

# Genome-wide identification and characterization of long non-coding RNAs in Longissimus dorsi skeletal muscle of Shandong black cattle and Luxi cattle

**Ruili Liu**

Qingdao Agricultural University <https://orcid.org/0000-0002-7518-7530>

**Xianxun Liu**

Qingdao Agriculture University: Qingdao Agricultural University

**Kun Yu**

Qingdao Agriculture University: Qingdao Agricultural University

**Xuejin Bai**

Qingdao Agricultural University

**Yajuan Dong** (✉ [etcenter@126.com](mailto:etcenter@126.com))

Qingdao Agricultural University <https://orcid.org/0000-0002-0911-4755>

---

## Research article

**Keywords:** Shandong black cattle, Luxi cattle, LncRNA, Longissimus dorsi, identification

**Posted Date:** December 7th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-120327/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

There is increasing understanding of the possible regulatory role of long non-coding RNAs (LncRNA). Studies on livestock have mainly focused on the regulation of cell differentiation, fat synthesis, and embryonic development. However, there has been little study of skeletal muscle of domestic animals and the potential role of LncRNA.

## Results

RNA samples were collected from longissimus dorsi muscle samples of Shandong black cattle and Luxi cattle and libraries were constructed and sequenced. A total of 1415 transcripts (of which 480 were LncRNAs) were differentially expressed ( $P < 0.05$ ) in the different breeds, and fourteen of these RNAs were randomly selected and validated by qPCR. We found that the most differentially expressed LncRNAs were found on chromosome 9, with 1164 within 50 kb of a protein-coding gene. In addition, Pearson's correlation coefficients of co-expression levels indicated a potential trans regulatory relationship between the differentially expressed LncRNAs and 43844 mRNAs ( $r > 0.9$ ). The identified co-expressed mRNAs (MYORG, DII1, EFNB2, SOX6, MYOCD, and MYLK3) are related to the formation of muscle structure, and enriched in muscle system process, strained muscle cell differentiation, muscle cell development, striated muscle tissue development, calcium signaling, and AMPK signaling. Additionally, we also found that some LncRNAs (LOC112444238, LOC101903367, LOC104975788, LOC112441863, LOC112449549, and LOC101907194) may interact with miRNAs related to cattle muscle growth and development. Based on this, we constructed a LncRNAs-miRNA-mRNA interaction network as the putative basis for biological regulation in cattle skeletal muscle. Interestingly, a candidate differential LncRNA (LOC104975788) and a protein-coding gene (Pax7) contain miR-133a binding sites and binding was confirmed by luciferase reporter assay. LOC104975788 may bind miR-133a competitively with Pax7, thus relieving the inhibitory effect of miR-133a on Pax7 to regulate skeletal muscle development. These results will provide the theoretical basis for further study of LncRNA regulation and activity in different cattle breeds.

## Conclusions

The data obtained in this study were used to predict muscle-related LncRNAs-miRNA-mRNA interaction networks, which can help elucidate the molecular mechanism of cattle muscle development. These results can be used to facilitate livestock breeding and improve livestock production.

## Background

As an important economic trait that affects the production efficiency of beef cattle, meat production traits are a research focus in the field of beef cattle genetics and breeding. The analysis of the molecular

regulation mechanisms of muscle growth can facilitate cattle breeding. Local cattle breeds in China have good meat quality but slow growth is slow, while imported cattle typically grow more quickly, but the meat quality is poor. These differences may be related to differences in skeletal muscle development among different breeds. The growth and development of beef cattle skeletal muscle can be affected by a variety of regulatory factors. Previous work mainly focused on contributions from DNA, mRNA, and miRNA, but regulatory effects of long non-coding RNA (LncRNA) on the growth and development of beef skeletal muscle remain poorly understood.

Previous studies reported contributions of LncRNA in the process of skeletal muscle proliferation and differentiation<sup>1,2</sup>. Liu identified, mapped, and determined the global and skeletal muscle expression patterns of 7188 LncRNAs, finding that these LncRNAs had similar open reading frames and expression levels as those of other mammalian LncRNAs. Subsequently, Yue identified a highly expressed LncRNA in muscle tissue, LncRNA-YYW. Microarray analysis showed that LncRNA-YYW positively regulated the expression of growth hormone-1 and its downstream genes AKT1 and PIK3CD in bovine myoblasts, and promoted myoblast proliferation<sup>3</sup>. RNA samples from the longissimus pectoralis muscle of Limousin beef cattle were subjected to paired-end RNA sequencing, and some of the identified expressed LncRNAs mapped to quantitative trait loci of meat quality traits<sup>4</sup>. These results suggest that LncRNA may regulate the growth and development of beef skeletal muscle and meat quality. Although LncRNAs lack coding capacity, many LncRNAs act in various biological processes, serving as an additional level of genomic regulation.

The first embryo transfer calf in China was obtained from a vitrified frozen somatic cell cloned embryo, which combined Japan black cattle, Luxi cattle, and Bohai black cattle. The new generation core breeding group of this new breed, Shandong black cattle, was established by filial generation. The new generation core breeding group hybridized with Shandong black cattle, and the second generation group of Japan black cattle (3/4) was crossed with luxi cattle (1/4) to produce the new generation (3/4) × bohai black cattle (1/4). Ideal black cattle and Shandong black cattle were selected from the second-generation group for cross-cross fixation. In this stage, several good lines were cross-bred with distant relatives, and then selected and the excellent lines were preserved. After four generations, this line was finally bred into a new Shandong black cattle variety matching line and used as a bull. In 2015, the Shandong black cattle variety was approved by the National Animal and Poultry Genetic Resources Committee as a new population and successfully established as a new cultivated variety in China<sup>5,6</sup>.

In recent years, there has been growing research of LncRNA, but most work has mainly focused on the effects of neurologic diseases, tumors, embryonic development, and cell differentiation in human and mouse, with little research on domestic animals. The growth and development of beef cattle skeletal muscle are highly regulated processes. Although the contributions of DNA, mRNA, and miRNA to this regulation have been studied, the extent to which LncRNA regulates beef skeletal muscle growth and development remains poorly understood. To investigate the potential contribution of LncRNA, RNA samples were prepared from longissimus dorsi muscle tissues of Shandong black cattle (hybrid offspring) and Luxi Cattle (the first maternal generation). RNA-seq technology was used to identify

LncRNA transcripts and their genomic characteristics, and LncRNAs differentially expressed in skeletal muscle tissues in different breeds of beef cattle were determined. Using these data, we constructed an interaction network of LncRNA-miRNA-mRNA and muscle development which can guide future breeding efforts.

## Methods

### 1. Ethics statement

The methods used in this study were performed in accordance with the guidelines of Good Experimental Practices adopted by the Institute of Animal Science (Qingdao

Agricultural University, Qingdao, China). All surgical procedures involving cattle were performed according to the approved protocols of the Biological Studies Animal Care and Use Committee, Shandong Province, China.

### 2. Animal and tissue preparation

Shandong black cattle and Luxi cattle were used in this study and were purchased from Shandong Black Cattle Technology Co., Ltd.. Cattle were fed three times and received approximately equal amounts of green coarse feed and concentrate feed according to standard NY5127–2002 pollution-free feeding management of beef cattle. Cattle were supplied sufficient drinking water. Six 18-month-old healthy beef cattle were randomly selected for this study. The selected cattle had no scratches, scars, and scabs on their bodies, with no fat deposits in the internal organs or abdomen. No disease was found during examination, and all physiological and biochemical indexes were normal. Longissimus dorsi muscle samples were collected and immediately frozen in liquid nitrogen for RNA extraction. The cattle after the experiment were euthanized. First, a 1-3% sodium pentobarbital solution was prepared with physiological saline, and then intravenously injected. The injection dose is 90~135mg/kg.

### 3. Hematoxylin and eosin (HE) staining of muscle tissue and fast/slow muscle fiber fluorescence staining

Paraffin sections were prepared from muscle tissue fixed with 4% paraformaldehyde. The HE staining protocol was performed as described previously<sup>7</sup>. Briefly, dewaxing, covering with water, haematoxylin staining, washing with water, 5% acetic acid differentiation, eosin staining, dehydration, natural drying, sealing, and image acquisition were performed.

Tissue sections were placed in a box filled with EDTA antigen repair buffer(Purchased from Shanghai Beyotime Biotechnology Co., Ltd) (ph 8.0) for antigen repair. After natural cooling, the slides were washed in PBS (pH7.4) with three washes of five minutes each. BSA was added for blocking, and then the antibody was added, followed by DAPI to stain the nucleus. An autofluorescence quenching agent was added to the slices for 5 min and the samples were then washed with running water for 10 min. After natural

drying, the film was sealed with an anti-fluorescence quenching sealing agent. Finally, the slices were observed under the fluorescence microscope and images were collected.

Image-Pro Plus software was used to count the images and measure the surface area. SPASS software was used for statistical analysis to determine significant differences.

#### **4. RNA extraction, library construction, and sequencing**

Total RNA samples were isolated using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. RNA degradation and contamination were monitored on 1 % agarose gels. RNA purity was checked using a NanoPhotometer® spectrophotometer (IMPLEN, Los Angeles, CA, USA). RNA concentration was measured using a Qubit® RNA Assay Kit in a Qubit® 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). RNA integrity was assessed using a RNA Nano 6000 Assay Kit in a Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Only samples with RNA Integrity Number (RIN) scores > 8 were used for sequencing. A total of 3 µg RNA per sample was used as the input material for RNA library preparation.

#### **5. Transcriptome assembly**

The original sequencing data were processed by Fastp software with parameters “-Q 20 -P90” with disjointing sequence and low-quality sequence. Clean data were obtained by removing reads containing adapters, reads containing over 10 % of poly(N), and low-quality reads (>50 % of the bases had Phred quality scores ≤ 10) from the raw data. The clean data were then calculated. All downstream analyses were based on the high quality clean data. The *Bos taurus* genome reference genome and gene model annotation files were downloaded from the NCBI database (CHIR\_1.0, NCBI) [60]. An index of the reference genome was built using Bowtie v2.0.6 [61, 62] and paired-end clean reads were aligned to the reference genome using HISAT v2-2.1.0 [63]. The mapped reads from each library were assembled with Cufflinks v2.2.1 [64], using Cufflinks with ‘min-frags-per-transfrag = 0’ and ‘-library-typefr-firststrand’, and other parameters set as default.

#### **6. Screening of LncRNA**

Stringtie was used to sort reads into different classes and then generate a mosaic map for each class. Based on the length of LncRNA, and the characteristics of non-protein coding sequences, we established strict screening conditions to screen LncRNA as follows:

- (1) Transcripts equal to or longer than 200 bp in length, and containing two or more exons;
- (2) Transcript read coverage of at least five reads;
- (3) No transcripts of known mRNA or other specific non-coding RNAs (rRNA, tRNA, snoRNA, or snRNA). This screening was based on gffcompare (<http://ccb.jhu.edu/software/stringtie/gff.shtml>) using the same species annotation file;

(4) According to the class in the comparison result\_ Code information ("U", "I", "X") was used to screen potential lincrna and intronic LncRNA anti-sense LncRNA.

Coding potential prediction screening:

To further assess if the identified transcripts are LncRNAs, a variety of coding potential analysis software was integrated, including CNCI analysis, CPC analysis, Pfam protein domain analysis, and CPAT analysis (only for animals)<sup>8</sup>. The transcripts identified as non-coding by several methods were the final Novell LncRNA dataset.

## **7. Differential expression analysis**

Comparison of FPKM<sup>9</sup> (fragments per kilobase of transcript per million mapped reads) for different genes is a very effective tool for quantitative estimation of gene expression based on RNA-Seq data. This method can eliminate the influence of gene length and sequencing quantity on the determination of gene expression levels. The calculated gene expression values allow direct comparison of gene expression differences among different varieties. Here, we used DEseq<sup>10</sup> to analyze the differential expression between the treatment group with the reference group, and selected  $|\text{Log}_2\text{ratio}| \geq 1$  and  $Q < 0.05$  genes as indicative of significant differential expression. The numbers of up-regulated and down-regulated genes were also obtained.

## **8. Prediction and analysis of target genes of differential LncRNA**

LncRNA can act on target genes by Cis or Trans. Cis effect refers to LncRNA acting on adjacent target genes<sup>11</sup>. To predict the cis regulatory genes of LncRNA, we screened the 50 kb of sequence upstream and downstream of LncRNA and looked for potential target genes. The coexpression relationship between LncRNA and mRNA was described by Trans. According to the correlation coefficient between LncRNA and mRNA expression (correlation coefficient  $\text{cor} \geq 0.9$ ); WGCNA (Weighted correlation network analysis) was used to predict the target genes of LncRNA for cluster analysis and functional enrichment analysis of LncRNA target genes. P-value  $< 0.05$  was set as the significance threshold.

## **9. Analysis of enrichment pathway of GO and KEGG**

GO (gene ontology) can be used for enrichment analysis of target genes with differential expression and Goseq<sup>12</sup> was used to analyze the target genes of differentially expressed LncRNA. A value of  $P < 0.05$  is considered as significant enrichment of differentially expressed genes.

The KEGG database describes advanced functions and utility of biological systems such as cells, organisms, and ecosystems (<http://www.genome.jp/kegg/>). We used KOBAS V3.0 software to detect the enrichment of LncRNA target genes differentially expressed in KEGG pathway.

## **10. Prediction of LncRNA-miRNA-mRNA interaction**

As competing endogenous RNAs (ceRNA), LncRNA can act as a sponge to adsorb miRNA and affect mRNA expression. Therefore, to further understand the function of LncRNA in the growth and development of skeletal muscle, we used TargetScan([http://www.targetscan.org/vert\\_71/](http://www.targetscan.org/vert_71/)) and miRanda(<http://www.microrna.org/microrna/home.do>) software to predict the differentially expressed miRNAs adsorbed by LncRNA and downstream-regulated mRNA. We then constructed an LncRNA-miRNA-mRNA interaction network map using Cytoscape version 3.5.1.

## 11. Luciferase reporter assay

Cells were seeded in 96-well plates at a density of  $5 \times 10^3$  cells (HEK-293T) per well, 24 h before transfection. The cells were co-transfected with a mixture of 50 ng Firefly luciferase (FL) reporter vectors, 5 ng Renilla luciferase (RL) reporter vectors (pRL-TK), and miRNA mimics at the indicated concentration. The miRNA mimics were obtained from Life Technologies. After 48 h, the luciferase activity was measured with a dual luciferase reporter assay system using the psiCHECK-2 vector (Promega, Madison, WI). The LOC104975788 and LOC536229 (Pax7) sequences were separately cloned into the reporter gene vector (psiCHECK-2) to synthesize the predicted miRNA mimics and control. The potential binding target of each miRNA was cloned into the 3'UTR region of r-luciferase (hrluc), and then co-transfected with the miRNA to determine the activity of R-Luciferase. F-Luciferase (hluc +) was used as an internal reference to correct for differences in transfection efficiency between different samples. The miRNA mimics and psicheck-LOC104975788 or psicheck-LOC536229 (Pax7) were co-transfected into 293T cells. The expression level of reporter genes was detected using a multi-functional enzyme labeling instrument, and the miRNAs that exhibited down-regulated reporter gene expression were further screened.

## 12. Validation of sequencing data by quantitative reverse transcription PCR (qRT-PCR)

The reaction system (20  $\mu$ L) consisted of the following: 1  $\mu$ L of template cDNA, 10  $\mu$ L each of the upstream and downstream primers, and 5 ml (5 $\times$ 1 mL vials) of RNase-free water. The thermal cycling procedure was as follows: 94  $^{\circ}$ C for 10 min, 94  $^{\circ}$ C for 30 s, 60  $^{\circ}$ C for 30 s, and 72  $^{\circ}$ C for 40 s, with 40 cycles. The expression of GAPDH was calculated by the  $2^{-\Delta\Delta CT}$  method. Primers used for qRT-PCR as shown in Supplementary documents Table S1.

