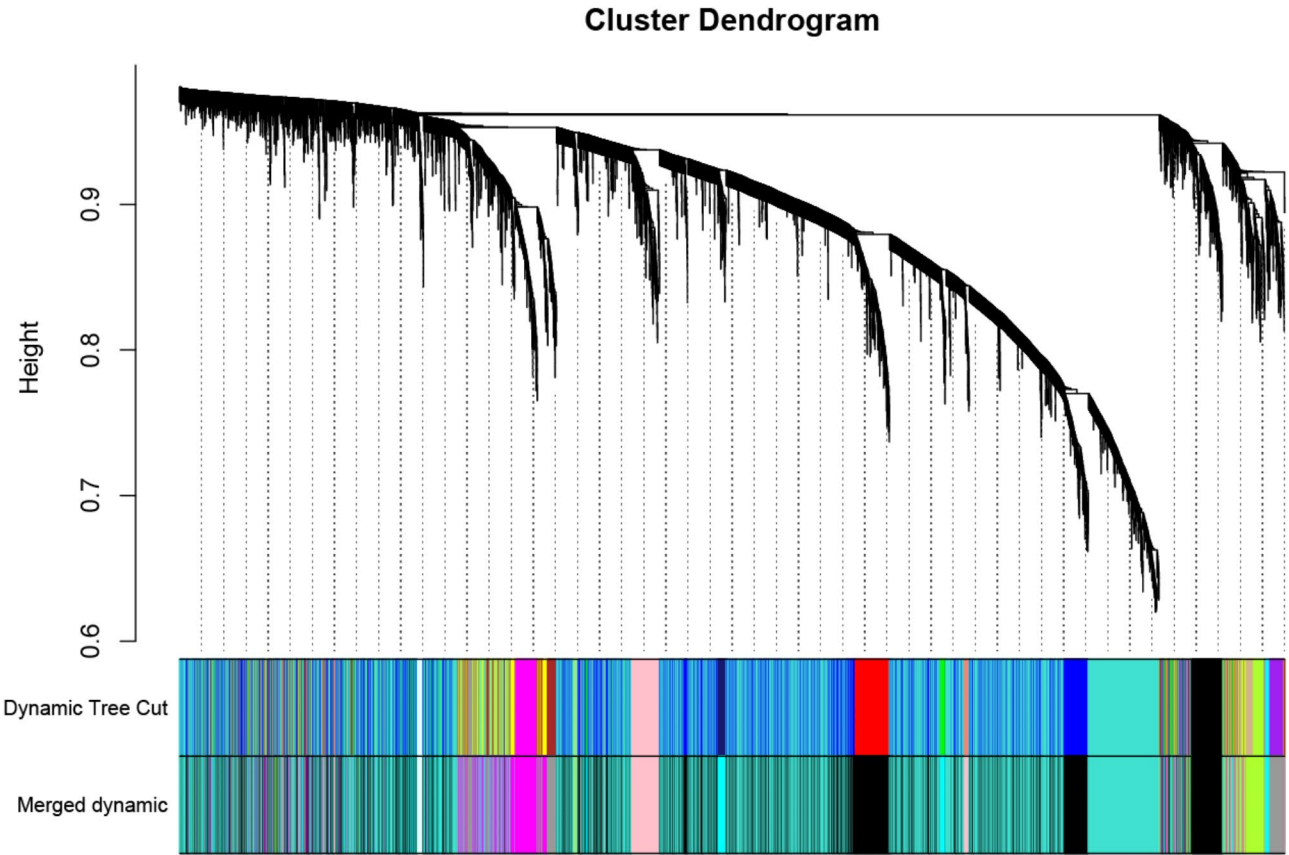


Fig. 2 WGCNA cluster dendrogram of all filtered mRNAs and lncRNA. The genes corresponding to branches of the same color are classified as the same gene module

Table 1 The number of mRNAs and lncRNAs in each of the nine modules

module	mRNA	lncRNA	Total
black	2363	56	2419
cyan	409	7	416
greenyellow	299	3	302
grey60	581	8	589
lightcyan	43	0	43
lightgreen	32	1	33
magenta	525	10	535
pink	300	1	301
turquoise	3536	112	3648

correlated with GS in both the magenta (corr=0.81, $P<0.001$) and turquoise (corr=0.86, $P<0.001$) modules. Hub genes were selected from the key modules for further analysis on the basis of the criteria of $GS>0.4$ and

$MM>0.8$ ($P<0.01$). A total of 50 hub genes, including 49 mRNAs and 1 lncRNA, were identified in the magenta module, while the turquoise module contained 407 hub genes, consisting of 396 mRNAs and lncRNAs. From these, all hub lncRNAs and the top 30 hub genes with the highest degrees of connectivity were selected from the two modules. Gene interaction networks were then constructed using Cytoscape software (Supplementary Fig. 2).

GO and KEGG pathway analyses of key modules

GO and KEGG enrichment analyses of the mRNAs in the magenta and turquoise modules were performed using the DAVID database (<https://davidbioinformatics.nih.gov/>). In total, 118 significantly enriched GO terms were identified in the magenta module, and 104 GO terms were identified in the turquoise module ($P<0.05$;

