Identification and co-expression analysis of long noncoding RNAs and mRNAs involved in the deposition of intramuscular fat in Aohan fine-wool sheep

Fuhui Han  
Qingdao Agricultural University

Jing Li  
Qufu animal husbandry and veterinary technical service center

Nan Liu  
Qingdao Agricultural University

Ranran Zhao  
Qingdao Agricultural University

Lirong Liu  
China Animal Health and Epidemiology Center

Lanlan Li  
Qingdao Agricultural University

Qian Li  
Qingdao Agricultural University

Jianning He (hexingxing104@163.com)  
Qingdao Agricultural University  https://orcid.org/0000-0002-9657-9752

Research article

Keywords: Intramuscular fat, Aohan fine-wool sheep, Lipid deposition, Long non-coding RNAs, Co-expression analysis

DOI: https://doi.org/10.21203/rs.3.rs-18941/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: Intramuscular fat (IMF) content has become one of the most important indicators for measuring meat quality, and levels of IMF are affected by various genes. Long non-coding RNAs (lncRNAs) are widely expressed non-coding RNAs that play an important regulatory role in a variety of biological processes; however, research on the lncRNAs involved in sheep IMF deposition is still in its infancy. Aohan fine-wool sheep (AFWS), one of China's most important meat-hair, dual-purpose sheep breed, provides a great model for studying the role of lncRNAs in the regulation of IMF deposition. We identified lncRNAs by RNA sequencing in longissimus dorsi muscle (LDM) samples of sheep at two ages: 2 months (Mth-2) and 12 months (Mth-12).

Results: We identified a total of 26,247 genes and 6,935 novel lncRNAs in LDM samples of sheep. Among these, 606 mRNAs and 408 lncRNAs were differentially expressed. We then compared the structural characteristics of lncRNAs and mRNAs. We obtained target genes of differentially expressed lncRNAs (DELs) and performed enrichment analyses using Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG). We found that target mRNAs were primarily enriched in lipid metabolism, lipid transport, regulation of primary metabolic processes and developmental pathways. Based on the results of important KEGG pathways, we obtained six candidate lncRNAs that potentially regulate lipid deposition and constructed an lncRNA-mRNA co-expression network that included MSTRG.792.1-SCD, MSTRG.8227.1-ACAA2, MSTRG.10679.1-FADS2, MSTRG.21942.1-PLA2G4E, MSTRG.21380.1-FZD4 and MSTRG.9270.1-ULK1. We speculated that these candidate lncRNAs might play a role by regulating the expression of target genes. We randomly selected five mRNAs and five lncRNAs to verify the accuracy of the sequencing data by qRT-PCR.

Conclusions: Our study provided a list of the lncRNAs and mRNAs related to intramuscular lipid deposition in sheep and laid a foundation for future research on regulatory mechanisms.

Background

High-quality lamb meat is becoming increasingly popular as living standards improve and dietary patterns change. Currently, evaluations of the meat quality of livestock have revealed that the content of intramuscular fat (IMF) is lower in carcass fats, yet IMF has a critically important influence on the edibility and flavor of muscle meat [1]. Indeed, the quantity of IMF has become one of the most critical parameters of meat quality indicators, as it is considered to be positively related to meat quality and texture [2, 3]. When a certain amount of fat is deposited between the muscle bundles and muscle fibers, the marbled section of the meat has a high score, and the meat is fresh and juicy, which is often considered ideal [4,5]. The selective deposition of fat can improve production efficiency and play a key role in improving meat quality. This practice is also a major focus and challenge of modern livestock breeding [6]. Therefore, ensuring the appropriate deposition of IMF in lean meat can enhance the future quality of sheep meat.
Studies have shown that intramuscular lipid deposition is affected by multiple genes and signaling pathways, such as the FAS, FAM134B, and HSL genes and the Wnt and AMPK signaling pathways [7-9]. Recently, long non-coding RNAs (lncRNAs) have received increased attention for their wide-ranging functions. LncRNAs refer to a class of non-coding RNAs longer than 200 nt in length [10]. Most lncRNAs have significant temporal and spatial expression specificity [11, 12] and have low sequence conservation among species [13-15]. LncRNAs can be divided into five types based on their positions relative to neighboring protein-coding genes: intronic lncRNAs, bidirectional lncRNAs, sense lncRNAs, intergenic lncRNAs and antisense lncRNAs [16].

LncRNAs can regulate various life activities of the body, including epigenetic regulation, transcriptional regulation and post-transcriptional regulation [17-19]. The most common regulation methods of lncRNAs include cis-regulation of the transcription of neighboring protein-coding genes and the trans-regulation of non-adjacent genes. In addition, lncRNAs can interact with miRNAs to affect the post-transcriptional translation of related mRNAs [20-22]. Studies have shown that lncRNAs can play direct or indirect roles in the process of lipid accumulation [23]. SRA (steroid receptor RNA activator) is one of the earliest discovered lncRNAs and plays an important role in fat metabolism. SRA can bind to peroxisome proliferator-activated receptor gamma (PPARγ) and enhance PPARγ activity, thereby promoting the differentiation of pre-adipocytes [24]. A study of the expression levels of lncRNAs in the IMF of Jinhua and Landrace pigs revealed a total of 119 differentially expressed lncRNAs (DEl.s), six of which were involved in fat deposition and lipid metabolism-related pathways [25]. Furthermore, an analysis of transcriptome data from IMF in Inner Mongolia goats revealed that 1,472 lncRNAs were involved in adipocyte growth regulation and morphological changes of adipocytes [26]. Another study has shown that lncRNAs can play a key regulatory role in fat deposition in sheep tails [27]. Overall, these findings demonstrate that lncRNAs can regulate lipid deposition through a variety of regulatory mechanisms. However, few studies have assessed the roles of lncRNAs in intramuscular lipid deposition in sheep.