# Results

To investigate the potential regulatory effects of LncRNA on the growth and development of beef skeletal muscle, we analyzed samples from muscles of different breeds of beef cattle.

## 1. Apparent differences in muscle fibres in different breeds of beef cattle

First, we obtained samples from Shandong black cattle and Luxi cattle and performed fluorescence staining sections of fast and slow muscle fibers (Fig. 1a) and HE staining (Fig. 1b). The results showed that significant differences in muscle fibers of longissimus dorsi muscles between Shandong black cattle and Luxi cattle. IPP software analysis showed that the average area of muscle fibres of Shandong black

cattle was significantly larger than that of Luxi cattle ( $P < 0.05$ ), with significant differences in the muscle fibre density and the ratio of fast-twitch fibers to slow-twitch fibers and slow-twitch fibers to muscle fiber area ( $P < 0.05$ ), but not in other muscle fibre properties ( $P > 0.05$ ) (Table 1).

**Table 1** Comparison of muscle fibre characteristics and growth characteristics of different breeds of cattle

Characteristics	Shandong black cattle	Luxi cattle
Area ( $\mu\text{m}^2$ )	5490.222 $\pm$ 184.649*	4869.008 $\pm$ 69.596
Diameter ( $\mu\text{m}$ )	106.837 $\pm$ 12.537	120.491 $\pm$ 4.324
Length ( $\mu\text{m}$ )	174.220 $\pm$ 7.395	142.435 $\pm$ 0.968
Density (Number of muscle fibres/ Muscle fiber area, EA/ $\mu\text{m}^2$ )	4887.848 $\pm$ 373.586**	7172.966 $\pm$ 319.501
Number of muscle fibres(EA)	57.667 $\pm$ 4.333	39 $\pm$ 1.732
Fast-twitch fibers/Slow-twitch fibers	0.412 $\pm$ 0.096**	3.280 $\pm$ 1.082
Fast-twitch fibers/Muscle fiber area	0.200 $\pm$ 0.0342	0.652 $\pm$ 0.110
Slow-twitch fibers/Muscle fiber area	0.508 $\pm$ 0.046*	0.246 $\pm$ 0.072
Weight (kg)	509.667 $\pm$ 2.026	489.333 $\pm$ 1.764

**Note:** In the table, \* indicates a significant difference ( $P < 0.05$ ) and \*\* indicates a extremely significant difference ( $P < 0.01$ ).

## 2. Overview of sequencing data of longissimus dorsi in beef cattle

RNA was extracted from longissimus dorsi muscle samples of Shandong black cattle and Luxi cattle and sequenced. Totals of 105345396 and 126735736 reads were obtained from Shandong black cattle and Luxi cattle, respectively, with 101145946 and 122463284 mapped reads. In the data for Luxi cattle longissimus dorsi muscle, the percentages of exons, introns, and genes were 43.26%, 41.34%, and 15.4%, respectively, with 38.16%, 48.3%, and 13.54%, respectively, for Shandong black cattle longissimus dorsi muscle tissue, as shown in Fig.2.

## 3. Analysis of differential mRNA expression

A total of 20104 transcripts were obtained from the two libraries, including 18859 and 18976 mRNAs for Luxi cattle and Shandong black cattle, respectively. According to the screening criteria of  $|FC| \geq 2$  and  $P < 0.05$ , 1415 differentially expressed mRNA (Table S2) were found, with 939 mRNAs significantly up-regulated and 476 down-regulated in Luxi cattle. The 92 mRNAs were specifically expressed in cattle,

with 30 and 62 mRNAs specifically expressed in Shandong black cattle and Luxi Cattle, respectively (Table S2). The top 40 genes with significant differential expression in the two breeds in cattle are listed in Table S3, and suggest genes that could be used for the functional comparison and analysis of differences among varieties.

#### **4. Functional analysis of differential mRNA**

Gene Ontology (GO) is an international standardized classification system of gene function that includes three categories: biological process (BP), molecular function (MF), and cell composition (CC). We performed GO enrichment analysis on 1415 differentially expressed genes (Fig. 3a and b), and the results showed 2662 items ( $P < 0.05$ ), with 388 were enriched in cell composition (CC), 361 in molecular function (MF), and 1913 in biological process (BP) in up-regulated genes. The enrichment results of down-regulated genes showed 1883 items in total ( $P < 0.05$ ), with 251 enriched in cell composition (CC), 247 in molecular function (MF), and 1385 in biological process (BP). In the muscle regulation system, muscle contraction is enriched in the first 50 GO terms, indicating that some differential genes may participate in muscle development and function and other life activities. Muscle-related GO terms are shown in Table 2.

**Table 2** Partial GO term related muscle

Term_type	GO accession	Description	Up-regulated genes	Down-regulated genes
Biological _process	GO:0060173	Limb development.	LMOD1,TNNI2,FGF10,N	AHI1,AQP2,CEBP
	GO:0010761	Myoblast migration. Fast muscle to slow muscle conversion.	OTCH1,ARMH2,HOXA11,MYH3,NOTCH4,NRARP,OSR1,PODXL,WNT4,AA	A,CITED2,CSRP3,CXCL10,EN1,FGF9,FHOD3,FRAS1,G
	GO:0030239	myofibrillar cones.	RD,ACVRL1,ADAMTS2,	DF11,GSC,ID2,KL
	GO:0035295	Hippo cascade signaling pathway. Myotube development.	ADGRG1,AGTR2,BMPER,ERRFI1,EYA1,FOXF2,GA	B,KRT19,LMOD2,LOC107133268,LO
	GO:0008543	Fibroblast growth factor receptor signaling pathway.	TA6,HIF1AN,HOXA10,H	C112445823
	GO:0007219	Notch signaling pathway.	HYAL2,IGFBP5,ITGA6,JA	H10MYOM3,MYO
	GO:0035148	Myotube formation.	G2,LAMA5,LMOD1,MDK,MYBPC2,MYH11,MYLK3,MYOZ3,NEBL,NEURL1,NFKBIA,NOTCH3,NR4A3,OSR2,PAK1,PDPN,PEA	Z2,PAX2,PERP,PO
	GO:0031432	Actin binding.	3,OSR2,PAK1,PDPN,PEA	H10MYOM3,MYO
	GO:0005523	Protomyosin binding.	R1,PITX1,PP3CA,PPP3R1,PRKD2,PTP4A3,RAB3A.	P63.
	GO:0004672	Muscle structural components.	DYRK3,MYH11,PLK1,TN	ANKRD2,AXL,DY
Molecular _function	GO:0004713	Protein kinase activity.	K2,ALK,BMX,CCDC88C,CLK1,CNTN3,EFNB2,EM	NC111,FGF9,GAS2,L2,KIF1A,KIF5C,
	GO:0008307	tyrosine kinase activity.	L1,EML6,EPHB1,KDR,KI	MST1R,MYH10,M
	GO:0004713	Microtubule binding.	F20A,KIF22,KIF26A,MY	YO1G,TRIM54,AL
	GO:0008017	Protein serine / threonine kinase activity.	H1,MYH3,MYH4,MYH8,MYO10,MYO5B,NEBL,N	PK1,ANKRD1,CSR
	GO:0004712	Insulin like growth factor binding.	TRK3,PRAG1,REEP1,RET	P3,FAM221A,KRT
	GO:0005520	Microtubule binding.	,S100A9,STIM1,TIE1,AC	19,LIG3,LMOD2,M
	GO:0015631	binding.	VRL1,BEAN1,CAMK1D,	YBPH,MYL6B,MY
Cellular_c omponent	GO:0005865	Myofibril	CDKL5,EEF2K,GRK5,HT	2,STK32A,TSSK3,
	GO:0030016		RA1,IGFBP7,IRAK2,ITG	
			A6,LMOD1,MYBPC2,MYLK3,MYLK4,MYLPF,NEK7,NRK,PAK1,PHKA1,PI	VRK2,WISP1.
			M1,PKN3,PNCK,PRKD2,PROX1,RPS6KA3,SBK2,SP	
			EG,TPM1,TRIM63,TRPM7,TSSK1B,TSSK2,TSSK6,WISP2,WNK2.	
			TNNT2,TNNT3,TPM1,CA	FHOD3,LMOD2,T
			CNA1S,FBXL22,FHL3,IGN1,MYBPC2,MYH1,MYH3,MYH4,MYH8,MYL3,MYOZ3,NEBL,NOS1,PDE	NNI1,ANKRD1,ANKRD2,CRYAB,CSR
			4B,PPP3CA,TRIM32,TRIM63.	P3,KRT19,LOC104969184,MYBPH,M
				YOM3,MYOZ2,NE
				XN,PDLIM1,SCN3
				B,SCN5A,SYNPO2
				L,TRIM54,TWF2.

Different genes coordinate with each other to perform their biological functions in organisms. We next performed KEGG pathway analysis of the differentially expressed genes to identify the related signal transduction and metabolic pathways (Fig. 3c and d). The results showed enrichment of 939 up-regulated genes in 243 pathways ( $P < 0.05$ ), and enrichment of 476 down-regulated genes in 239 pathways, including metabolism, disease, cell communication and endocrine regulation (Table S4). The enriched pathways ( $P < 0.05$ ) included some signaling pathways related to skeletal muscle growth, development, and function (Table 3), such as the Wnt signaling pathway, the calcium signal transduction

pathway, gap junction, the Hippo signaling pathway, the AMPK signaling pathway, and the thyroid hormone synthesis pathway.

**Table 3** Partial KEGG pathway related muscle

KEGG term (level 2)	UP-regulated genes	Down-regulated genes
Wnt signaling pathway	NFATC2,NKD2,TCF7L1,CTNNB1, WNT4 PPP3R1,PPP3CA	WNT5A,SFRP5
Gap junction	ADCY4,ADCY1,PDGFB	ADCY7
TGF-beta signaling pathway	LOC535280	
Hedgehog signaling pathway	GRK5	PTCH2
FoxO signaling pathway	S1PR1,FOXO6,PIK3R3,GADD45B ,GADD45G,PLK1	
Hippo signaling pathway	NKD2,TCF7L1,WNT4,TEAD1	WNT5A,ITGB2,DLG2,ID2,BMP8B
Thyroid hormone synthesis	ATP1A1,ADCY4,ADCY1,TSHR,C REB3L1	GPX3,ADCY7
Ras signaling pathway	RAPGEF5,PLA2G5,EFNA4,ANGP T2,PIK3R3,KDR,CALM3,ANGPT1 ,PAK1,FGF10,LOC521224,PLCE1, IGF2,PDGFB,VEGFA	PLA2G2D1,RASGRP1,RRAS2,FGF9,R ASGRP4
TNF signaling pathway	NFKBIA,PIK3R3,SOCS3,CREB3L 1,GRO1	CXCL10,CASP3
AMPK signaling pathway	PFKFB2,PIK3R3,EEF2K,PFKFB3P FKM,PPARGC1A,CREB3L1	
PI3K-Akt signaling pathway	COL6A2,COL6A1,COL4A5,COL4 A6,LOC530102,COL4A2,ITGA6,C OL1A1,LAMA5,NR4A1,ITGB8,C OL1A2,EFNA4,ANGPT2,PIK3R3, KDR,LAMA3,DDIT4,ANGPT1,PK N3,FGF10,RXRA,COL6A3,CREB3 L1,IGF2,PDGFB,VEGFA,LAMB2	PPP2R2C,HSP90AB1,FGF9,G6PC,PC K1
cGMP - PKG signaling pathway		PDE3B,ADRA1A,ADORA1,ADCY7
Rap1 signaling pathway		PFN2,LOC101903261,RAPGEF4,FGF9 ,ADCY7
Calcium signaling pathway		RYR3,ADRA1A,PTGER3,GNA15,AD CY7

## 5. Screening of differential LncRNAs

A total of 8427 LncRNAs were found in longissimus dorsi muscle of Shandong black cattle and Luxi cattle. Of these, 3498 LncRNAs were annotated in non-coding regions and the remaining 4929 LncRNAs

were in regions that are not annotated. According to the screening criteria of  $|FC| \geq 2$  and  $P < 0.05$ , 480 differentially expressed LncRNAs were found, with 245 and 235 LncRNAs significantly up-regulated and down regulated in Luxi Cattle compared to the levels in Shandong black cattle (Table S5). A total of 92 LncRNAs were only expressed in one cattle breed, with 30 specific LncRNAs in Shandong black cattle and 62 specific LncRNAs in Luxi cattle (Table S5). These breed-specific longissimus dorsi muscle LncRNAs included eight annotated LncRNAs in Shandong black cattle and 13 annotated LncRNAs in Luxi cattle, as shown in Table 4. The LncRNAs specifically expressed in Shandong black cattle and Luxi cattle, the first 40 unexplained LncRNAs and annotated LncRNAs are listed in Table S6.

**Table 4** LncRNA specially expressed in Shandong Black cattle and Luxi cattle

Breed	Genes	Genomic position	P-value	q-value
Luxi cattle	LOC112447810	chrNC_037335.1:59843881-59851000	0.002587626	0.03702486
	LOC112448752	chrNC_037338.1:13047319-13056419	3.14E-06	0.000158716
	LOC101905635	chrNC_037340.1:76723459-76739979	0.000910843	0.016809347
	LOC112449550	chrNC_037341.1:76478493-76487345	5.40E-06	0.000252129
	LOC104974324	chrNC_037342.1:75777823-75780010	0.000204918	0.005162159
	LOC104974692	chrNC_037344.1:67130432-67132302	2.33E-05	0.000858982
	LOC112442200	chrNC_037345.1:45609576-45616675	0.001584248	0.025977108
	LOC112442383	chrNC_037345.1:62159119-62160399	0.000214833	0.005332407
	LOC112443129	chrNC_037348.1:25091115-25128295	0.003878832	0.049737964
	LOC101903141	chrNC_037353.1:47669900-47687837	7.72E-13	1.95E-10
	LOC112444607	chrNC_037354.1:36287211-36298269	9.39E-07	5.76E-05
	LOC112446394	chrNC_037331.1:68871351-68873779	1.62E-08	1.64E-06
	LOC101907194	chrNC_037332.1:26047494-26049444	3.59E-12	8.06E-10
Shandong Black cattle	LOC112447868	chrNC_037335.1:410273-419534	4.43E-06	0.000212223
	LOC112449245	chrNC_037340.1:3165193-3220562	7.90E-15	2.68E-12
	LOC112441643	chrNC_037342.1:54862373-54875026	4.23E-05	0.001397951
	LOC112441659	chrNC_037342.1:78782537-78785977	0.000658605	0.01318643
	LOC104975164	chrNC_037346.1:61277209-61278623	6.05E-05	0.001918059
	LOC112442997	chrNC_037347.1:23104372-23114203	0.000324072	0.007462557

LOC100847412	chrNC_037348.1:32211463-32219891	1.13E-06	6.79E-05
LOC100847415	chrNC_037357.1:18645452-18649324	2.19E-06	0.000118565

We constructed a Circos diagram to visually display the distribution of differential LncRNAs and differential mRNAs (Fig. 4a). The results showed that most differentially expressed LncRNAs mapped to chromosome 9, with others mapping to chromosomes 3, 1, 11, and 7. Pearson correlation coefficient analysis of differentially expressed LncRNA and differential mRNA genes (Fig. 4b) showed 43844 pairs of co-expressed LncRNAs and differential mRNA genes with correlation coefficient >0.8 (Table S7). Based on genomic location, 387 differentially expressed LncRNAs were selected for further study.