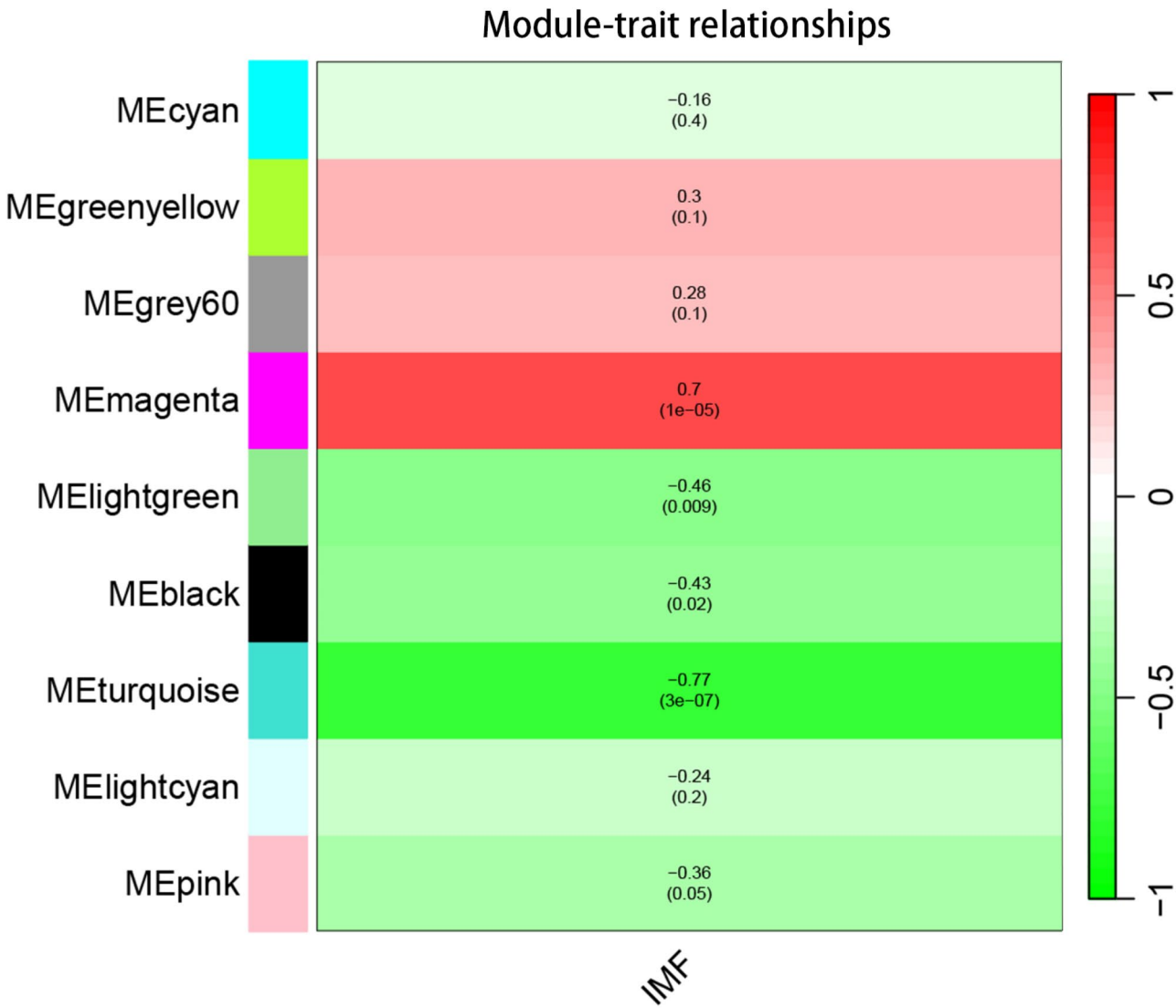


Fig. 3 Correlation analysis between gene module and IMF content

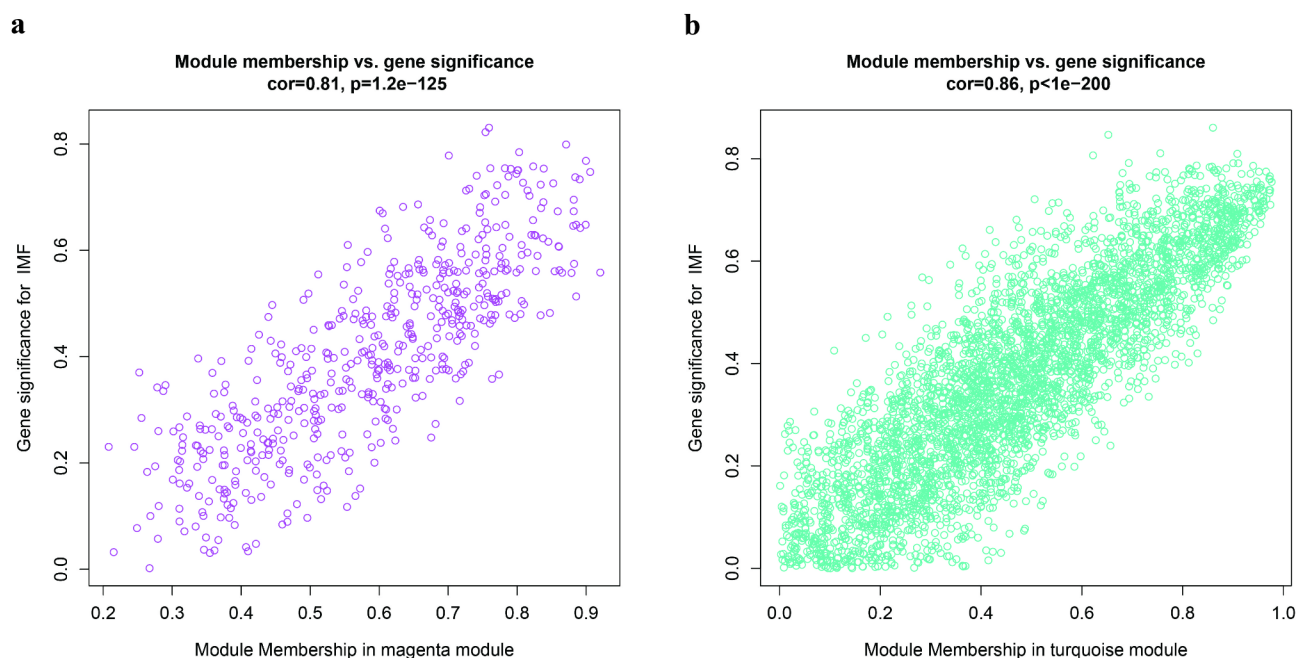


Fig. 4 Scatterplot of Gene module membership and gene significance within the magenta module (a) and the turquoise module (b)

Supplementary Table 3). The top 30 most significant GO terms and the top 10 KEGG pathways are presented in Fig. 5a-d. In both the magenta and turquoise modules, the majority of significantly enriched GO terms were found in the biological process (BP) category, accounting for 90% and 60%, respectively. The genes in the magenta module were predominantly associated with enzyme binding, intracellular signal transduction, and protein localization (Fig. 5a). The genes in the turquoise module were involved primarily in protein and energy metabolism and were localized mainly to intracellular ribonucleoprotein complexes, ribosomes, and mitochondria (Fig. 5b). KEGG pathway enrichment analysis revealed that the magenta module was predominantly associated with lipid synthesis, whereas the turquoise module was linked to energy metabolism. Genes in the magenta module were enriched in pathways such as focal adhesion, PI3K-Akt signalling, MAPK signalling, and insulin resistance (Fig. 5c). In contrast, pathways related to energy metabolism, including oxidative phosphorylation, thermogenesis, and fatty acid degradation, were enriched in the turquoise module (Fig. 5d). Additionally, eight hub mRNAs in the magenta module and thirty hub mRNAs in the turquoise module were identified in KEGG pathways potentially related to fat deposition (Supplementary Table 2).