Aohan fine-wool sheep (AFWS) is an important meat-hair, dual-purpose sheep breed in China that grows rapidly early in development. The elimination of male lambs for fat lamb production can increase both hair and meat gains as well as improve the overall benefits provided by fine wool sheep [28]. Exploring the developmental characteristics of IMF deposition and selecting candidate genes for AFWS provide references for future studies and applications in sheep breeding, improve the quality of mutton and accelerate the breeding process. The goal of our study was to systematically identify the profiles of differentially expressed mRNAs (DEMs) and DELs during intramuscular lipid deposition in sheep through high-throughput sequencing. We hoped that by studying the relationship between lncRNAs and lipid deposition, our findings would shed light on the mechanisms underlying selective muscle lipid deposition in sheep.

Results

Determination of IMF content
Results for the IMF content of sheep are shown in Table 1. The IMF content of the longissimus dorsi muscle (LDM) at 2, 4, 6 and 12 months was 2.202 ± 0.006, 4.566 ± 0.178, 10.685 ± 0.690 and 11.163 ± 0.878, respectively. We found that the IMF content of LDM at Mth-4 was significantly higher than that at Mth-2 (P < 0.01) and was significantly lower than that at Mth-6 and Mth-12 (P < 0.01). The IMF content of LDM in Mth-12 was also significantly higher than that observed in Mth-2 (P < 0.01). No significant differences were detected between Mth-6 and Mth-12. The same pattern was observed for the biceps femoris muscle (BFM). IMF content in the LDM was significantly higher than that in the BFM in the same month (P < 0.01). Thus, Mth-2 and Mth-12 were selected for RNA sequencing (RNA-seq).

Profiles of lncRNAs and mRNAs in sheep muscle

A total of six RNA expression profiles were generated in this experiment. The results are shown in Table 2. The average raw reading was 13.62 G. After preprocessing the raw data, the average value of the filtered data obtained in each library was 12.82 G. The data obtained from the six expression profiles were relatively average with Q20 ≥ 99% and G/C contents ranging from 49 to 53%, indicating that the quality of the filtered data was reliable. The filtered clean reads were compared with the reference genome with HISAT2. The comparison rate in all six samples was greater than 88%, indicating that the experiment was free of contamination and that the experimental results were robust.

An average of 24,384 expressed genes was identified in the six libraries, and a summary of the protein-coding genes identified is provided in Additional file 1 (Table S1). An average of 6,499 unique lncRNAs was identified in the libraries. The information associated with all identified lncRNAs is shown in Additional file 2 (Table S2A). We used circos (http://www.circos.ca) software to perform genomic mapping on the lncRNAs obtained by screening. We found that the number of reads was positively related to the length of the chromosome (Fig. 1a). Based on the locations of novel lncRNAs in the genome, we identified 525 antisense lncRNAs, 304 sense lncRNAs, 350 bidirectional lncRNAs, 1,710 intronic lncRNAs and 4,046 intergenic lncRNAs (Fig. 1b, Additional file 2: Table S2B). The sequence information for all identified lncRNAs is shown in Additional file 3 (Table S3).

The structural characteristics and the expression levels of lncRNAs and mRNAs were different. The average length of lncRNAs was 868 nt, which was shorter than the average length of mRNAs (2,131 nt) (Fig. 1c). LncRNAs consisted of 1.7 exons on average, while mRNAs had 9.9 exons on average (Fig. 1d); thus, lncRNAs had fewer exons than mRNAs. Meanwhile, lncRNAs had lower expression levels relative to mRNAs (Fig. 1e). Moreover, the length of the open reading frame (ORF) of lncRNAs tended to be shorter than that of mRNAs (Fig. 1f, 1g). Overall, lncRNAs were characterized by shorter lengths, fewer exons, lower expression levels and shorter ORF length distributions compared with mRNAs.

Identification of differentially expressed mRNAs and lncRNAs

A total of 606 DEMs were identified in muscle tissue (log2 (fold change) ≥ 1 or log2 (fold change) ≤ -1 and P <0.05). Of these differentially expressed genes (DEGs), 154 were up-regulated and 452 were down-regulated (Fig. 2a, 2c). A summary of DEGs is provided in Additional file 4 (Table S4). We identified 408
IncRNAs that were differentially expressed, of which 254 IncRNAs were up-regulated and 154 IncRNAs were down-regulated (Fig. 2b, 2d). The list of DELs is provided in Additional file 5 (Table S5). To illustrate the overall distribution of DEGs, we created clustering maps of DEMs and DELs (Fig. 2e, f). Red indicates that the gene had a higher expression level, and blue indicates that the gene had a lower level of expression.