## 6. Prediction of differentially expressed LncRNAs target genes

LncRNAs can regulate the expression of adjacent mRNAs. We analyzed the LncRNAs and protein-coding genes by analyzing mRNAs within 50 kb of LncRNAs. The results showed that 387 of the differentially expressed LncRNAs targeted 1164 target genes. In these predictions, one LncRNA targeted multiple mRNA and many LncRNAs targeted the same mRNA (Table S8). Of these, LOC112448071, LOC112444635, LOC112445963, LOC104975064, LOC101903261, LOC535280, and LOC521224 can target genes related to muscle development, including MYORG, Wnt4, PAK1, and ADCY7.

The potential target genes of LncRNA were next subjected to GO and KEGG analysis (Fig. 5). We found 130 significantly enriched GO items ( $P < 0.05$ ), with 20 related to the regulation of muscle development (Table 5). The 40 most enriched terms include multicellular organization development, single multicellular organization process, and multicellular organic process (Table S9). The GO terms related to muscle development are listed in Table 5. KEGG analysis showed enrichment of 1164 potential target genes in 299 pathways, with the 40 most significantly enriched pathways listed in Table S9. Some of these are related to muscle development (Table 6), including the calcium signaling pathway, the AMPK signaling pathway, the cGMP-PKG signaling pathway, and the PPAR signaling pathway.

**Table 5** Partial GO term related muscle

GO term (level 2)	Tagert genes
Muscle structure development	PITX1,MYOZ3,EPHB1,KY,MYORG,TRIM32,VGLL2,BVES,CITED2,DLL1,FHL2,LOC101903367,COL3A1,TRIM54,PLPP7,NOTCH1,RXRA,FGF9,HOXD9,CXADR,EFNB2,NEBL,GPCPD1,LAMA5,TCF15,JPH2,NFATC2,ADGRB1,CHD7,CRYAB,USP2,SOX6,MYBPH,ZBTB18,TBX1,LMOD1,TNNI1,NOS1,TBX1,MYLK3,IGFBP5,SPEG,MYBPC2,HLF,CHRND,MYH10,MYOCD,FGF10,MYOM3,WNT4,CACNA2D2,MYL3,MNF1,PIM1,JARID2,FHOD3,GATA6,MYH11,MYLPF,ANKRD1,ANKRD2,HIF1AN,NEURL1,MEF2D,CSR3,IGF2,MBNL3,LOC112445030,S1PR1,NEXN,COL6A3,HDAC9,LMOD2,MYF5,MYL6B,PKP2,MKL1,MYOZ2,PPP3CA,UCHL1,CXCL10,SHOX2,OLFM2
Muscle tissue development	PITX1,EPHB1,MYORG,VGLL2,BVES,CITED2,DLL1,FHL2,LOC101903367,COL3A1,OSR1,ID2,NOTCH1,RXRA,FGF9,HOXD9,CXADR,EFNB2,NEBL,GPCPD1,JPH2,CHD7,EYA1,SOX6,ZBTB18,PROX1,TNNI1,TBX1,MYLK3,IGFBP5,HLF,CHRND,MYH10,MYOCD,ADAMTS9,MYL3,MNF1,PIM1,JARID2,FHOD3,GATA6,MYH11,MYLPF,ANKRD1,ANKRD2,NEURL1,MEF2D,CSR3,S1PR1,NEXN,HDAC9,TP63,MYF5,MYL6B,PKP2,PPP3CA,SHOX2
Striated muscle tissue development	PITX1,EPHB1,MYORG,VGLL2,BVES,CITED2,DLL1,FHL2,LOC101903367,ID2,NOTCH1,RXRA,FGF9,HOXD9,CXADR,EFNB2,NEBL,GPCPD1,JPH2,CHD7,EYA1,SOX6,ZBTB18,PROX1,TNNI1,TBX1,MYLK3,HLF,CHRND,MYH10,MYOCD,ADAMTS9,MYL3,MNF1,PIM1,JARID2,FHOD3,GATA6,MYH11,MYLPF,ANKRD1,ANKRD2,NEURL1,MEF2D,CSR3,S1PR1,NEXN,HDAC9,MYF5,MYL6B,PKP2,PPP3CA,SHOX2
Muscle organ development	PITX1,EPHB1,KY,MYORG,VGLL2,BVES,CITED2,DLL1,LOC101903367,COL3A1,NOTCH1,RXRA,FGF9,HOXD9,CXADR,EFNB2,GPCPD1,LAMA5,TCF15,JPH2,ADGRB1,CHD7,CRYAB,USP2,SOX6,ZBTB18,PROX1,TNNI1,TBX1,SPEG,HLF,CHRND,MYOCD,MYL3,MNF1,PIM1,JARID2,GATA6,MYLPF,ANKRD1,ANKRD2,NEURL1,MEF2D,CSR3,LOC112445030,S1PR1,COL6A3,HDAC9,MYF5,MYL6B,PKP2,PPP3CA,SHOX2
Striated muscle cell differentiation	MYOZ3,MYORG,TRIM32,DLL1,FHL2,PLPP7,NOTCH1,RXRA,CXADR,EFNB2,NEBL,NFATC2,ADGRB1,SOX6,MYBPH,PROX1,LMOD1,NOS1,TBX1,MYLK3,IGFBP5,MYBPC2,MYH10,MYOCD,MYOM3,CACNA2D2,FHOD3,GATA6,MYH11,ANKRD2,CSR3,IGF2,NEXN,HDAC9,LMOD2,MYF5,MYOZ2,PPP3CA,UCHL1,CXCL10,SHOX2

**Table 6** Partial KEGG pathway related muscle

KEGG term (level 2)	Tagert genes
Calcium signaling pathway	ATP2A1,RYR3,MYLK3,P2RX6,CACNA1S,MYLK4,ADRA1A,PTGER3,PHKB,PLCD3,NOS1,CALM3,PLCD4,GNA15,ADCY4,STIM1,ORAI1,ATP2B2,HRH2,PPP3CA,ADCY1,PKA1,ADCY7,PLCE1
PPAR signaling pathway	PLIN2,ME3,SCD5,FABP3,FABP7,PLIN1,RXRA,FABP4,PCK1,GK,FADS2,A DIPOQ
Vascular smooth muscle contraction	PPP1R14A,MYLK3,CACNA1S,MYLK4,MYH11,ADRA1A,MYH10,CALM3,ADCY4,ADCY1,MYH4,MYH1,ADCY7,MYL6B
cGMP - PKG signaling pathway	ATP2A1,NFATC2,MYLK3,TRPC6,CACNA1S,MYLK4,PDE3B,ADRA1A,ERRI1,ATP1A1,CALM3,ADORA1,ADCY4,ATP2B2,PPP3CA,ADCY1,ADCY7,CREB3L1
Thyroid hormone signaling pathway	ATP2A1,ESR1,RCAN1,PLCD3,ATP1A1,WNT4,DIO2,PLCD4,PFKM,RXRA,SLC2A1,PLCE1,NOTCH1
AMPK signaling pathway	PPP2R2C,SCD5,CAB39,ADRA1A,LEPR,EEF2K,PFKFB3,PFKM,FBP1,LIPE,PPARGC1A,CREB3L1,PCK1,ADIPOQ
Cardiac muscle contraction	ATP2A1,CACNA1S,CACNA2D2,ATP1A1,MYL3,COX7A1,ACTC1,UQCRB,TPM1,LOC101902754,CACNA2D3

## 7. LncRNA-miRNA-mRNA network interaction analysis

We used Targetscan, Miranda, and PITA qtar software<sup>13</sup> to predict the target gene relationships between mRNA and miRNA, and used Miranda and PITA qtar to predict the target gene relationships between miRNA and LncRNA. The top 5% most related pairs were visualized using the average degree method and the Fruchterman- Reingold layout method in Gephi(Fig. 6a). LncRNA can function as ceRNA to regulate the expression of downstream genes at the post-transcriptional level by binding miRNAs. We selected several miRNAs related to muscle development (including miR-1, miR-133, miR-206, miR-181, miR-21, and mir-23a) to predict and screen downstream-regulated mRNA using Targetscan and Miranda software. We then constructed an LncRNA-miRNA-mRNA interaction network diagram using Cytoscape version 3.5.1 (Fig. 6b). LncRNAs that may be involved in the development of beef cattle muscle are summarized in Table 7.

**Table 7** LncRNA-miRNA-mRNA related to skeletal muscle development and growth

LncRNA	microRNA	Targets (including Potential Tissues & Cells)	(Targets-mi R)
LOC525506, LOC526226, LOC11244423, LOC101903367	miR-135a/b	Smad5, JAK2, MEF2C	C2C12, Hodgkin lymphoma
LOC540051, LOC112448318, LOC112441863, LOC112449549, LOC101907194	miR-143/145	SRF, myocardin, Nkx2-5	smooth muscle, skeletal muscle
LOC112448962, LOC104975788, LOC112449549	miR-133a/b	SRF, KLF15, Igf1R, dynamin 2, Pax7, MAML-1	Runx2, C2C12, skeletal muscle, cardiac muscle
LOC112449031	miR-365-3p	Set1	skeletal muscle
LOC101904174	miR-214	Nras	C2C12, skeletal muscle, cardiac muscle
LOC104970868	miR-449a	Cdk6, Cdc25a	C2C12, skeletal muscle
LOC112446607, LOC112441811, LOC101906545	miR-423	SRF, SUFU	skeletal muscle
LOC104975788, LOC112447073	miR-103	Wnt3a, RAI14	C2C12, skeletal muscle
LOC100847415	miR-17-5p	Mfn2, RB1, ENH1	Skeletal muscle satellite cells
LOC525506, LOC540051, LOC514189, LOC112444238, LOC112446526, LOC112448972, LOC112446078	miR-21	PTEN, PDCD4	skeletal muscle, cardiac muscle, smooth muscle
LOC525506	miR-1	KCNJ2, HSP60, HSP70, caspase-9, c-Met, Pax7, Pax3, IGF-1R, KLF4	C2C12, skeletal muscle, cardiac muscle, smooth muscle
LOC525506, LOC515772	miR-125a/b	IGF-II, Cbx7, SP7	skeletal muscle, smooth muscle, ESC
LOC112445239, LOC112441718, LOC112446120	miR-126	Spred-1, VCAM-1, IRS-1	cardiac muscle, ESC
LOC525506, LOC112449623	miR-128a	PPARγ, Runx1, Pax3	skeletal muscle
LOC112442443, LOC514189, LOC112448623, LOC100299745, LOC112443546, LOC112446047, LOC518869, LOC508468	miR-130a	GAX, HOXA5	vascular endothelial cells (ECs)
LOC525506, LOC112441673	miR-144	IRS1	Type II diabetes mellitus
LOC112441718, LOC112448623	miR-148a	ROCK1	skeletal muscle
LOC540051, LOC526226, LOC112444791, LOC112441811, LOC112448967, LOC615901, LOC508180, LOC617406	miR-15a	DLK1	3T3-L1 preadipocytes
LOC540051, LOC112444791, LOC11244896, LOC615901, LOC508180, LOC617406	miR-16a/b	cyclin D1	skeletal muscle
LOC540051, LOC526226, LOC112446505, LOC514189, LOC112442443, LOC112444239, LOC788100	miR-181	Hox-A11, Cbx7	skeletal muscle, ESC
LOC525506, LOC112446505, LOC514189, LOC112446526	miR-199a	Hif-1α, Sirt1	cardiac muscle
LOC525506, LOC787554	miR-204	Runx2	smooth muscle
LOC525506, LOC112448970	miR-208b/49	Sox6, Purβ, Sp3, HP-1β	muscle
LOC540051, LOC514189, LOC11244607, LOC112441937	miR-214	Ezh2, N-Ras	C2C12, skeletal muscle, ESC
LOC112442455, LOC112446845	miR-221	p27, Mdm2	skeletal muscle, mesenchymal cells
LOC525506, LOC526226, LOC112442877, LOC112449155	miR-23a	MAFbx, atrogin-1, Myh1/2/4	MuRF1, skeletal muscle

**Note:** Green represents up-regulated LncRNAs, red represents down-regulated LncRNAs, black represents LncRNAs with no significant difference.

## 8. Luciferase reporter assay

We next looked for potential miRNA binding sites of LOC104975788 and found potential miR-133a and miR-103 binding sites. Online miRNA binding site prediction software (RNA22 and RNAhybird) both predicted potential interaction of miR-133a with LOC104975788 and LOC536229(Pax7). The results showed that both LOC104975788 and LOC536229 (Pax7) contained miR-133a binding sites (Fig. 7a). Luciferase reporter assay showed significantly lower luciferase activity of miR-133a co transfected with

the psi-LOC112448162 vector compared to the control group ( $P < 0.01$ ). The luciferase activity of miR-133a co-transfected with psi-LOC536229 (Pax7) vector was also significantly decreased relative to the control group ( $P < 0.01$ ) (Fig. 7b and 7C). These results indicate that both LOC104975788 and LOC536229 (Pax7) contain miR-133a binding sites.

## 9. Verification of sequencing results

Nine differentially expressed LncRNAs and five differentially expressed mRNAs were verified by qRT-PCR. The qRT-PCR results (Fig. 7d and 7e) confirmed differential expression of these LncRNAs and mRNAs.

## Discussion

Muscle fiber is the basic unit of muscle, and the type of muscle fiber greatly shapes muscle characteristics<sup>14</sup>. There were obvious differences in appearance in the muscle fibres isolated from longissimus dorsi muscles of Shandong black cattle and Luxi cattle. The average muscle fiber area of Shandong black cattle was significantly larger than that of Luxi cattle ( $P < 0.05$ ), and the muscle fiber diameter was smaller than that of Luxi cattle. There was a significant positive correlation between muscle fiber area and carcass traits. Muscle fiber diameter is closely related to meat quality and taste and thicker muscle fiber diameter corresponds to decreased muscle tenderness. There is a negative correlation between muscle fiber diameter and muscle fiber density<sup>15</sup>. The amount of slow muscle fiber affects sarcomere length and has an important impact on meat quality<sup>16</sup>. There was a higher proportion of slow muscle fibers in longissimus dorsi muscle samples from Shandong black cattle than that of Luxi cattle. There were also significant differences in the muscle fiber density, the ratio of fast-twitch fibers/slow-twitch fibers, and weight ( $P < 0.05$ ) in this study. These differences may be key factors leading to the differences in meat production performance and meat quality of the two breeds of cattle after birth, which was also the research basis of this study to explore the underlying molecular regulatory mechanism.

With low conservation of LncRNA sequences among species, bioinformatics methods are required to screen and identify LncRNAs. This type of analysis is based on transcript length, the number of exons, and the coding potential. In this study, transcriptome data of skeletal muscle of Shandong black cattle and Luxi cattle were generated and analyzed. According to the screening criteria of  $|FC| \geq 2$  and  $P < 0.05$ , 1415 transcripts were found to be differentially expressed, with 939 and 476 transcripts were significantly up-regulated and down regulated in Luxi cattle, respectively. Further, 19 GO items (Table 2) and 14 regulatory pathways (Table 3) related to muscle development were screened by hierarchical GO and KEGG cluster analysis. Many genes involved in the regulation of muscle development were identified, including myosin protein family genes (MYH1, MYH3, MYH4, MYH8, and MYL3), myogenic regulatory factor family genes (MYO10 and MYO5B), Homeobox family genes (HOXC10, HOXC11, and HOXD9), Troponin T family genes (TNNT2 and TNNT3) and some regulatory transcription factors (WNT4 and TNNT2). It is worth noting that three LncRNAs (LOC107133268, LOC112445823, LOC104969184) are directly enriched into muscle regulation items. Further target gene prediction analysis shows that

LncRNAs may participate in the regulation of muscle development through bta-miR-2892 (LOC107133268), bta-miR-2360 and bta-miR-2449 (LOC112445823). The specific regulatory mechanism needs to be further verified and analyzed.