Construction of the PPI network and identification of key mRNAs

The hub mRNAs involved in KEGG pathways associated with IMF were selected to construct PPI networks using

the STRING database and subsequently analysed and visualized with Cytoscape 3.6.1 software. Only nine hub mRNAs in the magenta module were involved in IMF-related pathways, and STRING analysis did not identify *CASP7* proteins in *Sus scrofa*. As shown in Fig. 6a, proteins that did not interact with other proteins were excluded, and six hub mRNAs in the magenta module were identified as key mRNAs (*CBL*, *CRKL*, *PDGFRB*, *DOCK1*, *YWHAH*, and *HSP90AB1*). In the turquoise module, the PPI network consisted of 26 nodes and 201 edges (Fig. 6b). The top 12 hub mRNAs (*LOC100524873*, *NDUFB7*, *NDUFS5*, *NDUFA11*, *UQCR10*, *NDUFA12*, *NDUFS6*, *NDUFA2*, *NDUFA1*, *NDUFAB1*, *NDUFB10* and *NDUFB3*) with the highest degree of connectivity (degree > 38) were selected as key mRNAs.

Identification of key lncRNAs and lncRNA-mRNA pathway construction of the coexpression network

On the basis of the WGCNA results and the targeting relationships between lncRNAs and key mRNAs, we identified the key lncRNAs that may regulate IMF deposition. In the magenta module, *TGOLN2* was the only hub lncRNA that exhibited a strong correlation ($\text{corr} > 0.79$) with *CRKL* and *DOCK1*. *LOC100524915* demonstrated a strong correlation ($\text{corr} > 0.9$) with all 12 key mRNAs in the turquoise module, despite the high threshold set for correlation. Additionally, *LOC100622481* and *LOC100521518* were found to be associated with multiple key mRNAs on the basis of coexpression correlations. lncRNA-mRNA-pathway networks were constructed to elucidate the roles and functional mechanisms of

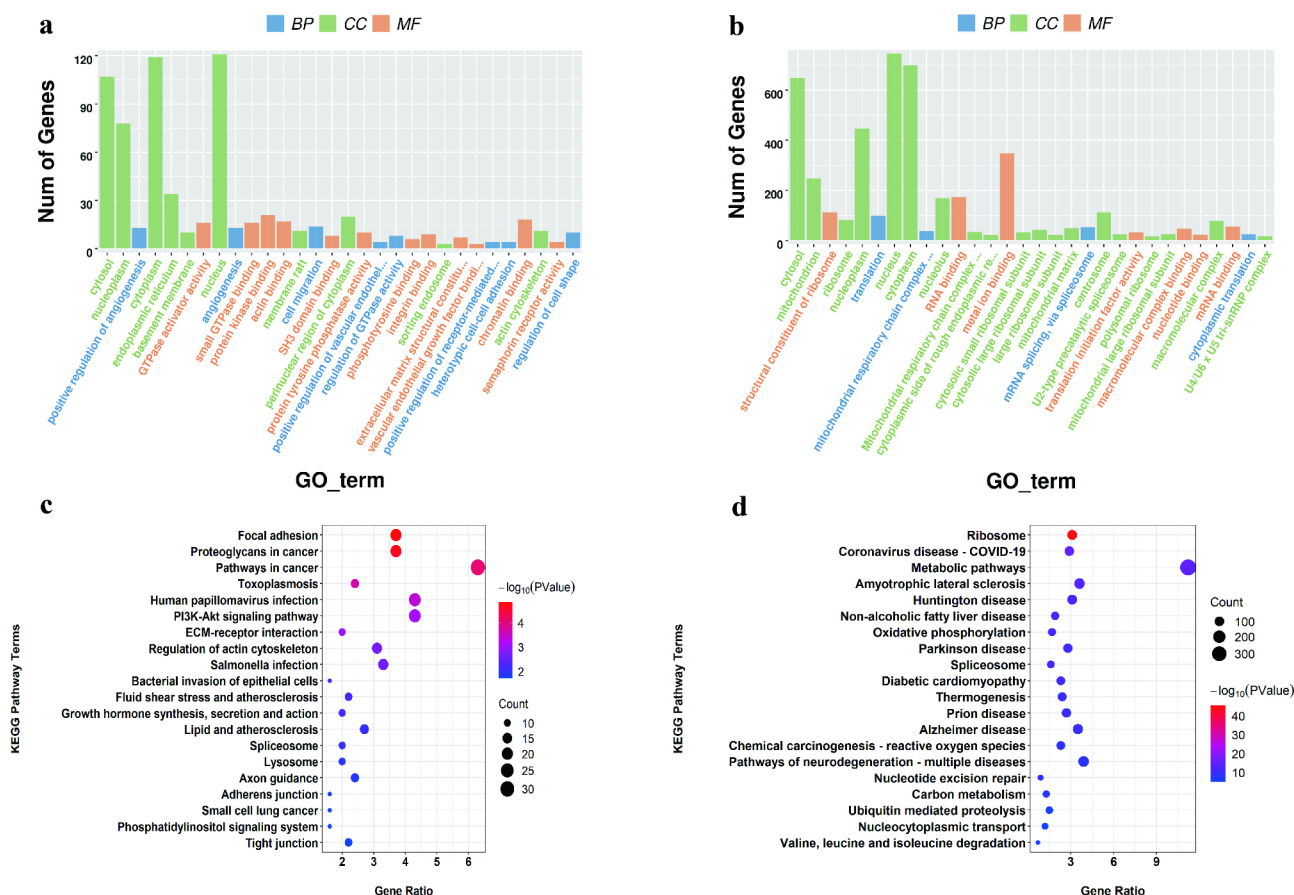


Fig. 5 The top 30 most significant GO terms and the top 10 KEGG pathways are shown in GO and KEGG analysis for genes in the magenta module and turquoise module. **(a), (c)**, GO terms and pathways of the magenta module; **(b), (d)**, GO terms and pathways of the turquoise module