Among the DEMs, we found that IGF2 was highly expressed at Mth-12. IGF2 has been reported to be a candidate gene, as it can play an important role in fat deposition in pig [29]. Some DEGs were also associated with lipid deposition in this study. As the target gene of miR-132-3p, UCP2 can regulate the differentiation of sheep precursor fat cells [30]. RNA-seq on the broiler pectoralis major muscles has revealed that ADIPOQ and CIDEC were key genes involved in lipid deposition [31]. SOCS2 can also act as a regulator of adipocyte size [32]. Transcriptome analyses examining IMF content in the LDM in heavy Iberian Pigs identified FOSB as a candidate gene and other regulatory factors [33]. We thus plan to focus on studying these genes in future analyses. Among the DELs, we found that MSTRG.8215.2, MSTRG.21380.1, MSTRG.21942.1, MSTRG.21719.1, MSTRG.8227.1, MSTRG.792.1 and MSTRG.19396.1 were highly expressed at Mth-12. We speculated that these novel IncRNAs might promote lipid deposition. Meanwhile, MSTRG.11343.4, MSTRG.13921.1, MSTRG.19788.2, MSTRG.2469.2, MSTRG.8912.2, MSTRG.2792.1 and MSTRG.21775.2 were highly expressed at Mth-2. We speculated that these novel IncRNAs might inhibit lipid deposition. However, the regulatory mechanisms underlying these IncRNAs require further study.

Enrichment analysis of differentially expressed mRNAs

GO functional enrichment analysis of DEGs revealed that these genes participated in a total of 419 significantly enriched functional classifications (P < 0.05), 282 of which were related to biological processes, 41 related to cellular components and 96 related to molecular functions (Additional file 6: Table S6A). The top 20 GO terms are shown in Fig. 3a. The most significantly enriched GO terms were: toll-like receptor 3 signaling pathway (GO: 0034138), synapse pruning (GO: 0098883), reverse cholesterol transport (GO: 0043691), protein hetero dimerization activity (GO: 0046982), Nucleosome (GO: 0000786), extracellular region (GO: 0005576), DNA binding (GO: 0003677) and chromosome (GO: 0005694).

In addition, results of the KEGG pathway analysis showed that these DEGs were involved in 232 biological pathways (Additional file 6: Table S6B), 18 pathways of which were significantly enriched (Additional file 7: Table S7), including cholesterol metabolism (ko04979), arachidonic acid metabolism (ko00590) and glycine, serine and threonine metabolism (ko00260), all of which were related to fat metabolism. Moreover, the top 20 signaling pathways are shown in Fig. 3b. Systemic lupus erythematosus (ko05322) showed the highest level of significance with 23 DEGs. The results indicated that these pathways may have significantly contributed to the deposition of IMF.

Comprehensive analysis of candidate IncRNAs and mRNAs
To understand the potential function of novel lncRNAs, we performed cis-regulation and trans-regulation analyses on candidate lncRNAs. A total of 183 DELs regulated 218 DEMs, five lncRNAs of which acted on four mRNAs through cis-regulation and 183 lncRNAs that acted on 218 mRNAs through trans-regulation (Additional file 8: Table S8). GO analysis of targets of lncRNAs revealed that these genes participated in a total of 1,840 GO terms, 546 of which were significantly enriched (P < 0.05) (Additional file 9: Table S9A). In GO annotation, these DEGs primarily played a role in biological processes. For example, positive regulation of phospholipid translocation (GO: 0061092), regulation of phospholipid catabolic process (GO: 0060696), Wnt signaling pathway, calcium modulating pathway (GO: 0007223), regulation of intracellular cholesterol transport (GO: 0032383), phospholipase C-activating dopamine receptor signaling pathway (GO: 0060158) and regulation of phospholipid biosynthetic process (GO: 0071071). The KEGG pathway enrichment analysis of target genes revealed a total of 214 annotated pathways (Additional file 9: Table S9B). Of these pathways, 28 were significantly enriched (P < 0.05) (Additional file 10: Table S10). Among them, eight pathways were related to lipid deposition and metabolism, including alpha-Linolenic acid metabolism (ko00592), FoxO signaling pathway (ko04068), Biosynthesis of unsaturated fatty acids (ko01040), Cell adhesion molecules (ko04514), Phosphonate and phosphinate metabolism (ko00440), Ether lipid metabolism (ko00565), Tight junction (ko04530) and Arachidonic acid metabolism (ko00590). Although some pathways were not significantly enriched, such as PPAR, Wnt, AMPK and mTOR signaling pathways, these pathways still played an important role in lipid deposition [31, 34-36]. Overall, the KEGG enrichment analysis of target genes revealed 16 critical pathways with 17 target genes (Table 3). We selected the DEMs and DELs with Pearson correlation coefficients ≥ 0.8 or ≤ -0.8 (Additional file 11: Table S11). Among these genes, six were associated with lipid deposition, including SCD, ACAA2, FADS2, PLA2G4E, FZD4 and ULK1 [37-41], and were used to construct the lncRNA-mRNA co-expression network (Fig. 4). We speculate that MSTRG.792.1, MSTRG.8227, MSTRG.10679.1, MSTRG.21942.1, MSTRG.21380.1 and MSTRG.9270.1 might significantly contribute to the deposition of IMF. However, the regulatory mechanisms underlying these lncRNAs require further study.