Compared with studies in human<sup>17</sup> and other model organisms<sup>18</sup>, there has been limited identification and characterization of beef LncRNAs, especially ones related to skeletal muscle development, with most studies of cattle focused on the identification of genes and miRNAs<sup>19,20</sup>. In this study, we identified 8427 multiple exon LncRNAs in beef skeletal muscle, and 480 differentially expressed LncRNAs were identified. More LncRNAs were detected in this study than previously reported in goats<sup>21,22</sup>. Fifteen randomly selected differentially expressed transcripts were verified by qPCR, and the results were consistent with the results of RNA sequencing. In conclusion, these results confirm the reliability of the identification of LncRNAs<sup>23</sup>.

Although LncRNAs can act on gene sites far from their chromosomal location<sup>24</sup>, genes in close proximity on a chromosome may participate in the same cellular metabolic pathways and have similar biological functions<sup>25</sup>. Therefore, the distribution of differentially expressed LncRNAs on chromosomes and the linkage differential expression of nearby genes may have biological significance that can help us to determine the function of a gene. The most differentially expressed LncRNAs were found on chromosome 9, followed by chromosomes 3, 1, 11, and 7. In co-expression analysis, we detected 387 differentially expressed LncRNA transcripts related to protein-coding genes according to the expression correlation coefficient values ( $r > 0.9$ ), and predicted 1164 target genes. GO and KEGG analysis showed significant enrichment of 130 GO items ( $P < 0.05$ ), with 20 of them related to the regulation of genes involved in muscle development. KEGG analysis showed enrichment of 1164 potential target genes in 299 pathways, with some related to muscle development (Table 6), such as the calcium signaling pathway, the AMPK signaling pathway, the cGMP - PKG signaling pathway, and the PPAR signaling pathway. Interestingly, we also found MYORG, DII1, EFNB2, SOX6, PROX1, MYOCD, NEBL and MYLK3 genes, annotated as related to muscle development (Table 5). Overall, LncRNAs may play a regulatory role in skeletal muscle biological development through cis or trans mechanisms.

At the post transcriptional level, LncRNA binds miRNA competitively with mRNA for protein-coding genes, thus relieving the inhibitory effect of miRNA on protein-coding genes to promote expression of these genes<sup>26</sup>. We analyzed the expression patterns of LncRNAs differentially expressed among different breeds, and constructed an LncRNA-miRNA-mRNA interaction network related to skeletal muscle development. A total of 52 LncRNAs related to muscle development were identified, with nine up-regulated and five down-regulated. All of these LncRNAs contain one or more putative miRNA binding sites with 48 LncRNAs predicted to interact with 1–2 miRNAs related to muscle development, and other LncRNAs predicted to interact with more than three miRNAs, such as LOC525506, LOC540051, LOC514189. With the miRNA seed sequences, LncRNAs can bind to miRNA to act as a sponge, preventing miRNA from binding to its target mRNA. As a classic example, M. Cesana<sup>27</sup> confirmed that linc-MD1, a long non-coding RNA specifically expressed during myoblast differentiation, regulates the expression of

muscle specific transcription regulators MAML1 and MEF2C through miR-133 and miR-135. In particular, LOC525506 may be involved in the development of beef skeletal muscle by regulating miR-1, miR-23a, miR-378, let-7, miR-483, and miR-21. This is the most predicted LncRNA, but little is known about its expression and function. Both LOC112447073 and LOC104975788 are involved in the skeletal muscle development of beef cattle by interaction with miR-103. More interestingly, some target genes of LOC112447073, LOC104975788, and LOC101903367 directly regulate the development of muscle fiber and maintain the stability and development of muscle, such as miRna-103, miR-133a, miR-145, MEF2C, myocardin, and Pax7<sup>28,29</sup>. Choi identified 11 LncRNAs in bovine transcripts by studying skeletal muscle and adipose tissue of Korean cattle, with four related to muscle function<sup>30</sup>. These data provide new insight into the role of LncRNA in muscle development and enrich the existing LncRNA mammalian skeletal muscle resources.

The second to eighth bases of the sequence at the 5' end of miRNA is called the seed sequence. Complete complementarity of this sequence and that of the target gene indicates the potential for binding of the target gene by miRNA<sup>31,32</sup>. We predicted the miRNA binding sites of functional LncRNAs that may be related to skeletal muscle development. Using a double luciferase test, we found a recognition site of bta-mir-133a in the sequence of LOC104975788 and a binding site for bta-mir-133a in the downstream target gene Pax7. We speculate that LOC104975788 may be involved in the regulation of skeletal muscle development by competitively inhibiting the expression of the target gene Pax7 through binding of bta-mir-133a. Future work should test this proposed regulatory mechanism.

## Conclusion

The expression patterns of LncRNAs in skeletal muscle of Shandong black cattle and Luxi cattle were elucidated by RNA sequence analysis, and the LncRNAs that may be involved in the skeletal muscle development of beef cattle were identified. The results allowed construction of interaction networks of LncRNAs-miRNA-mRNA regulated by muscle biology, which can provide insight into the regulatory mechanisms of LncRNAs involved in skeletal muscle development. Interactions among mRNA, miRNA, and LncRNA may regulate muscle development. This work provides a framework to explore the molecular mechanisms used by LncRNA to regulate important economic traits of muscle development, and should help to accelerate the livestock breeding process and improve animal production performance.

## Abbreviations

DEGs:differentially expressed genes; KEGG:Kyoto Encyclopedia of Genes and Genomes

GO:The Gene Ontology; LncRNA:Long noncoding RNA; miRNA: MicroRNA (miRNA) is a class of noncoding single stranded RNA molecules with a length of about 22 nucleotides encoded by endogenous genes; mRNA: Messenger RNA(mRNA) that encodes proteins.

## Declarations

## **Ethics approval and consent to participate**

All experimental design and procedures of this study was performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Committee on the Ethics of Animal Experiments of Qingdao Agricultural University IACUC (Institutional Animal Care and Use Committee) .

## **Consent for publication**

Not applicable.

## **Availability of data and materials**

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## **Data availability**

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. We have stored the RNA sequence data in the public domain GEO NCBI to obtain the GEO accession numbers: GSM4904154, GSM4904155, GSM4904156, GSM4904157, GSM4904158, GSM4904159.

## **Competing interests**

The authors declare that they have no competing interests.

## **Funding**

This work was supported by Shandong modern agricultural industry technology system cattle industry innovation team (668-2216030). The team participated in the collection and analysis of data.

## **Authors' contributions**

YJD, XJB, CZX and RLL designed the study. RLL,XXL and KY conducted the experiment. RLL, XXL and KY performed and collected the data. RLL analyzed the data and wrote the manuscript. All authors read and approved the final manuscript.

## **Acknowledgements**

The authors would like to thank Shanghai Sangon Biotechnology Co., Ltd. and Beijing AnnoroadGene Co., Ltd. for their assistance in the original data processing and related bioinformatics analysis for RNA-sequencing and GC-MS. The authors also thank Dr. Liu for her meticulous revision of language. More importance, special thanks to Yue Liu for his contribution to data processing. Additionally, the authors would like to particularly thank her parents and Dr. Dong , for their continuing love and support.

## Author details

<sup>1</sup>Laboratory of Animal Physiology and Biochemistry, Animal embryo Center, College of Animal Science, Qingdao Agricultural University, Qingdao, Shandong, People's Republic of China200109. <sup>2</sup>Laboratory of Animal molecular , Shandong black cattle breeding Engineering Technology Center, College of Animal Science, Qingdao Agricultural University, Qingdao, Shandong, People's Republic of China200109.

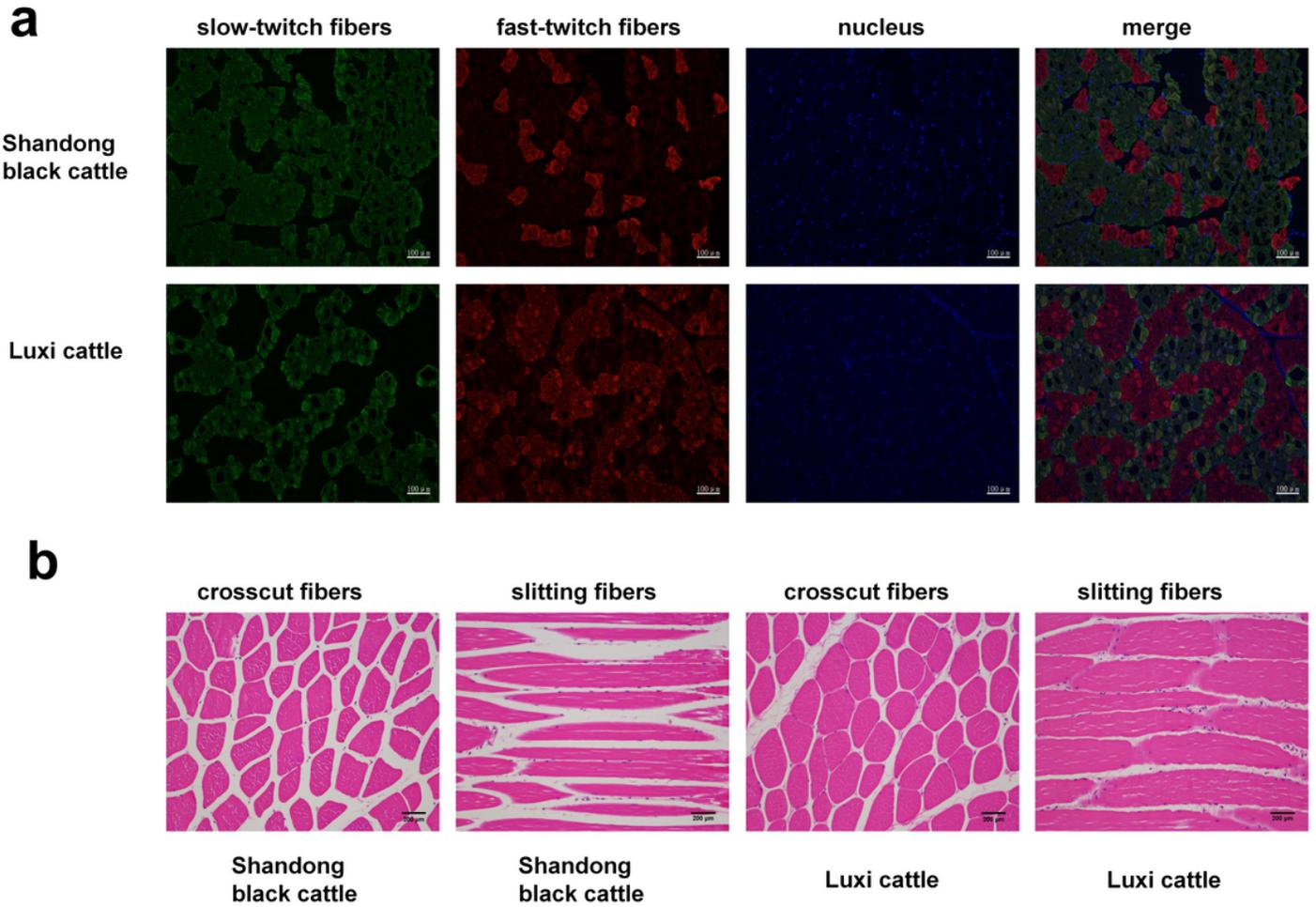
## References

- [1] Li Y , Chen X , Sun H , et al. Long non-coding RNAs in the regulation of skeletal myogenesis and muscle diseases. *Cancer Letters*. 417(2017)58-64.
- [2] Sun, X, Li,et al. The developmental transcriptome sequencing of bovine skeletal muscle reveals a long noncoding RNA, lncMD, promotes muscle differentiation by sponging miR-125b. *Biochimica Et Biophysica Acta*. (2016).
- [3] Yingwei, Yue, Congfei, et al. A lncRNA promotes myoblast proliferation by up-regulating GH1.. In vitro cellular & developmental biology. *Animal*. (2017).
- [4] Billerey C , Boussaha M , Diane Esquerré, et al. Identification of large intergenic non-coding RNAs in bovine muscle using next-generation transcriptomic sequencing. *BMC Genomics*. 15(2014)499.
- [5] Lu W Q. Feeding and management techniques of Luxi cattle in fattening period. *Beijing Agriculture*. 25(2015)130-131 .
- [6] Liu R L , et al. Screening of Skeletal Muscle Differential Genes Based on Transcriptome. *North China Agricultural Journal*. 33 (2018)64-72.
- [7] Cao, T, Shi, L.G, Zhang, L.L,et al. Comparative Study on Fetal Muscle Fiber of Wuzhishan Pig and Changbai Pig during 65d Gestation. *Journal of Animal Ecology*. 35(2014)37-40.
- [8] Zhao J, Ohsumi TK, Kung JT, Ogawa Y, Grau DJ, Sarma K, Song JJ, Kingston RE,et al. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol Cell*. 40(2010)939-53.
- [9] Wang L , Feng Z , Wang X , et al. DEGseq: an R package for identifying differentially expressed genes from RNA SEQ data . *Bioinformatics*. 26 (2010)136-138.
- [10] Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 11(2010)R106.
- [11] Cabili M N , Trapnell C , Goff L , et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & Development*. 25( 2011)1915-1927.

- [12] Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11(2010)R14.
- [13] Li D, Bao P, Yin Z, Sun L, Feng J, He Z, Jin M, et al. Exploration of the involvement of LncRNA in HIV-associated encephalitis using bioinformatics. *PeerJ.* 6(2018)e5721.
- [14] Steinbacher P, Haslett J R, Sanger A M, et al. Evolution of myogenesis in fish: a sturgeon view of the mechanisms of muscle development . *Anat Embryol (Berl).* 211(2006)311-322.
- [15] Zhou Zhijia. Relationship between muscle histological characteristics and meat quality . *Breeding technology consultant.* 8( 2010) 57-58.
- [16] Hendricks H B. Relation of porcine muscle fiber type and size to postmortem shortening. *Anim Sci.* 32(1971)57.
- [17] Agliano F, Rathinam V.A, Medvedev A.E, Vanaja S.K, Vella A.T. Long Noncoding RNAs in Host-Pathogen Interactions. *Trends Immunol.* (2019).
- [18] Briggs J.A, Wolvetang E.J, Mattick J.S, Rinn J.L, Barry G. Mechanisms of Long Non-coding RNAs in Mammalian Nervous System Development, Plasticity, Disease, and Evolution. *Neuron.* 88(2015)861-877.
- [19] Peters E.L, Van der Linde S.M, Vogel I.S.P, Haroon M, Offringa C, De Wit G.M.J, Koolwijk P, Van der Laarse W.J, Jaspers R.T. IGF-1 Attenuates Hypoxia-Induced Atrophy but Inhibits Myoglobin Expression in C2C12 Skeletal Muscle Myotubes. *Int. J. Mol. Sci.* 18(2017)1889.
- [20] Xue Z, Huang K, Cai C, Cai L, Jiang C.Y, Feng Y, Liu Z, Zeng Q, Cheng L, Sun Y.E, et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature.* 500(2013)593-597.
- [21] Sun M, Gadad S.S, Kim D.S, Kraus W.L. Discovery, Annotation, and Functional Analysis of Long Noncoding RNAs Controlling Cell-Cycle Gene Expression and Proliferation in Breast Cancer Cells. *Mol. Cell.* 59(2015)698-711.
- [22] Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, Liu Y, Chen R, Zhao Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 41(2013)e166.
- [23] Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: The Dynamic Tree Cut package for R. *Bioinform.* 24(2008)719-720.
- [24] Cohen B A , Mitra R D , Hughes J D , et al. A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nature Genetics.*26(2000), 183-186.