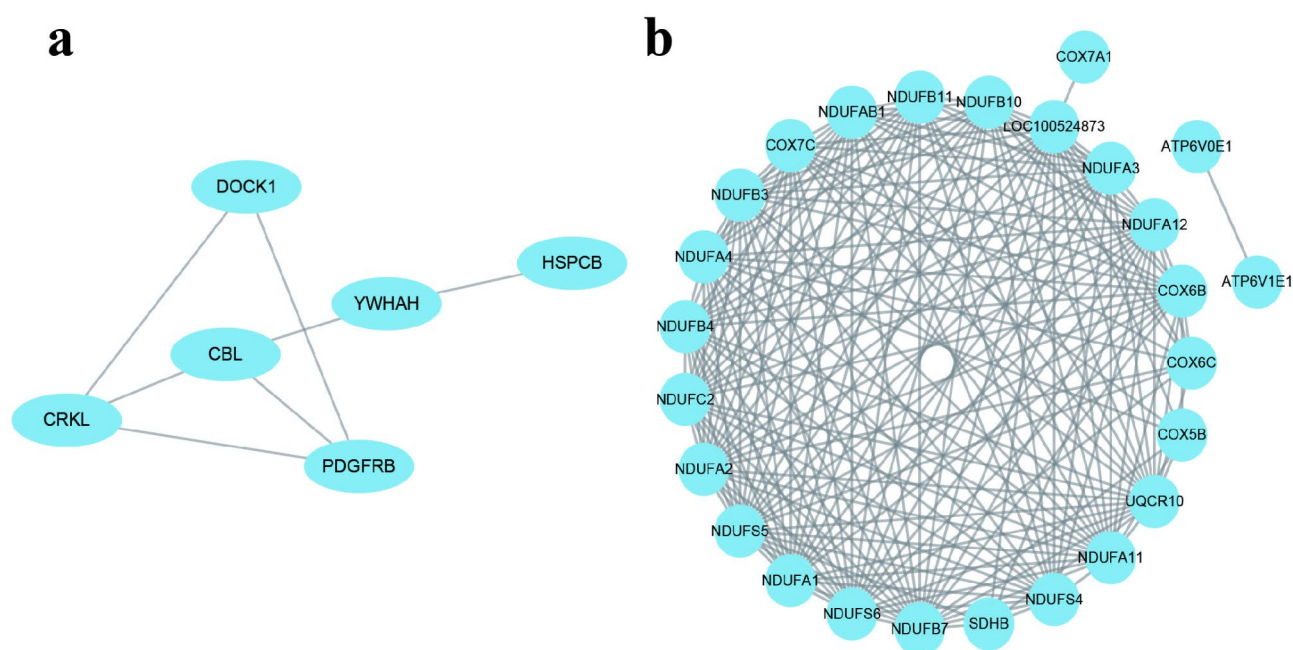


Fig. 6 PPI network of key mRNA in key modules. **a.** The PPI network of 6 IMF-related hub mRNAs in the magenta module. **b.** PPI network of the 26 IMF-related hub mRNAs in the turquoise module

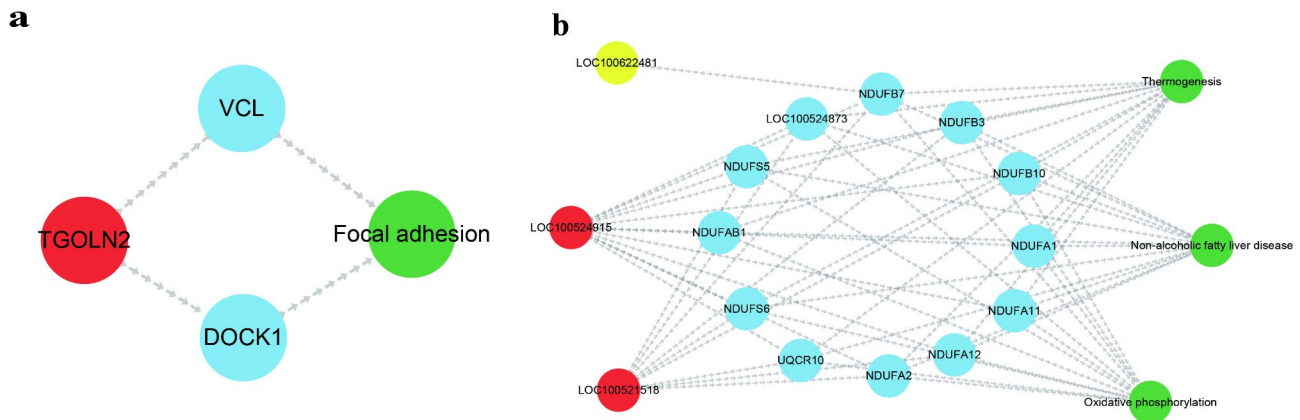


Fig. 7 Co-expression mRNA-lncRNA-pathway of the magenta module (a) and turquoise module (b). Blue nodes represent mRNAs, red nodes represent lncRNAs, and green nodes represent KEGG pathways

key lncRNAs in the magenta and turquoise modules. In the magenta module, the lncRNA-mRNA-pathway network contained two key mRNAs, one key lncRNA, and one KEGG pathway (Fig. 7a). *TGOLN2* may regulate the expression of *DOCK1* and *CRKL*, potentially involving focal adhesion. In the turquoise module, the mRNA-lncRNA-pathway network included 12 key mRNAs, three key lncRNAs, and three KEGG pathways (Fig. 7b). *LOC100524915* was linked to all key mRNAs and enriched in pathways related to thermogenesis, oxidative phosphorylation, and nonalcoholic fatty liver disease. *LOC100521518* was associated with eight mRNAs (*NDUFA2*, *NDUFS6*, *NDUFB7*, *NDUFB3*, *UQC10*, *LOC100524873*, *NDUFB10*, and *NDUFS5*) and was enriched in the oxidative phosphorylation, thermogenesis, and nonalcoholic fatty liver disease pathways. *LOC100622481* was associated with only *NDUFB7* and was enriched primarily in nonalcoholic fatty liver disease.