Validation of lncRNA and mRNA expression by qRT-PCR

To validate the expression levels of DELs and DEMs, we randomly selected five DELs and five DEMs and detected their expression levels by qRT-PCR (Fig. 5a). The results of RNA-seq are shown in Fig. 5b. Comparison of the two sets of results above revealed consistent regulatory trends of genes detected by the two methods, indicating that the RNA-seq data were accurate.

Discussion

IMF content increased gradually with growth, as significant differences were detected between Mth-2 and Mth-12. These findings were consistent with a previous study showing that the IMF content of sheep increased from 0 to 6 months but remained stable thereafter until 12 months of age [42]. Furthermore, these findings are consistent with the characteristics of muscle growth and the development of experimental sheep. The sheep switched to a fattening phase after weaning at 2 months. The weight of sheep increased rapidly between the ages of 4 to 6 months, after which weight gain stabilized. IMF is an
important feature contributing to meat quality. Therefore, we selected the LDM samples at Mth-2 (less lipid deposition) and Mth-12 (more lipid deposition) for RNA-seq to provide a robust test of gene expression differences.

Overall, we identified a total of 26,247 genes and 6,935 predicted novel IncRNAs in LDM samples of sheep by RNA-seq. Among these, 606 mRNAs (154 up-regulated and 452 down-regulated) and 408 IncRNAs (254 up-regulated and 154 down-regulated) were differentially expressed. To further characterize the mechanisms underlying DEGs, we performed GO and KEGG analysis of DEMs. In GO annotation, these DEGs primarily played a role in biological processes. These processes were closely related to fat formation and deposition, such as reverse cholesterol transport (GO: 0043691), positive regulation of cholesterol efflux (GO: 0010875) and positive regulation of cholesterol esterification (GO: 0010873) [43-45]. In addition, we found that many genes were enriched in biological processes, such as signal transmission, organ development, biosynthesis and cell proliferation, and these are also important processes in muscle development [46-48]. KEGG pathway analysis revealed that the DEMs were significantly enriched in the immune system, inflammatory response and infectious diseases pathways, demonstrating that signal transmission between adipocytes and immune cells can greatly affect the function of adipose tissue [49]. This result was consistent with the fact that inflammatory cell infiltration has been documented to commonly occur in adipose tissue and stimulate the activation of the immune defense system [50]. In addition, we also found that many pathways related to lipid metabolism (Cholesterol metabolism and Arachidonic acid metabolism) were significantly enriched for many DEMs, such as ADIPOQ, CIDEC, SOCS2, SCD and ACAA2, that have been reported to participate in lipid metabolism [51-55]. Many DEMs, such as C1QC, AGPAT4, SNCA and CAV3, were found to be related to lipid deposition for the first time [56-58]. Based on the GO and KEGG analysis, we obtained DEM expression profiles that affected the IMF deposition of sheep. However, further research is required to identify the underlying mechanisms.

In our study, 183 DELs participated in the regulation of mRNAs, five IncRNAs of which acted on four mRNAs through cis-regulation and 183 IncRNAs on 218 mRNAs through trans-regulation. The functionality of IncRNA is reflected through the study of their target genes [59]. We performed GO and KEGG enrichment analysis on these target genes. We focused on the GO terms related to lipid deposition. These included terms under biological processes, such as lipoprotein metabolic process (GO: 0042157), long-chain fatty acid biosynthetic process (GO: 0042759) and unsaturated fatty acid biosynthetic process (GO: 0006636), as well as molecular functions, such as Wnt-protein binding (GO: 0017147), phosphatidic acid binding (GO: 0070300) and phosphodiesterase I activity (GO:0004528). We found that the identified IncRNAs might be related to these GO terms; however, further research is required to identify the precise mechanisms.

The KEGG enrichment analysis revealed that target genes were significantly enriched in lipid metabolism pathways, such as alpha-linolenic acid metabolism (ko00592), phosphonate and phosphininate metabolism (ko00440) and ether lipid metabolism (ko00565). Cell adhesion molecules, tight junction and the FoxO signaling pathway also have important regulatory effects on lipid deposition [60-62]. Although
some signaling pathways, such as the “Wnt signaling pathway,” “MAPK signaling pathway,” “AMPK signaling pathway” and “PPAR signaling pathway” were not significantly enriched in our study, they are critically important in the process of lipid deposition [63-66]. The above analysis and KEGG pathways led to the identification of six novel lncRNAs: MSTRG.21380.1, MSTRG.21942.1, MSTRG.792.1, MSTRG.8227.1, MSTRG.10679.1 and MSTRG.9270.1. These lncRNAs might play important roles in lipid deposition and deserve further study.