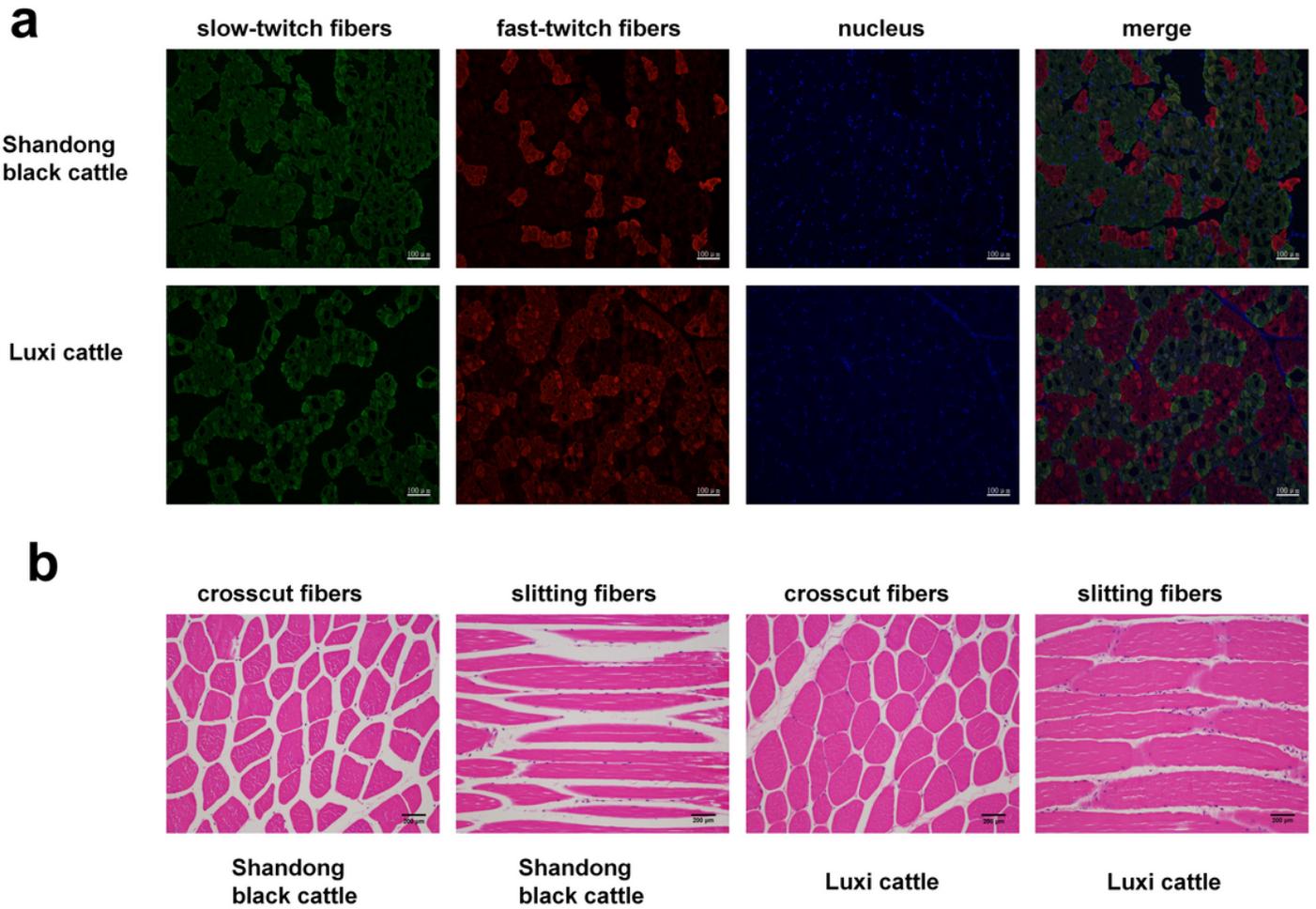
- [25] Salmena L, Poliseno L, Tay Y, et al. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language?. *Cell*. 146(2011) 353-358.
- [26] Kartha R V, Subramanian S. Competing endogenous RNAs (ceRNAs): new entrants to the intricacies of gene regulation. *Frontiers in genetics*. 5 (2014) 8.
- [27] Cesana M, Cacchiarelli D, Legnini I, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*. 147(2011) 358-369.
- [28] Mathew S.J., Hansen J.M., Merrell A.J., Murphy M.M., Lawson J.A., Hutcheson D.A., Hansen M.S., Angus-Hill M., Kardon G. Connective tissue fibroblasts and Tcf4 regulate myogenesis. *Dev. (Camb. Engl.)*. 138(2011)371-384.
- [29] Bennett R.D., Caride A.J., Mauer A.S. Interaction with the IQ3 motif of myosin-10 is required for calmodulin-like protein-dependent filopodial extension. *Febs Lett*. 582( 2008)2377-2381.
- [30] Choi J Y , Shin D , Lee H J , et al. Comparison of long noncoding RNA between muscles and adipose tissues in Hanwoo beef cattle. *Animal cells and systems the official publication of the Zoological Society of Korea*. 23(2018)1-9.
- [31] Lewis B P, Burge C B, Bartel D P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are micro RNA targets. *Cell*. 120(2005)15-20.
- [32] Grimson A, Farh K K, Johnston W K, et al. Micro RNA targeting specificity in mammals: determinants beyond seed pairing. *Molecular Cell*. 27(2007) 91-105.

## Figures



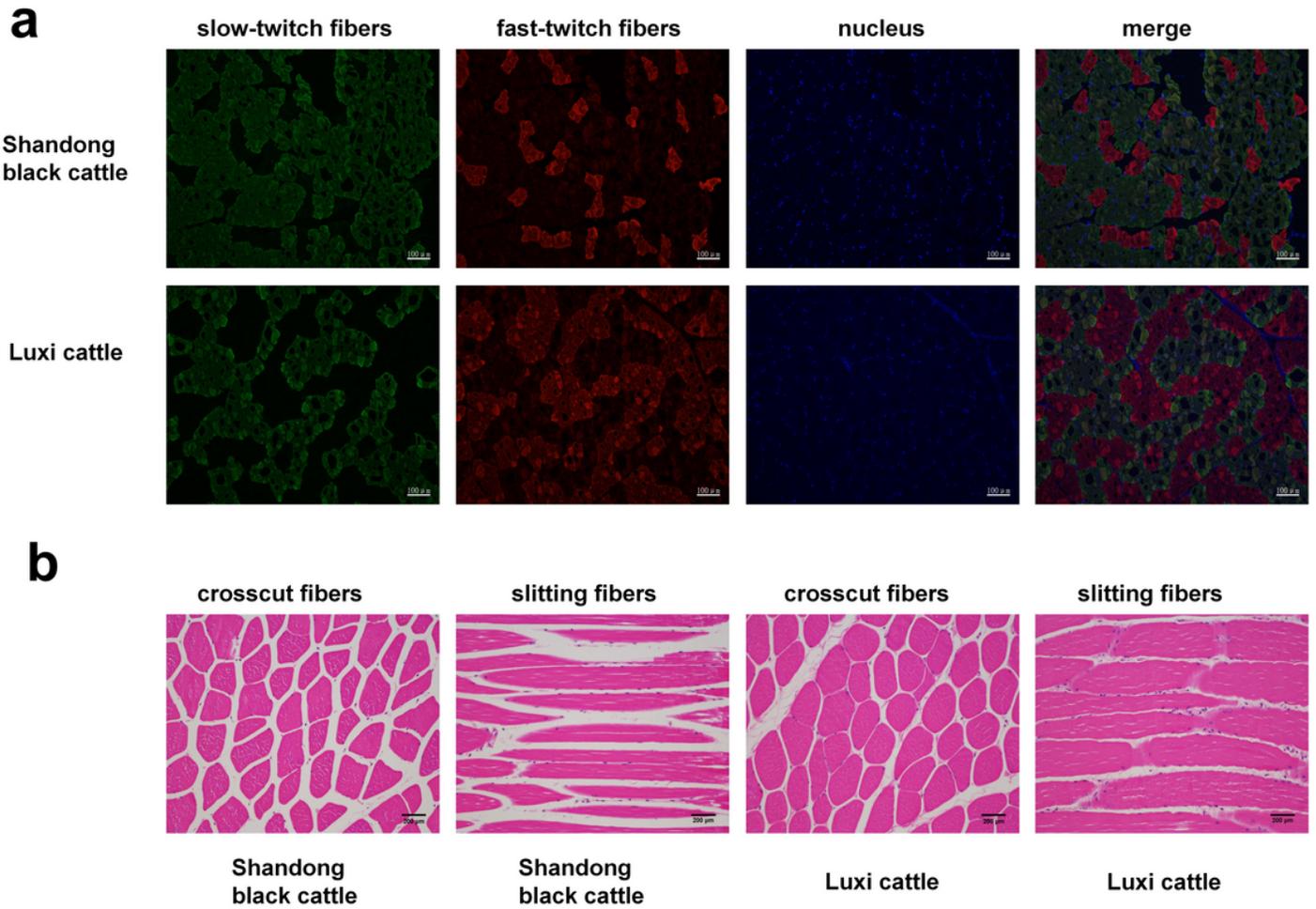
**Figure 1**

Fast/slow muscle fluorescence staining and paraffin section HE staining of muscle tissue. a Green represents slow muscle fiber, red represents fast muscle fiber, and blue represents nucleus of muscle cell. b Red represents a single muscle fibre, Blue represents a nucleus.



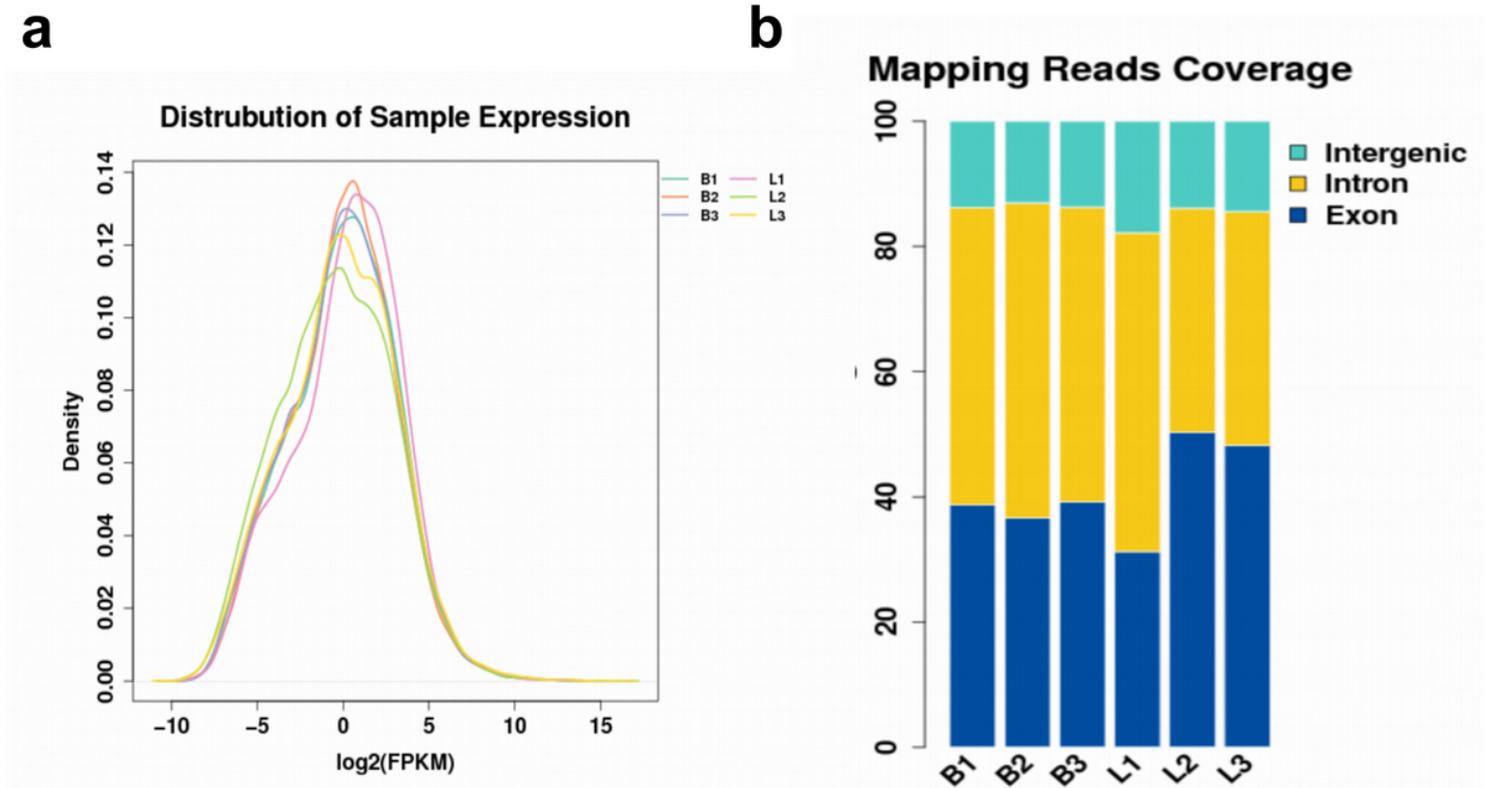
**Figure 1**

Fast/slow muscle fluorescence staining and paraffin section HE staining of muscle tissue. a Green represents slow muscle fiber, red represents fast muscle fiber, and blue represents nucleus of muscle cell. b Red represents a single muscle fibre, Blue represents a nucleus.