Discussion

Multiple genes are involved in the regulation of porcine fat deposition, with complex interactions between them. Traditional unidimensional approaches have limited the identification of key genes and regulatory mechanisms underlying this trait. However, WGCNA provides an effective solution by modularizing thousands of genes on the basis of their expression patterns, thereby facilitating the identification of key genes that influence a trait [35]. In our study, we used 8,093 mRNAs and 198 lncRNAs from 31 Jiangquan black pigs to construct a WGCNA. The average connectivity was 243 when R^2 was set to 0.9, indicating both the high biological significance of the network and strong gene correlations within the modules, which facilitated the identification of key mRNAs and lncRNAs. Following the WGCNA, nine coexpression modules were identified, with the magenta module showing a positive correlation with IMF and the

turquoise module exhibiting a negative correlation with IMF (corr = 0.68, $P < 0.01$).

The results of the functional enrichment analysis revealed significant differences in the biological functions associated with the magenta and turquoise modules. The magenta module was associated primarily with IMF-related pathways, including focal adhesion, the PI3K-Akt signalling pathway, the MAPK signalling pathway, and the JAK-STAT signalling pathway. Focal adhesion plays a crucial role in cell adhesion and migration, processes essential for adipocyte differentiation and lipid storage [36–38]. The PI3K-Akt pathway is a key regulator of lipid metabolism in the liver and diabetic kidney disease and influences the proliferation of intramuscular fat precursor cells in pigs [38]. The MAPK pathway is involved in cell proliferation, differentiation, and stress responses, regulating adipogenesis at multiple stages, from stem cells to mature adipocytes [39]. The JAK-STAT pathway is involved in inflammation and the regulation of adipokines in adipocytes, which are hormones that influence fat metabolism [36–37]. According to the WGCNA results, the magenta module was positively correlated with the IMF content, and the hub genes within this module are involved in pathways related to adipocyte differentiation. These findings suggest that the hub genes may play pivotal roles in regulating IMF deposition. Through KEGG and PPI analyses, *PDGFRB*, *CRKL*, *DOCK1*, *CBL*, *YWHAH*, and *HSP90AB1*, which may be key regulators of IMF deposition, were identified as key genes in the magenta module. Platelet-derived growth factor receptor beta (*PDGFRB*, GS = 0.48, MM = 0.83) encodes a cell surface tyrosine kinase receptor that binds members of the platelet-derived growth factor family. *PDGFRB* is involved in several IMF-related pathways, including focal adhesion, the JAK-STAT signalling pathway, and the MAPK signalling pathway. In addition to promoting adipocyte differentiation in the myocardium, *PDGFRB* has been shown to contribute to adipocyte formation

following rotator cuff tears in a mouse model [40]. *CRKL* is an adapter protein involved in the MAPK signalling pathway, focal adhesion, and insulin signalling pathway, all of which regulate preadipocyte differentiation. Several studies have demonstrated that *CRKL* plays a critical role in cancer signalling and promotes adipogenesis in adipose-derived stem cells [41]. *HSP90AB1* encodes the constitutive form of the 90 kDa cytosolic heat shock protein and is involved in the PI3K–Akt signalling pathway, lipid metabolism and atherosclerosis. *HSP90AB1* has been shown to promote proliferation, migration, and glycolysis in cancer cells; however, its association with IMF deposition has not been previously reported [42]. *DOCK1* functions primarily as a small GTPase activator, playing crucial roles in cell migration, morphological changes, and signal transduction. These processes are essential for the differentiation of adipocytes from precursor cells to mature adipocytes. However, the precise mechanisms underlying its role in adipogenesis and lipid metabolism require further investigation [43]. lncRNAs play critical roles in regulating adipogenesis, thermogenesis, and lipid metabolism. In the magenta module, *TGOLN2* was identified as the sole hub lncRNA. Two key mRNAs, *CRKL* and *DOCK1*, were identified as trans-target genes of *TGOLN2*, suggesting that *TGOLN2* may indirectly regulate key processes in lipid metabolism through its interaction with these genes. In the lncRNA–mRNA–pathway network, *TGOLN2* is involved in the focal adhesion signalling pathway by regulating the expression of *CRKL* and *DOCK1*. However, the specific role of *TGOLN2* in lipid metabolism requires further investigation, particularly regarding its impact on adipocyte differentiation and function.

The turquoise module was negatively correlated with fat deposition. Enrichment analysis revealed that the hub genes in this module are involved primarily in energy metabolism, including processes such as oxidative phosphorylation, thermogenesis, the citric acid cycle, and fatty acid degradation [44, 45]. Fat serves as a critical energy source for cells, and its accumulation in adipocytes is typically promoted under conditions of energy surplus, facilitating adipocyte growth and differentiation, which leads to IMF deposition. Oxidative phosphorylation, thermogenesis, the citric acid cycle, and fatty acid degradation are classic pathways involved in energy metabolism and play key roles in cellular energy production and lipid consumption. For example, in brown adipose tissue, the thermogenesis pathway regulates body temperature through fat oxidation, while simultaneously consuming excess fat, thereby reducing fat storage [46]. Hub genes in the turquoise module may regulate oxidative metabolic processes, thereby promoting lipid catabolism and ultimately reducing IMF levels. These findings suggest that the genes within the turquoise module play a key