To facilitate future studies of the mechanisms underlying lncRNAs, we constructed the lncRNA-mRNA co-expression network based on important KEGG pathways. The network contained 127 lncRNAs and eight mRNAs. We obtained six pairs of lncRNAs-mRNAs, including MSTRG.792.1-SCD, MSTRG.8227.1-ACAA2, MSTRG.10679.1-FADS2, MSTRG.21942.1-PLA2G4E, MSTRG.21380.1-FZD4 and MSTRG.9270.1-ULK1. Target genes of these lncRNAs have been reported to be involved in lipid deposition. For example, a study examining gene expression differences in metabolism and function between intramuscular and subcutaneous adipocytes in cattle found that SCD was highly expressed in adipocytes and closely associated with fat formation [53]. ACAA2 can play an important role in fatty acid metabolism by promoting the differentiation of sheep precursor adipocytes into adipocytes [38, 67]. The inhibition of FADS2 in essential fatty acid deficiency induces hepatic lipid accumulation via impairment of very low-density lipoprotein secretion [68]. PLA2G4E was significantly enriched in alpha-linolenic acid metabolism pathways (linolenic acid is closely related to the flavor of lamb) and is associated with lipid metabolism or cholesterol metabolism [40, 69]. FZD4 is highly expressed during fat production [41], and ULK1 participates in lipid metabolism [62, 63]. These results provide information for future studies examining how lncRNAs regulate IMF deposition in sheep. The specific regulatory mechanisms require further study and testing.

**Conclusions**

Our study systematically identified mRNA and lncRNA expression profiles during intramuscular lipid deposition in sheep. We obtained a total of 606 DEMs and 408 DELs and identified some important lncRNAs related to lipid deposition through GO and KEGG enrichment analysis. In addition, co-expression network analysis of lncRNAs and mRNAs involving 127 lncRNAs and eight mRNAs was conducted based on significant KEGG pathways. Six pairs of lncRNA-mRNA, including MSTRG.792.1-SCD, MSTRG.8227.1-ACAA2, MSTRG.10679.1-FADS2, MSTRG.21942.1-PLA2G4E, MSTRG.21380.1-FZD4 and MSTRG.9270.1-ULK1, were selected for further research. Our study provided a list of the lncRNAs and mRNAs related to intramuscular lipid deposition and laid a foundation for future research on the regulatory mechanisms of lncRNA on sheep muscle lipid deposition.

**Methods**

**Sample preparation**
All experimental sheep came from the AFWS Stud Farm (Chifeng, Inner Mongolia, China). All sheep were fed under the same feeding and management conditions. A total of 12 healthy AFWS rams (3 individuals for each stage) at 2, 4, 6 and 12 months of age were killed for the sample collection. AFWS rams were obtained from 12 ewes of a similar age and weight that were in estrus simultaneously and were artificially inseminated from the same ram. The 12 healthy AFWS rams were placed in a closed chamber and anesthetized with sodium pentobarbital at a dose of 25 mg/kg by intravenous injection. Rams were anesthetized and eventually sacrificed in the enclosed chamber by having it filled with 20% carbon dioxide every minute until the gas concentration had reached 80%. Experimental animal handling procedures were performed following published protocols [72, 73]. Samples of the LDM and BFM were collected, placed in RNAase-free Eppendorf tubes and stored immediately in liquid nitrogen. Likewise, the LDM and BFM samples (150 g for each muscle) were quickly collected and stored at -20 °C for the determination of IMF content using Soxhlet petroleum-ether extraction.

**Determination of IMF content**

After removing the white intermuscular fat from the muscle samples, the samples were minced thoroughly with a meat grinder, loaded into glassware and dried at 105 °C until completely dry. Samples were then weighed after crushing (marked as z), wrapped with quantitative filter paper and baked at 105 °C until samples were dry and their weight did not change. Samples were then weighed in paper bags after drying (marked as x). The dried paper bag was then placed in the Soxhlet extraction bottle, and the ether reflux device was used to reflux the sample at 65 °C until the drops are transparent. The paper bag was then placed in a fume hood to fully volatilize the ether, followed by drying at 105 °C until the weight did not change (marked as y). The measurement was repeated three times for each sample. The following formula was used to calculate the IMF content: IMF content (%) = (x−y) / z × 100%.

**RNA extraction and quality assessment**

Total RNA of the longissimus dorsi muscles was extracted using Trizol reagent (Invitrogen, Carlsbad, CA, USA) per the manufacturer’s instructions. RNA purity was measured at an OD 260/280 with a NanoDrop ND-2000 instrument (Thermo Fisher Scientific, MA, USA). RNA integrity (RIN) was evaluated by 1% agarose gel electrophoresis and Aglian 2100 (Agilent, Santa Clara, CA, USA). RNA samples with OD\text{260}/OD\text{280} ratio greater than 1.8 and RIN value greater than 7.5 were selected for sequencing.

**Library preparation and sequencing**

First-strand complementary DNA (cDNA) was synthesized using random hexamer primers and M-MuLV reverse transcriptase (RNase H-) [74], with rRNA-depleted RNA used as a template. Second-strand cDNA was then synthesized with dNTPs, DNA polymerase I and RNase H. Next, T4 DNA polymerase and Klenow DNA polymerase were used to repair and modify the ends to add an A base and ligate the sequencing adapter. The cDNA products were then purified using AMPure XP beads (Beckman Coulter, Brea, CA, USA). Finally, uracil DNA glycosylase (NEB, Ipswich, MA, USA) was used to degrade the U-containing chain to remove second-strand cDNA. The purified first-strand cDNA was enriched by PCR to obtain a cDNA
library. The quality of the libraries was assessed using an Agilent 2100 Bioanalyzer, and sequencing was performed using paired-end sequencing (2*150 bp) with the Illumina HiSeq 4000 platform (LC Sciences, Houston, TX, USA).