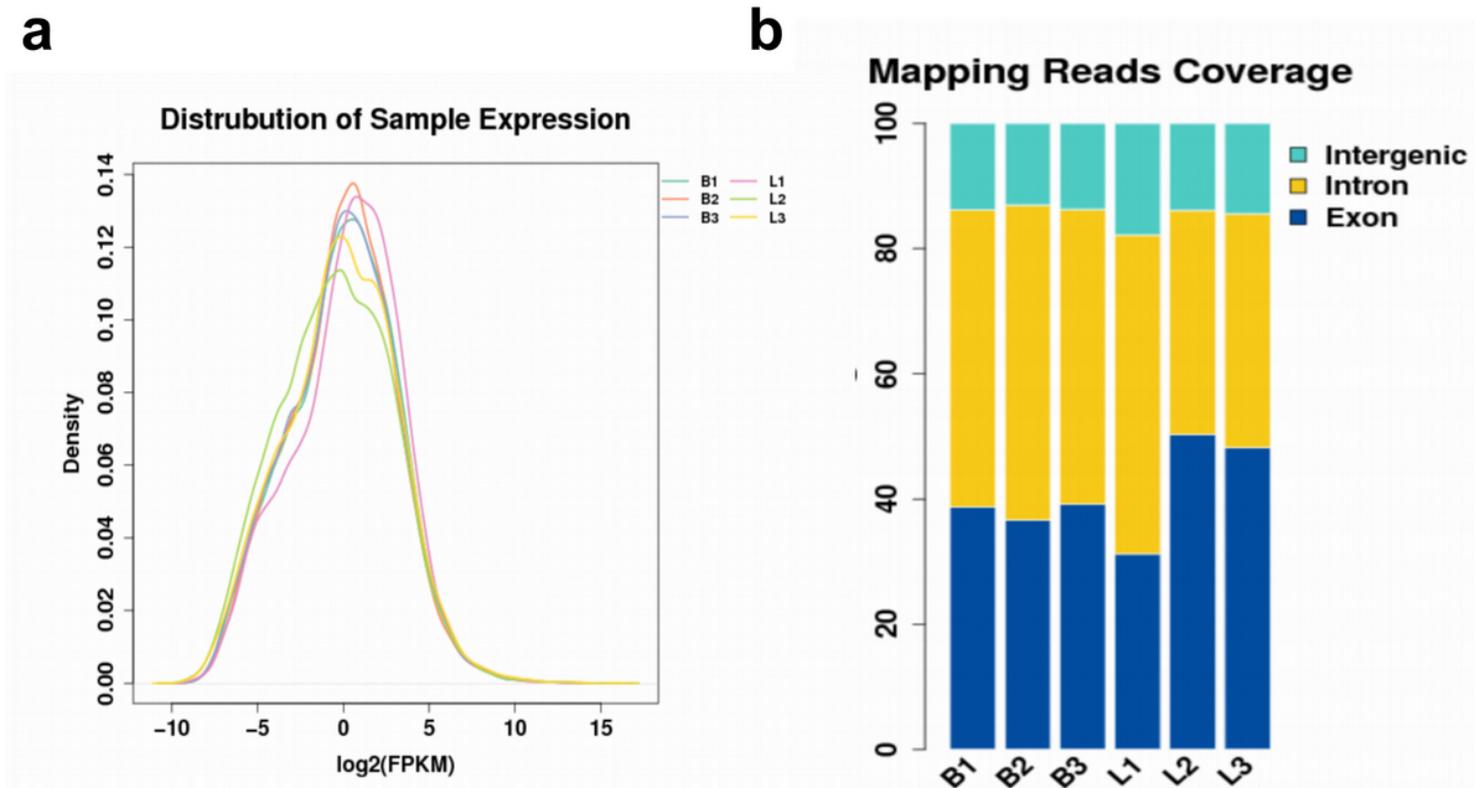
**Figure 1**

Fast/slow muscle fluorescence staining and paraffin section HE staining of muscle tissue. a Green represents slow muscle fiber, red represents fast muscle fiber, and blue represents nucleus of muscle cell. b Red represents a single muscle fibre, Blue represents a nucleus.



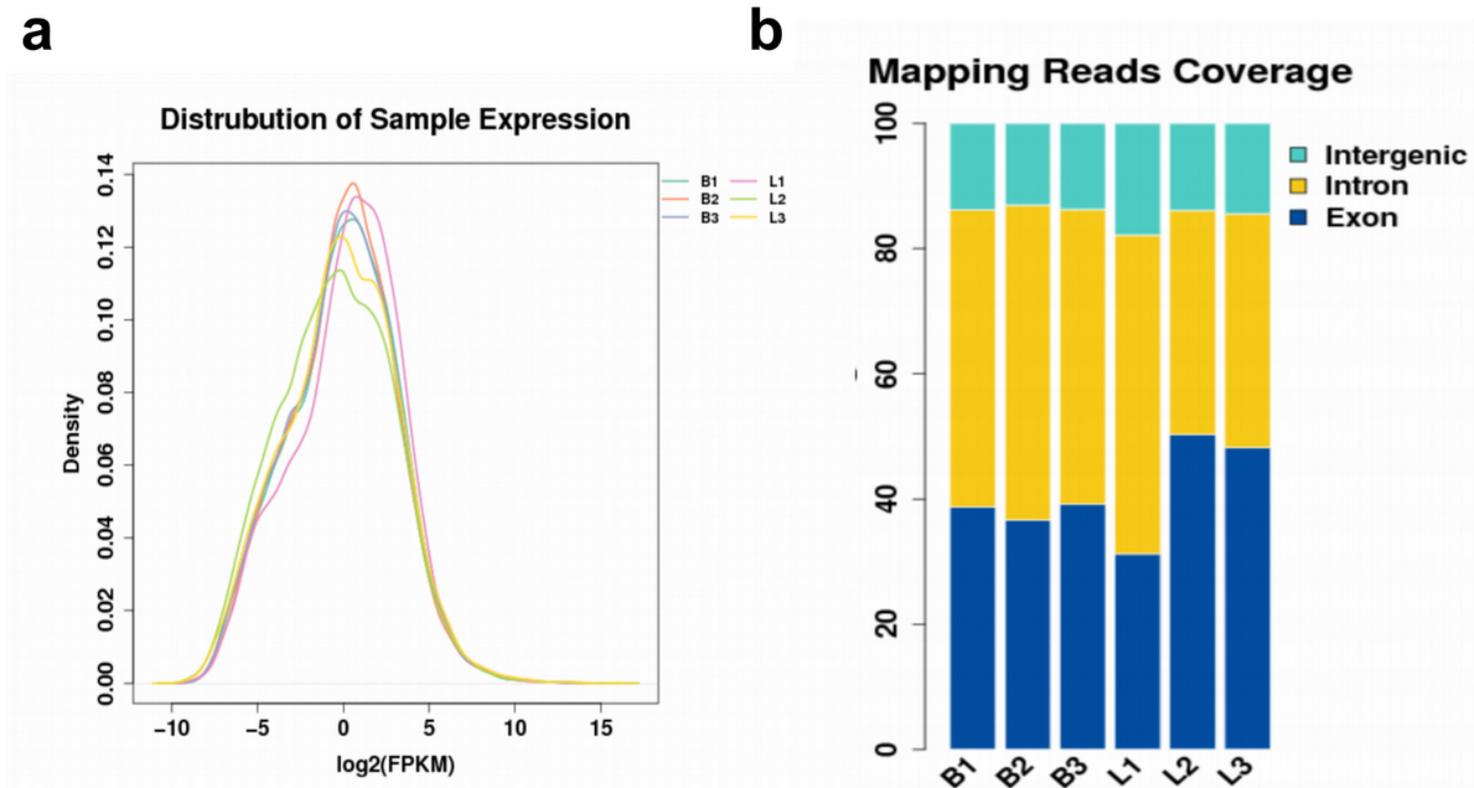
**Figure 2**

The relative expression and the genome distribution of novel transcripts and clean reads. a The new transcripts were relatively expressed in longissimus dorsi muscle of Shandong black cattle and Luxi cattle, and the density distribution map was made after taking logarithm of 2 as the base. The abscissa was  $\log_2(\text{FPKM} + 0.0001)$ , and the ordinate was gene density. Different colors represent different samples. b The results showed that the genomic distribution of clean data of longissimus dorsi muscle tissue of Shandong black cattle and Luxi cattle was genome distribution; the abscissa was each sample, and the ordinate was the ratio of the sequence number of exon, intron and intergenic region between the reference genome.



**Figure 2**

The relative expression and the genome distribution of novel transcripts and clean reads. a The new transcripts were relatively expressed in longissimus dorsi muscle of Shandong black cattle and Luxi cattle, and the density distribution map was made after taking logarithm of 2 as the base. The abscissa was  $\log_2(\text{FPKM} + 0.0001)$ , and the ordinate was gene density. Different colors represent different samples. b The results showed that the genomic distribution of clean data of longissimus dorsi muscle tissue of Shandong black cattle and Luxi cattle was genome distribution; the abscissa was each sample, and the ordinate was the ratio of the sequence number of exon, intron and intergenic region between the reference genome.



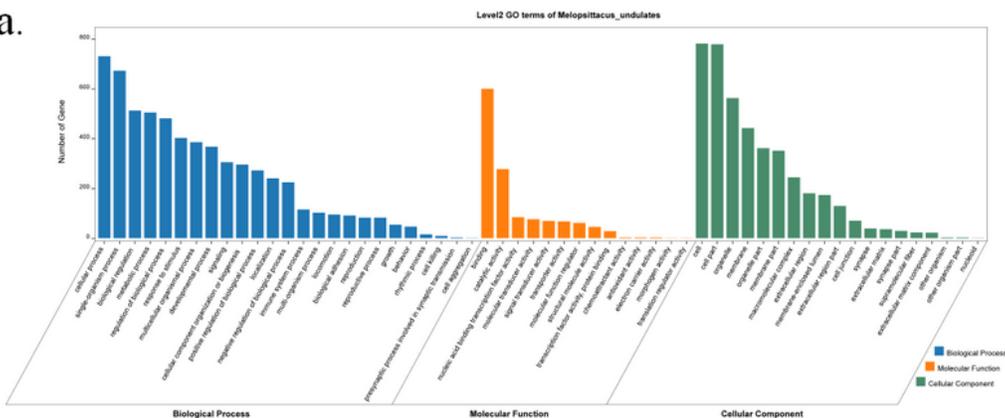
**Figure 2**

The relative expression and the genome distribution of novel transcripts and clean reads. a The new transcripts were relatively expressed in longissimus dorsi muscle of Shandong black cattle and Luxi cattle, and the density distribution map was made after taking logarithm of 2 as the base. The abscissa was  $\log_2(\text{FPKM} + 0.0001)$ , and the ordinate was gene density. Different colors represent different samples. b The results showed that the genomic distribution of clean data of longissimus dorsi muscle tissue of Shandong black cattle and Luxi cattle was genome distribution; the abscissa was each sample, and the ordinate was the ratio of the sequence number of exon, intron and intergenic region between the reference genome.

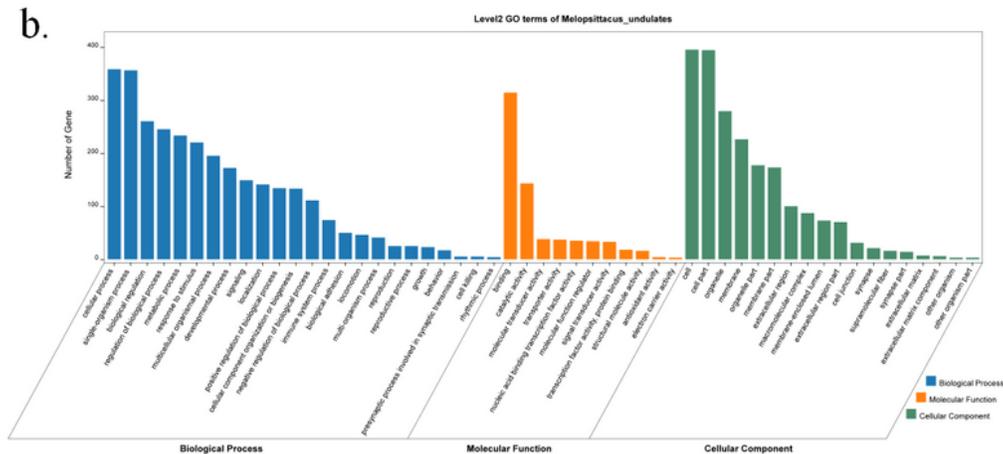


enrichment degree of the KEGG entry, and the closer the color to green, the higher the enrichment degree. The size of each dot indicates the number of genes enriched in the KEGG entry. The larger the dot, the more genes are enriched in the KEGG entry, and vice versa.

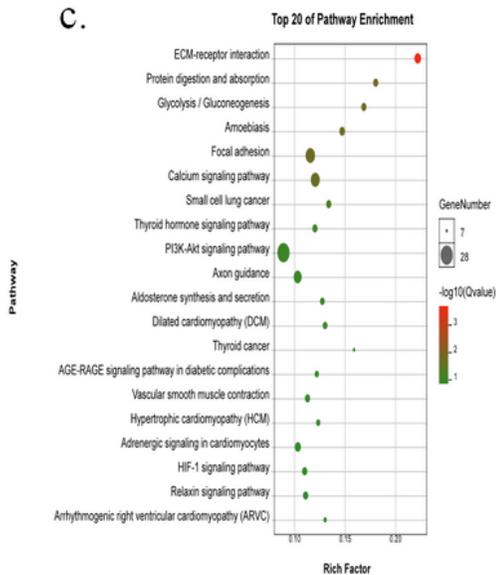
a.



b.



c.



d.

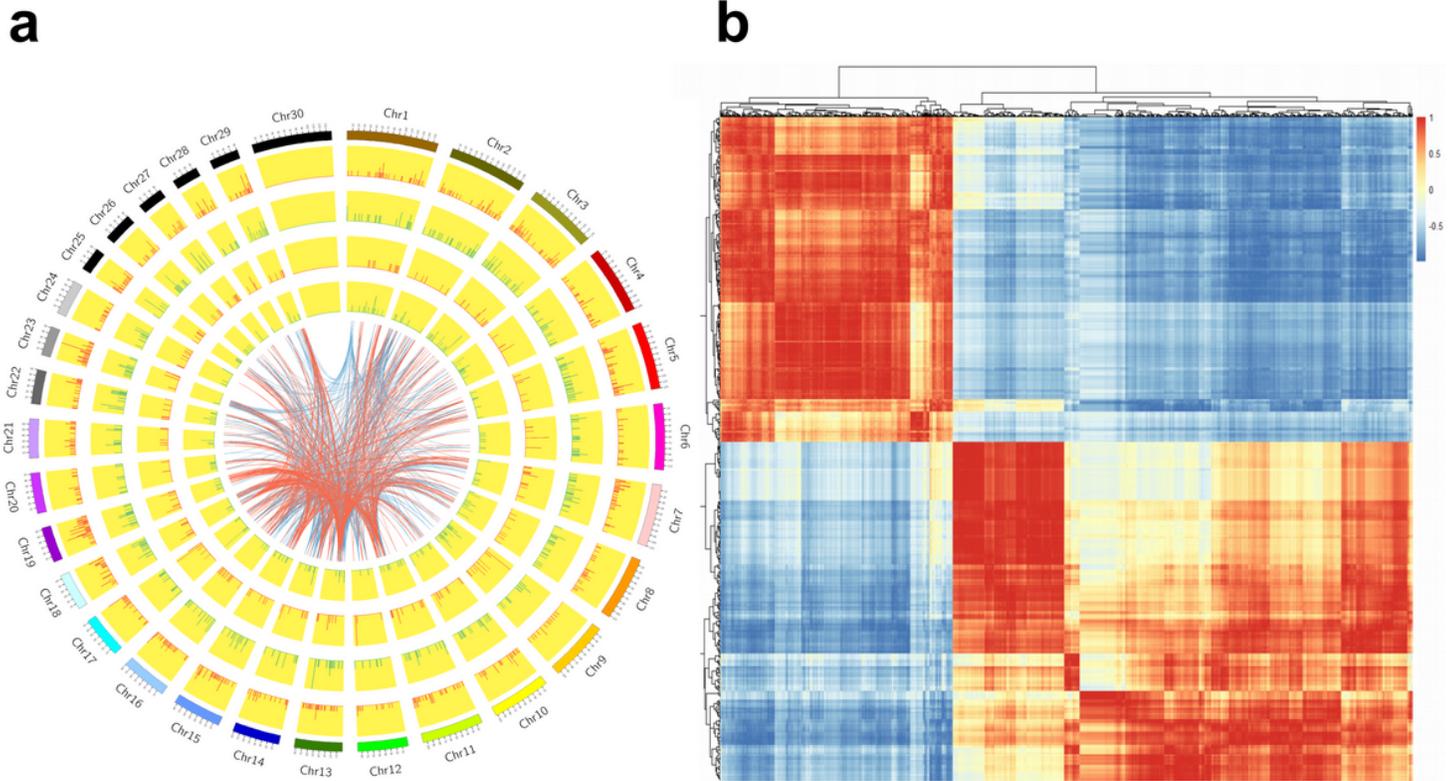


Figure 3

Functional analysis of differential genes. a and b The results of go cluster enrichment analysis for up-regulated and down regulated genes: the horizontal axis represents the name of go entry, and the vertical