role in regulating lipid metabolism, particularly in energy expenditure and lipid breakdown, which in turn modulates the mechanisms of fat storage [47]. A total of 35 hub genes in the turquoise module are involved in energy metabolism-related pathways, 16 of which are subunits of NADH: ubiquinone oxidoreductase, including *NDUFA2*, *NDUFB4*, and *NDUFC2*. Ubiquinone oxidoreductase subunits are essential for the assembly of mitochondrial respiratory chain complex I and play a critical role in energy metabolism [48]. PPI network analysis revealed 12 key mRNAs with the highest degree values among the 35 hub genes associated with energy metabolism, suggesting that these genes may play a central regulatory role within the PPI network. *LOC100524873* encodes a subunit of the cytochrome bc1 complex, which plays a crucial role in mitochondrial function [49]. *UQCRI0*, a subunit of mitochondrial complex III, is an essential component of the mitochondrial membrane respiratory chain [50]. The remaining 10 key mRNAs are part of the NADH dehydrogenase family, with these subunits playing a role in electron transfer and mitochondrial energy metabolism [51]. For example, *NDUFAB1* contributes to the breakdown of fat by enhancing mitochondrial metabolism, potentially preventing obesity [52], whereas *NDUFA2* is involved in fatty acid β -oxidation and oxidative phosphorylation, processes that contribute to fatty acid catabolism and may reduce the intramuscular fat content [53].

The lncRNA–mRNA–pathway regulatory network constructed in the turquoise module includes three key lncRNAs (*LOC100524915*, *LOC100622481*, and *LOC100521518*). Of these, 12 key mRNAs are target genes of *LOC100524915* and are involved in processes such as oxidative phosphorylation, thermogenesis, and nonalcoholic fatty liver disease [54]. Although the specific function of *LOC100524915* has not been reported in the literature, the pathways associated with its target genes identified in this study suggest that *LOC100524915* may play a critical role in energy metabolism and lipid metabolism. *NDUFB7* is the sole target gene of *LOC100622481*. In a previous WGCNA of beef quality, *NDUFB7* was identified as a potential key gene associated with intramuscular fat deposition [55]. In humans, *NDUFB7* is expressed at higher levels in adipose tissue than in other tissues [56]. The interaction between *LOC100622481* and *NDUFB7* suggests that *LOC100622481* may play a role in regulating lipid metabolism and IMF deposition [51].

In this study, we have provided preliminary evidence from sequencing analysis, suggesting that key lncRNAs and key mRNAs may be involved in IMF deposition and energy metabolism. However, these hypotheses have not been experimentally validated. Owing to the lack of experimental support, the findings are subject to limitations, and the precise roles of these molecules in IMF deposition and energy metabolism cannot be definitively

confirmed. Future research should include functional experiments to further validate the specific functions and mechanisms of these lncRNAs and mRNAs in lipid metabolism. Additionally, the small sample size and the data derived from a specific population may limit the generalizability of the results. We recommend further validation in a broader cohort to enhance the external validity and reliability of the findings.

Conclusions

In this study, 18 key mRNAs and four key lncRNAs potentially involved in IMF deposition in pigs were identified through WGCNA, enrichment analysis, and PPI analysis. The identified mRNAs, including *CRKL*, *CBL*, *PDGFRB*, and various subunits of NADH dehydrogenase, likely play significant roles in regulating IMF deposition. Additionally, the constructed lncRNA-mRNA-pathway network highlighted the potential regulatory mechanisms of lncRNAs, such as *TGOLN2*, *LOC100521518*, *LOC100524915*, and *LOC100622481*, in influencing IMF deposition. These findings provide valuable insights into the complex regulatory networks governing IMF deposition and offer a foundation for further studies to elucidate the molecular mechanisms underlying fat deposition in pigs, with potential implications for breeding strategies aimed at improving pork quality.

Abbreviations

DAVID	Database for Annotation, Visualization and Integrated Discovery
GO	Gene Ontology
IMF	Intramuscular fat
KEGG	Kyoto Encyclopedia of Gene and Genome
LDM	Longissimus dorsi muscle
lncRNA	Long non-coding RNA
NGS	Next-generation sequencing
PCC	Pearson correlation coefficient
PPI	Protein-protein interaction
RNA-seq	RNA sequencing
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
WGCNA	Weighted gene co-expression network analysis

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11427-x>.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5

Author contributions

HT and WL conceived the study. SY, ZC, HL and FX collected the samples and recorded their phenotypes. WL, SY and ZC extracted RNA. WL analyzed the RNA data. CN helped with data visualization. DW, HT and QZ supervised the study and revised the manuscript. WL and DW wrote and revised the manuscript. All authors read and approved the manuscript.

Funding

The study was funded by Agricultural Animal Breeding Project of Shandong Province (2022LZGC003), the 14th Five-Year Plan national key research Project (2021YFD1301203), the National Natural Science Foundation of China (32102526 and 32002172), the Shandong Provincial Natural Science Foundation (ZR2020QC176 and ZR2020QC175), the National Key Research and Development Program of China (2022YFF1000900), the National Natural Science Foundation of China (32172711 and 32370675) and the Introduction and Cultivation Plan of Youth Innovation Talents for Universities of Shandong Province.

Data availability

Sequence data that support the findings of this study have been deposited in the European Nucleotide Archive with the primary accession code PRJNA997007.

Declarations

Ethics approval and consent to participate

The authors confirm that this study complies with the ARRIVE guidelines of BMC Genomics and was approved by the Ethics Committee of Shandong Agricultural University (No. SDAUA-2018-018).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Shandong Provincial Key Laboratory of Animal Biotechnology and Disease Control and Prevention, Key Laboratory of Efficient Utilization of Non-grain Feed Resources (Co-construction by Ministry and Province), Ministry of Agriculture and Rural Affairs, College of Animal Science & Technology, Shandong Agricultural University, 61 Daizong Street, Taian City, Shandong Province 271018, China