**Mapping of reads and transcriptome assembly**

Cutadapt was used to remove the reads that were contaminated by adapters, low-quality bases and undetermined bases. The clean reads were mapped to *Ovis aries* Ensembl release 96 (ftp://ftp.ensembl.org/pub/release-96/fasta/ovis_aries/dna/) using HISAT2 [75]. Mapped reads of each sample were assembled by StringTie [76]. The R package “edge” was used for difference statistics and visual drawing. Gffcompare was used to combine all transcripts from samples to reconstruct a comprehensive transcriptome. StringTie was used to determine the expression level for all transcripts by calculating FPKM (FPKM = [total exon fragments/mapped reads (millions)×exon length(kb)]). The R package edgeR was used to select the differentially expressed transcripts that satisfied the condition of \( \log_2 \) (fold change) ≥ 1 or \( \log_2 \) (fold change) ≤ -1 and \( P < 0.05 \).

**Identification of IncRNAs**

Known transcripts and transcripts less than 200 bp in length were removed from the data set. CPC (Coding Potential Calculator) and CNCI (Coding-Non-Coding Index) were then used to screen IncRNAs [77, 78]. When the CPC software score was less than 0.5 and the CNCI software score was less than 0, a transcript was considered a novel IncRNA.

**Enrichment analysis of differentially expressed mRNAs**

We used the Gene Ontology database (http://www.geneontology.org) and the Kyoto Encyclopedia of Genes and Genomes (http://www.kegg.jp/kegg) to annotate DEGs. The genes were mapped to GO terms and KEGG pathways based on annotation information and then the hypergeometric test was performed. The clustering map was drawn by the R package. GO terms and KEGG pathways were defined as significantly enriched when \( P < 0.05 \).

**Prediction of IncRNA target genes**

Based on the cis-and trans-regulation mechanisms of IncRNA, we identified the protein-coding genes (100-kb upstream and downstream) located on the same chromosome as the IncRNA that was a target for cis-regulation. RIsearch was used to predict the free energy of IncRNA-mRNA gene combinations on different chromosomes; combinations of IncRNA and mRNA with free energies below -11 kcal/mol were identified as trans target genes of IncRNA [79]. The results of the cis and trans-regulation were used to calculate Pearson correlations between IncRNA and mRNA expression. Cytoscape was used to plot the co-expression network.

**Verification of sequencing data**
We randomly selected five lncRNAs and five mRNAs to validate their expression using SYBR Green PCR Master Mix (Takara, Dalian, China). Primer 5 was used to design primers for the candidate genes. The sequences of the primers used are listed in Additional file 12 (Table S12). A 20-μL PCR mixture consisted of 10 μL SYBR® Premix Ex Taq II (2×), 0.5 μL forward primer (10 μM/L), 0.5 μL reverse primer (10 μM/L), 1 μL cDNA and 8 μL ddH₂O. The PCR parameters were as follows: 95 °C for 30 s; 40 cycles of 95 °C for 5 s; 60 °C for 30 s; 72 °C for 30 s; and 72 °C for 5 min. Three replicates were conducted for each sample. The 2⁻ΔΔCt method was used to quantify relative expression levels [80].

**Statistical analysis**

Data on IMF content were expressed as means± standard deviation. One-way analyses of variance in SPSS 17.0 were used to analyze experimental results. Independent sample t-tests were used to compare the IMF content of muscles at the same age. All the data from the qRT-PCR were obtained using at least three independent replicates. Differences were deemed statistically significant if p-values are smaller than 0.05.

**Abbreviations**

AFWS: Aohan fine wool sheep; IMF: Intramuscular fat; LncRNAs: long noncoding RNAs; Mth-2, Mth-4, Mth-6, Mth-12: Rams at 2, 4, 6, 12 months; LDM: Longissimus dorsi muscle; BFM: Biceps femoris muscle; RNA-seq: RNA sequencing; DEGs: Differentially expressed genes; DEMs: Differentially expressed mRNAs; DELs: Differentially expressed lncRNAs; ORF: open reading frame; GO: Gene ontology; KEGG: Kyoto encyclopedia of genes and genomes; qRT-PCR: Quantitative real-time PCR; CPC: Coding Potential Calculator; CNCI: Coding-Non-Coding Index; RIN: RNA integrity

**Declarations**

**Ethics approval and consent to participate**

All the experimental operation have conforming to the Guidelines for Experimental Animals of the Ministry of Science and Technology (Beijing, China) and were approved by the Experimental Animal Ethics Committee of Qingdao Agricultural University. The managements of laboratory animal has in keeping withLaboratory Animal-Requirements of Environment and Housing Facilities(GB 14925-2001). The written informed consent to participate was obtained from the AFWS Stud Farm in Inner Mongolia Autonomous Region. All efforts were made to minimize suffering.

**Consent for publication**

Not applicable.

**Availability of data and materials**
Additional data can be found in supplementary files. The RNA-Seq data is being uploaded and can be uploaded to the database before publication.

Competing interests

The authors declare that they have no competing interests.

Funding

The work was supported by the Earmarked Fund for Modern China Wool & Cashmere Technology Research System (CARS-39); the Project of Shandong Province Agricultural Variety Program (2019LZGC012); the National Natural Science Foundation of China (31402047), and the Project of Shandong Province Higher Educational Science and Technology Program(J18KA136). The funding body had no role in the design of the study, collection, analysis, interpretation of data, decision to publish or writing the manuscript.