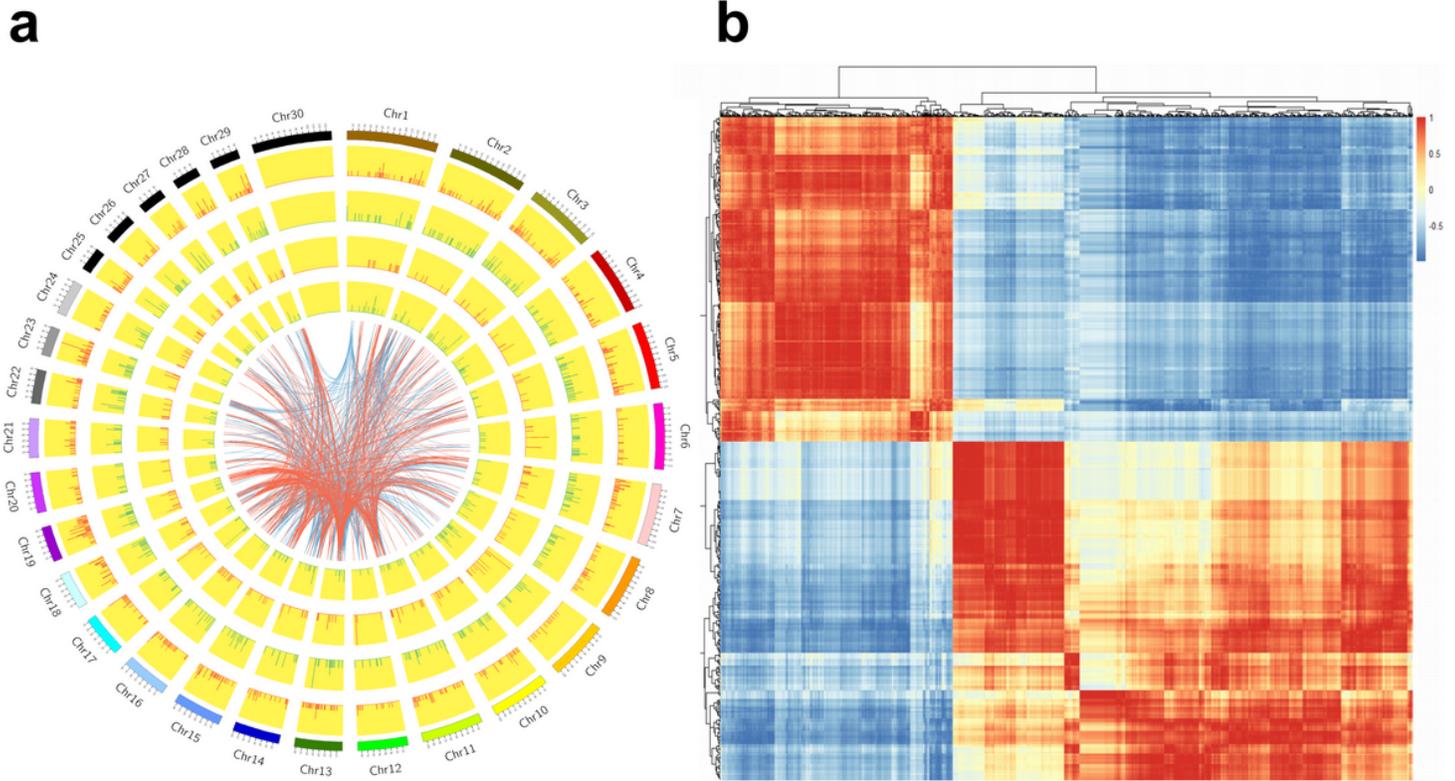


Functional analysis of differential genes. a and b The results of go cluster enrichment analysis for up-regulated and down regulated genes: the horizontal axis represents the name of go entry, and the vertical axis represents the number of genes enriched in the go entry. c and d. Up-regulated genes and down regulated genes were analyzed by go cluster analysis. The horizontal axis represents the name of the sample group, and the vertical axis represents the KEGG biological pathway. Each point indicates the enrichment degree of the KEGG entry, and the closer the color to green, the higher the enrichment degree. The size of each dot indicates the number of genes enriched in the KEGG entry. The larger the dot, the more genes are enriched in the KEGG entry, and vice versa.



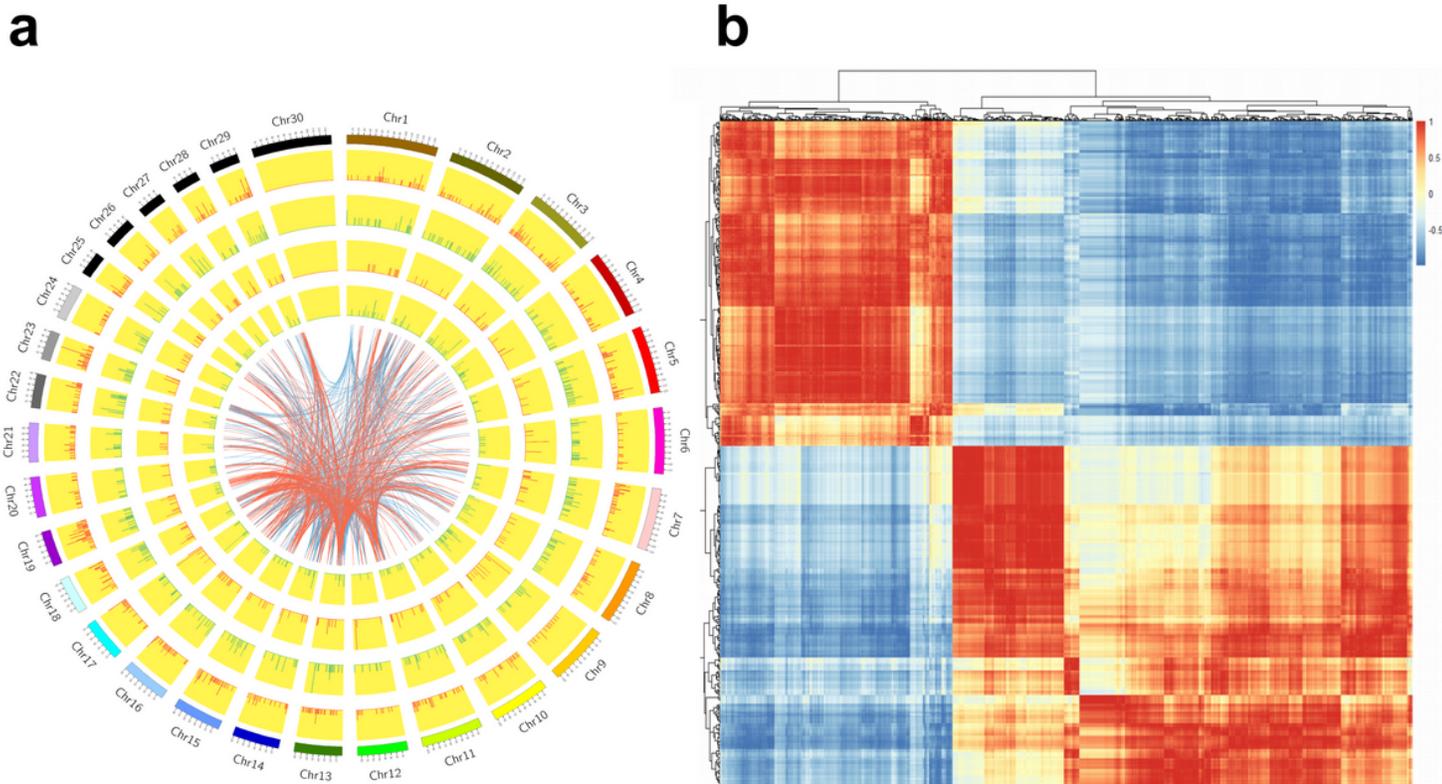
**Figure 4**

a. Circos map of differentially expressed genes and lncRNAs on cattle chromosomes. The outer ring is the diagram of autosomal distribution of the species; the first and second circles show the distribution of differentially expressed genes on chromosomes, with red lines indicating up-regulation, green lines indicating down-regulation, and the higher the column, the more differentially expressed genes in this region; the fourth and fifth circles show the distribution of lncRNAs on chromosomes, with the same expression pattern as genes. The internal lines indicate the corresponding relationship between lncRNAs and genes coexpressed by TOP 500. b. The Pearson correlation coefficient heat map of differential lncRNAs and mRNAs: Red represents positive correlation and blue represents negative correlation.



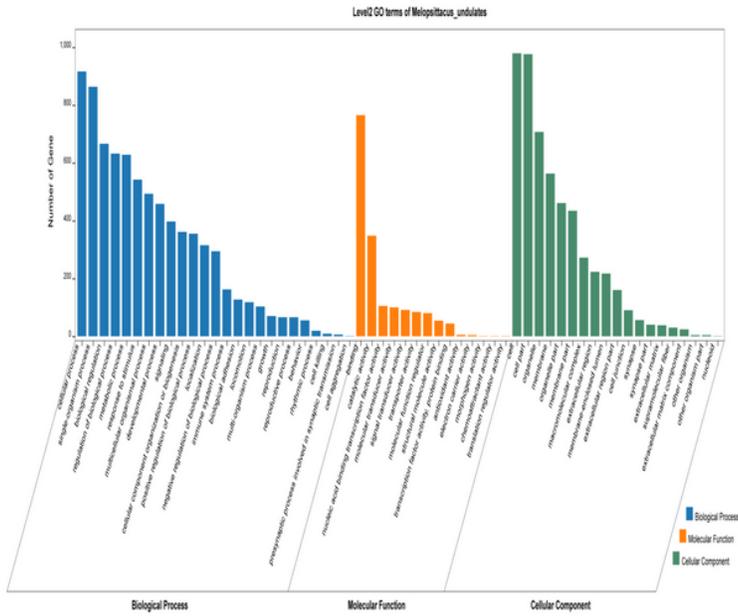
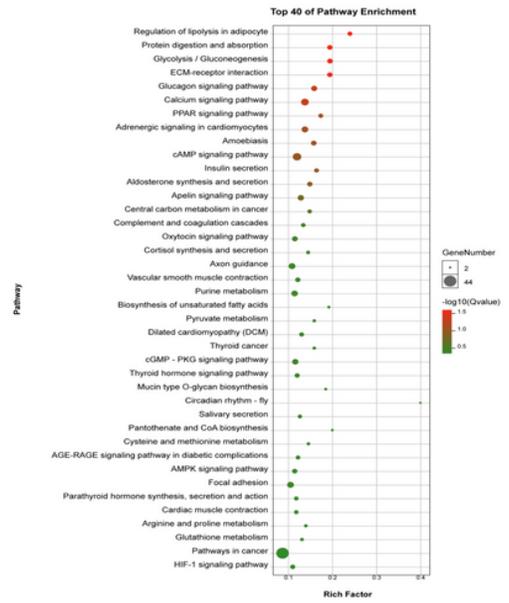
**Figure 4**

a. Circos map of differentially expressed genes and LncRNAs on cattle chromosomes. The outer ring is the diagram of autosomal distribution of the species; the first and second circles show the distribution of differentially expressed genes on chromosomes, with red lines indicating up-regulation, green lines indicating down-regulation, and the higher the column, the more differentially expressed genes in this region; the fourth and fifth circles show the distribution of LncRNAs on chromosomes, with the same expression pattern as genes. The internal lines indicate the corresponding relationship between LncRNAs and genes coexpressed by TOP 500. b. The Pearson correlation coefficient heat map of differential LncRNAs and mRNAs: Red represents positive correlation and blue represents negative correlation.

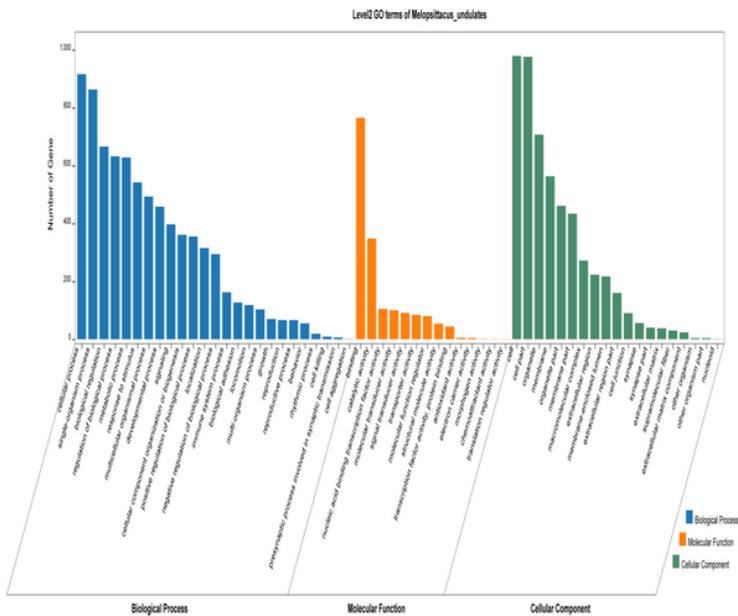
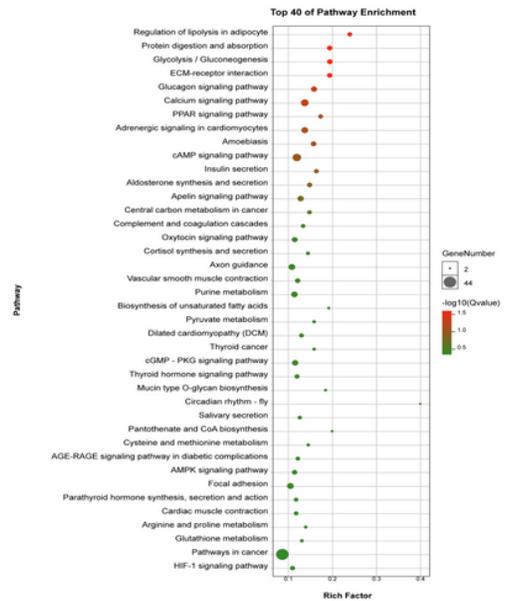


**Figure 4**

a. Circos map of differentially expressed genes and LncRNAs on cattle chromosomes. The outer ring is the diagram of autosomal distribution of the species; the first and second circles show the distribution of differentially expressed genes on chromosomes, with red lines indicating up-regulation, green lines indicating down-regulation, and the higher the column, the more differentially expressed genes in this region; the fourth and fifth circles show the distribution of LncRNAs on chromosomes, with the same expression pattern as genes. The internal lines indicate the corresponding relationship between LncRNAs and genes coexpressed by TOP 500. b. The Pearson correlation coefficient heat map of differential LncRNAs and mRNAs: Red represents positive correlation and blue represents negative correlation.

**a****b****Figure 5**

Analysis results of potential target genes GO (a) and KEGG (b) of LncRNA differentially expressed in beef cattle muscle growth and development. The legend is the same as Fig 3.

**a****b****Figure 5**

Analysis results of potential target genes GO (a) and KEGG (b) of LncRNA differentially expressed in beef cattle muscle growth and development. The legend is the same as Fig 3.