Received: 6 January 2024 / Accepted: 2 March 2025

Published online: 11 March 2025

References

1. Fernandez X, Monin G, Talmant A, Mourot J, Lebret B. Influence of intramuscular fat content on the quality of pig M.at-2. Consumer acceptability of M. longissimus lumborum. *Meat Sci.* 1999;53(1):67–72.
2. Hocquette JF, Gondret F, Baeza E, Medale F, Jurie C, Pethick DW. Intramuscular fat content in meat-producing animals: development, genetic and nutritional control, and identification of putative markers. *Animal.* 2010;4(2):303–19.
3. Suzuki K, Inomata K, Katoh K, Kadowaki H, Shibata T. Genetic correlations among carcass cross-sectional fat area ratios, production traits, intramuscular fat, and serum leptin concentration in duroc pigs. *J Anim Sci.* 2009;87(7):2209–15.
4. Zhang FL, Li WD, Zhang G, Zhang M, Liu ZJ, Zhu KX, Liu QC, Zhang SE, Shen W, Zhang XF. Identification of unique transcriptomic signatures through integrated multispecies comparative analysis and WGCNA in bovine oocyte development. *BMC Genomics.* 2023;24(1):265.
5. Zhu J, Wang Y, Su Y, Zheng M, Cui H, Chen Z. RNA sequencing identifies key genes involved in intramuscular fat deposition in chickens at different developmental stages. *BMC Genomics.* 2024;25(1):219.
6. Wang D, Qin P, Zhang K, Wang Y, Guo Y, Cheng Z, Li Z, Tian Y, Kang X, Li H, et al. Integrated LC/MS-based lipidomics and transcriptomics analyses revealed lipid composition heterogeneity between pectoralis intramuscular fat and abdominal fat and its regulatory mechanism in chicken. *Food Res Int.* 2023;172:113083.
7. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;10(1):57–63.
8. Xu J, Wang C, Jin E, Gu Y, Li S, Li Q. Identification of differentially expressed genes in longissimus dorsi muscle between Wei and Yorkshire pigs using RNA sequencing. *Genes Genom.* 2018;40(4).