Authors’ contributions

NL and JNH designed this study; FHH, RRZ, LRL, LLL, and QL participated in sample collection; FHH, RRZ, LLL and QL performed qRT-PCR validation; FHH, JL, NL, LRL and JNH analyzed the RNA-Seq data; FHH wrote the manuscript with contribution from JL, NL, LRL and JNH. All authors reviewed and approved the final manuscript.

Acknowledgments

The authors thank the staff of the Laboratory of Animal Genetics, Breeding and Reproduction, Qingdao Agricultural University for providing the research laboratories.

Authors’ information

1 College of Animal Science and Technology, Qingdao Agricultural University, Qingdao 266109, China.

2 Qufu Animal Husbandry and Veterinary Technical Service Center, Qufu 273100, China.

3 China Animal Health and Epidemiology Center, Qingdao 266032, China.

References


Tables

Table 1 IMF content of sheep (%)
**IMF content in different parts of muscles among sheep with the same age.** ** indicates that means were highly significantly different ($P < 0.01$); * indicates significant differences ($P < 0.05$); different lowercase letters indicate that means differ significantly ($P < 0.05$) between the same muscles groups of sheep of different ages; different capital letters indicate that means were highly significantly different ($P < 0.01$) between the same muscles groups of sheep of different ages.

### Table 2 Statistical data derived from RNA Sequencing

<table>
<thead>
<tr>
<th></th>
<th>Mth-2-1</th>
<th>Mth-2-2</th>
<th>Mth-2-3</th>
<th>Mth-12-1</th>
<th>Mth-12-2</th>
<th>Mth-12-3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Raw reads</strong></td>
<td>90397074</td>
<td>89528496</td>
<td>88813860</td>
<td>90619672</td>
<td>95989952</td>
<td>89577658</td>
</tr>
<tr>
<td><strong>Valid reads</strong></td>
<td>85098620</td>
<td>84675502</td>
<td>81473556</td>
<td>85252798</td>
<td>91240846</td>
<td>85201776</td>
</tr>
<tr>
<td><strong>Valid Ratio</strong></td>
<td>94.14</td>
<td>94.58</td>
<td>91.74</td>
<td>94.08</td>
<td>95.05</td>
<td>95.11</td>
</tr>
<tr>
<td><strong>Q30%</strong></td>
<td>98.02</td>
<td>97.98</td>
<td>97.83</td>
<td>98.03</td>
<td>98.15</td>
<td>98.17</td>
</tr>
<tr>
<td><strong>GC content</strong></td>
<td>49.5</td>
<td>51</td>
<td>53</td>
<td>49</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td><strong>Mapped reads</strong></td>
<td>76713303</td>
<td>74525753</td>
<td>69966338</td>
<td>77129925</td>
<td>83264100</td>
<td>77731307</td>
</tr>
<tr>
<td></td>
<td>(90.15%)</td>
<td>(88.01%)</td>
<td>(85.88%)</td>
<td>(90.47%)</td>
<td>(91.26%)</td>
<td>(91.23%)</td>
</tr>
<tr>
<td><strong>Expressed genes</strong></td>
<td>24537</td>
<td>24229</td>
<td>24295</td>
<td>24605</td>
<td>24563</td>
<td>24077</td>
</tr>
<tr>
<td><strong>Unique IncRNAs</strong></td>
<td>6495</td>
<td>6492</td>
<td>6507</td>
<td>6532</td>
<td>6541</td>
<td>6425</td>
</tr>
</tbody>
</table>
### Table 3 Critical mRNAs based on the KEGG pathways related to lipid deposition

<table>
<thead>
<tr>
<th>Critical pathways</th>
<th>critical mRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoxO signaling pathway</td>
<td>HOMER2, AGAP2, FBXO32, BNIP3, ARAF</td>
</tr>
<tr>
<td>Biosynthesis of unsaturated fatty acids</td>
<td>FADS2, SCD</td>
</tr>
<tr>
<td>Phosphonate and phosphinate metabolism</td>
<td>CHPT1</td>
</tr>
<tr>
<td>Cell adhesion molecules (CAMs)</td>
<td>CDH5, SIGLEC1</td>
</tr>
<tr>
<td>Ether lipid metabolism</td>
<td>CHPT1, PLA2G4E</td>
</tr>
<tr>
<td>PPAR signaling pathway</td>
<td>FADS2, SCD</td>
</tr>
<tr>
<td>Fatty acid elongation</td>
<td>ACAA2</td>
</tr>
<tr>
<td>mTOR signaling pathway</td>
<td>ULK1, FZD4</td>
</tr>
<tr>
<td>Fatty acid degradation</td>
<td>ACAA2</td>
</tr>
<tr>
<td>Wnt signaling pathway</td>
<td>FZD4</td>
</tr>
<tr>
<td>AMPK signaling pathway</td>
<td>ULK1, SCD</td>
</tr>
<tr>
<td>alpha-Linolenic acid metabolism</td>
<td>PLA2G4E, FADS2</td>
</tr>
<tr>
<td>Tight junction</td>
<td>CTTN, TUBA4A</td>
</tr>
<tr>
<td>Arachidonic acid metabolism</td>
<td>PLA2G4E</td>
</tr>
</tbody>
</table>