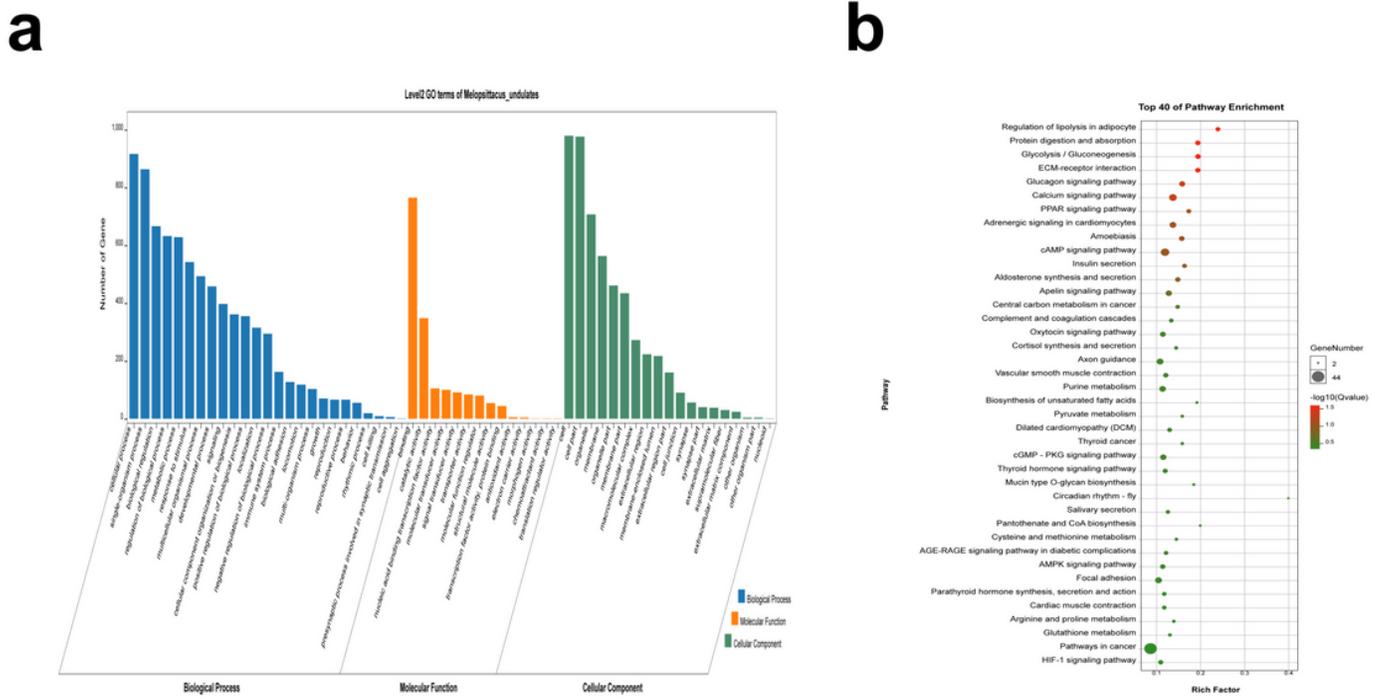


Figure 5

Analysis results of potential target genes GO (a) and KEGG (b) of LncRNA differentially expressed in beef cattle muscle growth and development. The legend is the same as Fig 3.

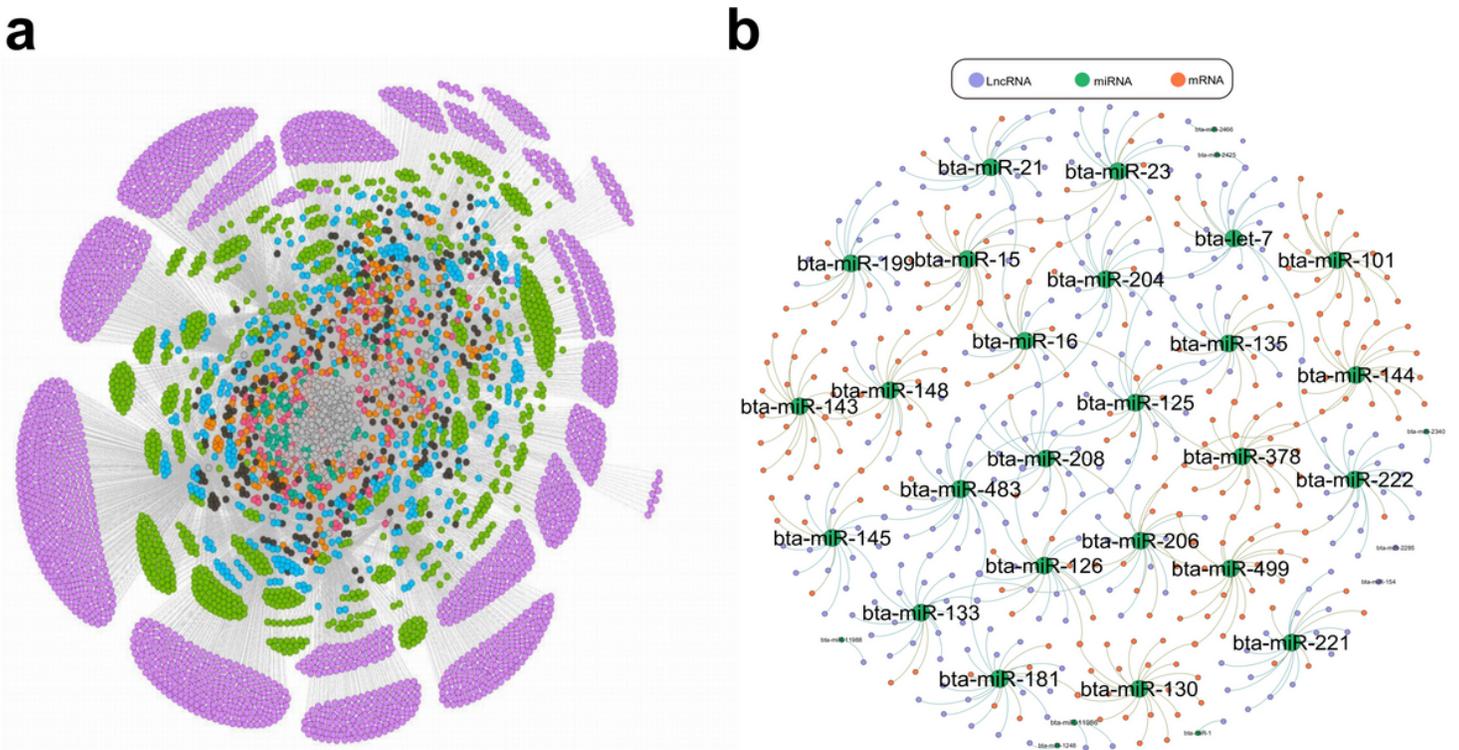


Figure 6

LncRNA-miRNA-mRNA network interaction diagram. a visualization of LncRNA-miRNA-mRNA interaction b LncRNA-miRNA-mRNA interaction network diagram related to muscle, green represents miRNA, orange represents mRNA, purple represents LncRNA.

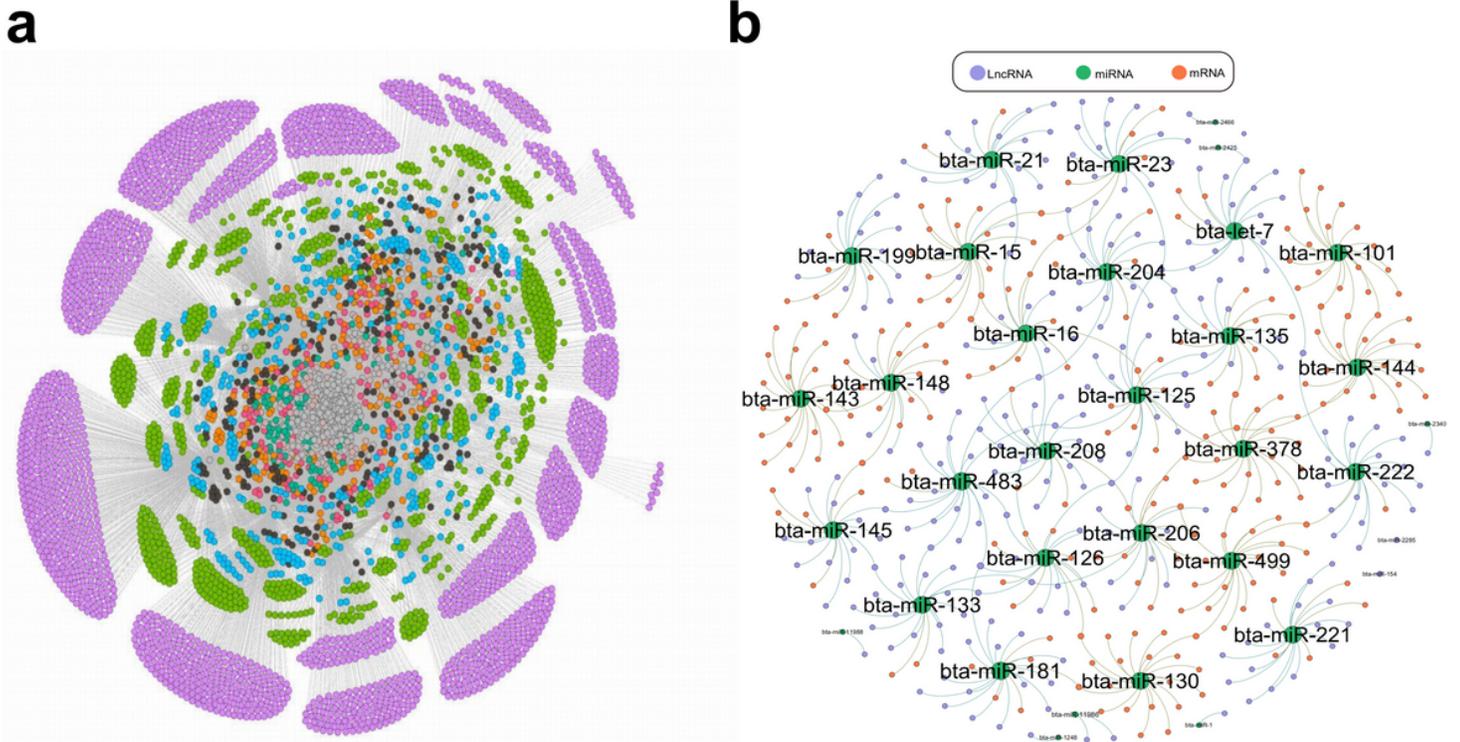
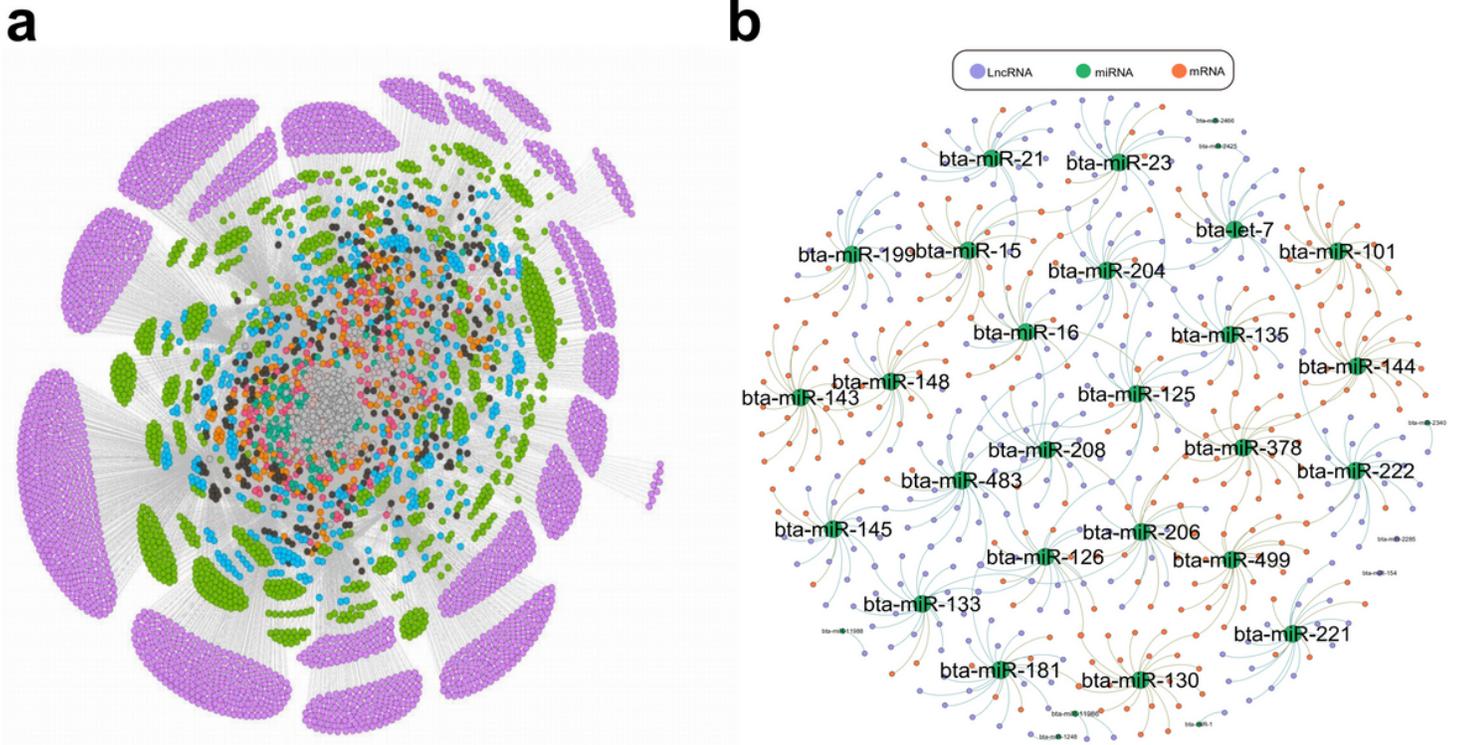


Figure 6

LncRNA-miRNA-mRNA network interaction diagram. a visualization of LncRNA-miRNA-mRNA interaction b LncRNA-miRNA-mRNA interaction network diagram related to muscle, green represents miRNA, orange represents mRNA, purple represents LncRNA.



**Figure 6**

LncRNA-miRNA-mRNA network interaction diagram. a visualization of LncRNA-miRNA-mRNA interaction b LncRNA-miRNA-mRNA interaction network diagram related to muscle, green represents miRNA, orange represents mRNA, purple represents LncRNA.







- [TableS1PrimersusedforqRTPCR.xls](#)
- [TableS1PrimersusedforqRTPCR.xls](#)
- [TableS1PrimersusedforqRTPCR.xls](#)
- [TableS2DifferentiallyexpressedmRNAs.xls](#)
- [TableS2DifferentiallyexpressedmRNAs.xls](#)
- [TableS2DifferentiallyexpressedmRNAs.xls](#)
- [TableS3.xls](#)
- [TableS3.xls](#)
- [TableS3.xls](#)
- [TableS4DifferentiallyenrichedpathwayinKEGG.xls](#)
- [TableS4DifferentiallyenrichedpathwayinKEGG.xls](#)
- [TableS4DifferentiallyenrichedpathwayinKEGG.xls](#)
- [TableS5DifferentiallyexpressedlncRNAs.xls](#)
- [TableS5DifferentiallyexpressedlncRNAs.xls](#)
- [TableS5DifferentiallyexpressedlncRNAs.xls](#)
- [TableS6Thetop40significantlydiffexpressionofLncRNAs.xls](#)
- [TableS6Thetop40significantlydiffexpressionofLncRNAs.xls](#)
- [TableS6Thetop40significantlydiffexpressionofLncRNAs.xls](#)
- [TableS7.xls](#)
- [TableS7.xls](#)
- [TableS7.xls](#)
- [TableS9.xls](#)
- [TableS9.xls](#)
- [TableS9.xls](#)