9. Xu Y, Qi X, Hu M, Lin R, Hou Y, Wang Z, Zhou H, Zhao Y, Luan Y, Zhao S et al. Transcriptome analysis of adipose tissue indicates that the cAMP signaling pathway affects the feed efficiency of pigs. *Genes (Basel)*. 2018;9(7).
10. Li X, Zhou J, Liu L, Qian K, Wang C. Identification of genes in longissimus dorsi muscle differentially expressed between Wannanhu and Yorkshire pigs using RNA-sequencing. *Anim Genet*. 2016;47(3).
11. Junxing Z, Kan L, Qiyuan Y, Min D, Xiangdong L, Guoqing C. Enhanced adipogenesis in Mashen pigs compared with large white pigs. *Ital J Anim Sci*. 2017;16(2).
12. Li XJ, Zhou J, Liu LQ, Qian K, Wang CL. Identification of genes in longissimus dorsi muscle differentially expressed between Wannanhu and Yorkshire pigs using RNA-sequencing. *Anim Genet*. 2016;47(3):324–33.
13. Bridges MC, Daulagala AC, Kourtidis A. LNCcation: LncRNA localization and function. *J Cell Biol*. 2021;220(2).
14. Wang L, Xie Y, Chen W, Zhang Y, Zeng Y. Identification and functional prediction of long noncoding RNAs related to intramuscular fat content in Laiwu pigs. *Anim Biosci*. 2021.
15. Zhao YH, Chen SK, Yuan JN, Shi YM, Wang Y, Xi YF, Qi XL, Guo Y, Sheng XH, Liu JF et al. Comprehensive analysis of the lncRNA-miRNA-mRNA regulatory network for intramuscular fat in pigs. *Genes-Basel*. 2023;14(1).
16. Xing K, Liu H, Zhang F, Liu Y, Shi Y, Ding X, Wang C. Identification of key genes affecting Porcine fat deposition based on co-expression network analysis of weighted genes. *J Anim Sci Biotechnol*. 2021;12(1):100.
17. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
18. Joshua MS, Eran S, Daphne K, Stuart KK. A gene-coexpression network for global discovery of conserved genetic modules. *Science*. 2003;302(5643).
19. Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet*. 2004;5(2):101–13.
20. Zappaterra M, Gioiosa S, Chillemi G, Zambonelli P, Davoli R. Muscle transcriptome analysis identifies genes involved in ciliogenesis and the molecular cascade associated with intramuscular fat content in large white heavy pigs. *PLoS ONE*. 2020;15(5).
21. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884–90.
22. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60.
23. Pertea M, Pertea GM, Antonescu CM, Chang T, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*. 2015;33(3).
24. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, stringtie and ballgown. *Nat Protoc*. 2016;11(9):1650–67.
25. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9(1).
26. Puig DLBR, Karachaliou N, Estrada-Tejedor R, Teixido J, Costa C, Borrell JI. ALK and ROS1 as a joint target for the treatment of lung cancer: a review. *Transl Lung Cancer Res*. 2013;2(2):72–86.
27. Himansu K, Krishnamoorthy S, Woncheol P, Seung-Hoon L, Bong-Hwan C, Hana K, Yong-Min K, Eun-Seok K, Jin HK, Jang HL et al. Transcriptome analysis to identify long non coding RNA (lncRNA) and characterize their functional role in back fat tissue of pig. *Gene*. 2019;703.
28. Wang X, Li H, Lu X, Wen C, Huo Z, Shi M, Tang X, Chen H, Peng C, Fang Y, et al. Melittin-induced long non-coding RNA NONHSAT105177 inhibits proliferation and migration of pancreatic ductal adenocarcinoma. *Cell Death Dis*. 2018;9(10):940.
29. Wilhelm SM, Carter C, Tang L, Wilkie D, McNabola A, Rong H, Chen C, Zhang X, Vincent P, McHugh M, et al. BAY 43-9006 exhibits broad spectrum oral antitumor activity and targets the RAF/MEK/ERK pathway and receptor tyrosine kinases involved in tumor progression and angiogenesis. *Cancer Res*. 2004;64(19):7099–109.
30. Mu T, Hu H, Ma Y, Yang C, Feng X, Wang Y, Liu J, Yu B, Zhang J, Gu Y. Identification of critical lncRNAs for milk fat metabolism in dairy cows using WGCNA and the construction of a CeRNAs network. *Anim Genet*. 2022;53(6).
31. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504.
32. Wucher V, Legeai F, Hédan B, Rizk G, Lagoutte L, Leeb T, Jagannathan V, Cadieu E, David A, Lohi H et al. FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic Acids Res*. 2017;45(8).
33. Xu M, Ouyang T, Lv K, Ma X. Integrated WGCNA and PPI network to screen hub genes signatures for infantile hemangioma. *Front Genet*. 2020;11:614195.
34. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, et al. Long noncoding RNAs with enhancer-like function in human cells. *Cell*. 2010;143(1):46–58.
35. Liu X, Hu AX, Zhao JL, Chen FL. Identification of key gene modules in human osteosarcoma by co-expression analysis weighted gene co-expression network analysis (WGCNA). *J Cell Biochem*. 2017;118(11):3953–9.
36. Zhou M, Ren P, Li S, Kang Q, Zhang Y, Liu W, Shang J, Gong Y, Liu H. Danhong injection attenuates high-fat-induced atherosclerosis and macrophage lipid accumulation by regulating the PI3K/AKT insulin pathway. *J Cardiovasc Pharmacol*. 2019;74(2):152–61.
37. Huang H, Liu L, Li C, Liang Z, Huang Z, Wang Q, Li S, Zhao Z. Fat mass- and obesity-associated (FTO) gene promoted myoblast differentiation through the focal adhesion pathway in chicken. *3 Biotech*. 2020;10(9):403.
38. Liu Y, Wang R, Zhang L, Li J, Lou K, Shi B. The lipid metabolism gene FTO influences breast cancer cell energy metabolism via the PI3K/AKT signaling pathway. *Oncol Lett*. 2017;13(6):4685–90.
39. Xin F, Cheng Y, Wen X, Zhang J, Shi X, Liu P, Ren J, Lu W, Liu F, Li Z, et al. BK channel depletion promotes adipocyte differentiation by activating the MAPK/ERK pathway. *Stem Cells*. 2024;42(2):146–57.
40. Jiang Z, Feng T, Lu Z, Wei Y, Meng J, Lin C, Zhou B, Liu C, Zhang H. PDGFR β + mesenchymal cells, but not NG2 + mural cells, contribute to cardiac fat. *Cell Rep*. 2021;34(5).
41. Mazzu YZ, Hu Y, Soni RK, Mojica KM, Qin LX, Agius P, Waxman ZM, Mihailovic A, Socci ND, Hendrickson RC, et al. miR-193b-regulated signaling networks serve as tumor suppressors in liposarcoma and promote adipogenesis in adipose-derived stem cells. *Cancer Res*. 2017;77(21):5728–40.
42. Wang H, Deng G, Ai M, Xu Z, Mou T, Yu J, Liu H, Wang S, Li G. Hsp90ab1 stabilizes LRP5 to promote epithelial-mesenchymal transition via activating of AKT and Wnt/beta-catenin signaling pathways in gastric cancer progression. *Oncogene*. 2019;38(9):1489–507.
43. Laurin M, Fradet N, Blangy A, Hall A, Vuori K, Cote JF. The atypical Rac activator Dock180 (Dock1) regulates myoblast fusion in vivo. *Proc Natl Acad Sci U S A*. 2008;105(40):15446–51.
44. Takeda Y, Harada Y, Yoshikawa T, Dai P. Mitochondrial energy metabolism in the regulation of thermogenic brown fats and human metabolic diseases. *Int J Mol Sci*. 2023;24(2).
45. Ikeda K, Kang Q, Yoneshiro T, Camporez JP, Maki H, Homma M, Shinoda K, Chen Y, Lu X, Maretich P, et al. UCP1-independent signaling involving SERCA2b-mediated calcium cycling regulates beige fat thermogenesis and systemic glucose homeostasis. *Nat Med*. 2017;23(12):1454–65.
46. Chouchani ET, Kajimura S. Metabolic adaptation and maladaptation in adipose tissue. *Nat Metab*. 2019;1(2):189–200.
47. Sarantopoulos CN, Banyard DA, Ziegler ME, Sun B, Shaterian A, Widgerow AD. Elucidating the preadipocyte and its role in adipocyte formation: a comprehensive review. *Stem Cell Rev Rep*. 2018;14(1):27–42.
48. Crane FL, Low H. NADH oxidation in liver and fat cell plasma membranes. *Febs Lett*. 1976;68(2):153–6.
49. Goetzl EJ, Srihari VH, Guloksuz S, Ferrara M, Tek C, Heninger GR. Decreased mitochondrial electron transport proteins and increased complement mediators in plasma neural-derived exosomes of early psychosis. *Transl Psychiatry*. 2020;10(1):361.
50. Esposti MD, Lenaz G. Kinetic indication for multiple sites of ubiquinol-1 interaction in ubiquinol-cytochrome C reductase in bovine heart mitochondria. *Arch Biochem Biophys*. 1982;216(2):727–35.
51. Correia SP, Moedas MF, Naess K, Bruhn H, Maffezzini C, Calvo-Garrido J, Lesko N, Wibom R, Schober FA, Jemt A, et al. Severe congenital lactic acidosis and hypertrophic cardiomyopathy caused by an intronic variant in NDUFB7. *Hum Mutat*. 2021;42(4):378–84.
52. Zhang R, Hou T, Cheng H, Wang X. NDUFB1 protects against obesity and insulin resistance by enhancing mitochondrial metabolism. *Faseb J*. 2019;33(12):13310–22.
53. He P, Jiang F, Guo W, Guo YF, Lei SF, Deng FY. PhosSNPs-regulated gene network and pathway significant for rheumatoid arthritis. *Hum Hered*. 2021;86(1–4):10–20.
54. Rak M, Rustin P. Supernumerary subunits NDUFA3, NDUFA5 and NDUFA12 are required for the formation of the extramembrane arm of human mitochondrial complex I. *Febs Lett*. 2014;588(9):1832–8.

55. Yu H, Yang Z, Wang J, Li H, Li X, Liang E, Mei C, Zan L. Identification of key genes and metabolites involved in meat quality performance in Qinchuan cattle by WGCNA. *J Integr Agr*. 2024;23(11):3923–37.
56. Fagerberg L, Hallstrom BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, Habuka M, Tahmasebpoor S, Danielsson A, Edlund K, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteom*. 2014;13(2):397–406.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.