### Additional Files

**Additionnal file 1:** Table S1. Summary of protein-coding genes identified in the libraries. (XLSX 7373 kb)

**Additionnal file 2:** Table S2. (A). Summary of lncRNAs identified in the libraries. (B). Type statistics of lncRNAs identified in the libraries. Class code: “x” represents the antisense lncRNAs, “o” represents the sense lncRNAs, “j” represents the bidirectional lncRNAs, “i” represents the intronic lncRNAs, and “u” represents the intergenic lncRNAs. (XLSX 5253 kb)

**Additionnal file 3:** Table S3. Sequence information of all expressed lncRNAs found in the study. (FA 157696 kb)

**Additionnal file 4:** Table S4. The summary of differentially expressed protein-coding genes. (XLSX 212 kb)

**Additionnal file 5:** Table S5. The summary of differentially expressed lncRNAs. (XLSX 77 kb)

**Additionnal file 6:** Table S6. (A). GO enrichment analysis of the differentially expressed mRNAs. S gene number: the number of significant differentially expressed mRNAs which match to a GO term; TS gene
number: the number of significant differentially expressed mRNAs which have GO annotations; B gene number: the number of detected mRNAs which match to a GO term; TB gene number: the number of all detected mRNAs which have GO annotations. (B). KEGG enrichment analysis ($P < 0.05$) of the differentially expressed mRNAs. (XLSX 77 kb)

**Additional file 7: Table S7.** Significantly enriched KEGG pathways ($P < 0.05$) of the differentially expressed mRNAs. (XLSX 10 kb) S gene number: the number of significant differentially expressed mRNAs that match to a pathway; TS gene number: the number of significant differentially expressed mRNAs that have pathway annotations; B gene number: the number of detected mRNAs that match to a pathway term; TB gene number: the number of all detected mRNAs that have pathway annotations.

**Additional file 8: Table S8.** The differentially expressed target mRNAs of the differentially expressed IncRNA in either trans- or cis- regulatory roles. (XLSX 5919 kb)

**Additional file 9: Table S9. (A).** GO enrichment analysis of differentially expressed target mRNAs of differentially expressed IncRNAs in the study. (B) KEGG enrichment analysis of differentially expressed target mRNAs of differentially expressed IncRNAs in the study. (XLSX 138 kb)

**Additional file 10: Table S10.** The significantly enriched KEGG pathway analysis ($P < 0.05$) of target genes of differentially expressed IncRNAs. (XLSX 11 kb)

**Additional file 11: Table S11.** Critical mRNAs and their IncRNAs that were related to the lipid deposition. (XLSX 102 kb)

**Additional file 12: Table S12.** Primers used in the qRT-PCR analysis. (XLSX 99 kb)

**Figures**
Figure 1

Five differentially expressed mRNAs and five differentially expressed IncRNAs, which were detected by qRT-PCR. a: Expression levels of genes by qRT-PCR b: Expression levels of genes by RNA-seq.
**Figure 2**

LncRNA-mRNA co-expression network. The red points represent mRNAs in critical pathways. The blue points represent candidate lncRNAs, which either cis- or trans-regulate target genes.
Figure 3

GO terms and pathways analysis of differentially expressed mRNAs. a: Top 20 GO terms of differentially expressed mRNAs. The size of the points represents the number of significant differentially expressed mRNAs that matched to a GO term. The color of the points represents the significance of enrichment. b: Top 20 pathways of differentially expressed mRNAs.
Figure 4

The differentially expressed mRNAs and IncRNAs in the intramuscular fat of sheep. a: and b: The number of up-regulated and down-regulated differentially expressed mRNAs and IncRNAs. The left red bars represent the number of genes up-regulated; the right blue bars represent the number of genes down-regulated. c and d: The volcano of differentially expressed mRNAs and IncRNAs. The left blue points represent significantly decreased mRNAs and IncRNAs; gray points represent mRNAs and IncRNAs without significant changes. The right red points represent significantly increased mRNAs and IncRNAs. e and f: The heat map of differentially expressed mRNAs and IncRNAs. Red indicates that the gene had a higher expression level, and blue indicates that the gene had a lower level of expression.
Characteristics of IncRNAs and mRNAs in the intramuscular fat of sheep. a: Density distribution of IncRNAs. b: Distribution of different types of IncRNAs. c: Length distribution of IncRNAs and mRNAs. d: Distribution of exon number for IncRNAs and mRNAs. e: Expression levels (log10FPKM) and numbers of IncRNAs and mRNAs. f and g: Length distribution of ORFs of IncRNAs and coding genes.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile12PrimersusedintheqRTPCRanalysis.xlsx
- Additionalfile11CriticalDEMsandtheirIncRNAs.xlsx
- Additionalfile10SignificantKEGGoftargetsofDELs.xlsx
- Additionalfile9GOandKEGGoftargetsofIncRNAS.xlsx
- Additionalfile8TheDEMs oft heDELs.xlsx
• Additional file 7: Significantly pathway of the DEmRNAs.xlsx
• Additional file 6: GO and KEGG of DEmRNAs.xlsx
• Additional file 5: Differentially expressed lncRNAs.xlsx
• Additional file 4: Summary of differentially expressed mRNA.xlsx
• Additional file 3: merged.fa
• Additional file 2: Summary of lncRNAs.xlsx
• Additional file 1: Summary of lncRNAs.xlsx