Transcriptome Analysis Reveals that Long Noncoding RNAs Contribute to Development Differences of Medium-sized Ovarian Follicle between Meishan and Duroc Sows

Mengxun Li
Shihezi University https://orcid.org/0000-0002-7477-969X
Yi Liu
Shihezi University
Lipeng Ma
Shihezi University
Zhichao Zhao
Shihezi University
Su Xie
Shihezi University
Hongbin Gong
Shihezi University
Yishan Sun
Shihezi University
Tao Huang (2009tao@shzu.edu.cn)
College of Animal Science and Technology, Shihezi University https://orcid.org/0000-0003-2943-5131

Research article

Keywords: Duroc sows, Meishan sows, lncRNAs, Ovarian follicle, reproduction loci, PI3K-Akt

Posted Date: October 5th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-53459/v1

Abstract

Background: Ovulation rate is an extremely important factor of litter size in sows. It differs greatly among pig breeds of different genetic backgrounds. Long non-coding RNAs (lncRNAs) can regulate follicle development, granulosa cell(GC) growth and hormone secretion, which in turn affects sow litter size.

Results: In our research, we identified 3554 lncRNAs and 25491 mRNAs in M2 follicle from Meishan and Duroc pigs. lncRNAs sequence and open reading frame(ORF) length is shorter than mRNAs, and it has fewer exons, lower abundance and conserved than protein-coding RNAs. Furthermore, 201 lncRNAs were differentially expressed in breeds, and quantitative trait loci (QTL) analysis of differential expression (DE) lncRNAs were performed. 127 DE lncRNAs are located in 119 reproduction trait-related loci. In addition, the lncRNAs potential target genes (PTGs) in cis or trans were predicted. Gene ontology(GO) and KEGG pathway analysis revealed that some PTGs were include some follicular development and hormone secretion-related biological processes or pathways, such as regulation of progesterone biosynthetic process, oestrogen metabolic process and ovarian steroidogenesis and PI3K-Akt signalling pathway. Furthermore，We also screened 19 differentially expressed lncRNAs of PI3K-Akt signalling pathway as candidates.

Conclusions: This study provided a new significance on the roles of lncRNAs in follicular growth and development and porcine reproduction.

Background

Ovulation rate is an important limiting factor of litter size in pig reproductive performance[1]. Meishan sows ovulate 25–29 eggs per oestrus, whereas Duroc sows ovulate only 12.3[2]. The difference in ovulation rates causes the difference in litter size between the two breeds. Therefore, these two breeds can be compared as prolific and ordinary sows.

Ovarian follicles are the basic functional unit in oocyte development, and they are round, vesicular and located in the ovarian cortex. Many studies on the regulating mechanisms of animal follicle development focus on protein-coding RNA and microRNA(miRNA) in molecular genetic level. Lee et al. found that retinoblastoma (Rb) protein was closely related to the growth state of follicles; the expression of Rb protein was significantly reduced, and follicles underwent rapid atresia[3]. The growth differentiation factor-9 (GDF-9) gene promoted the follicles growth and differentiation, knocking GDF-9 out caused infertility[4]. Inhibition of miR-145 expression reduced the primordial follicles proportion or number and increased growing follicles numbers[5]. MiR-92a inhibits the apoptosis of follicular granulosa cells by binding to its target gene drosophila mothers against decapentaplegic 7 (Smad 7)[6]. However, studies on the effect of lncRNAs on follicle growth and development are still few.

In recent years, numerous lncRNAs have been identified, and some of it have important functions in different biological processes, such as regulation on gene expression [7], genomic imprinting[8] and individual growth regulation[9, 10]. However, lncRNAs identified in pig are relatively few compared with those identified in humans and mice[11, 12], and lncRNAs that affect follicle development in pig are yet to be elucidated.

In this research, RNA samples from medium follicles(M2) of Duroc and Meishan sows were used to perform in high-throughput RNA-Seq. Plenty of lncRNA and mRNA transcripts were identified，differentially expressed lncRNAs and mRNAs between the Meishan ans Drouc sows were detected. GO and KEGG analysis were carried in the potential target genes (PTGs) of lncRNAs. Some PTGs significantly take part in the follicle growth-related biological processes and signalling pathways. Combined with the results of DE genes analysis, we found that most of the DE lncRNAs can regulate the expression of their PTGs positively. This study gaves the way to future research to explore the functions of specific lncRNAs that may be involved in follicular growth and development.

Results

Overview of RNA sequencing

We analysed the RNA data of six libraries of M2 follicles from three Meishan and three Duroc pigs, and obtained 111,275,282 to 142,313,048 raw reads and 101,760,644 to 138,324,398 clean reads from per sample, respectively. The clean reads were used in the after discarding transcripts with adapters, the low quality reads or other possible contaminants. Tophat2 software was used to perform reference genomic alignment analysis on Clean reads. The results showed that the six sequencing libraries' mapping ratio were all greater than 70%, indicating that the sequencing accuracy is high and can be used for subsequent analysis (Table 1).

Table 1
The six library simples information.

Sample name	Raw reads	Clean reads	Clean bases	Mapped Reads	Mapping Ratio	Error rate(%)	Q20(%)	Q30(%)	GC content(%)
MFM2DY4_1	111656928	107238580	16.09G	80514662	75.08%	0.02	96.59	91.73	47.14
MFM2DY4_2	121457850	116315790	17.45G	88556504	76.13%	0.02	96.78	92.03	52.18
MFM2DY4_3	111275282	106746570	16.01G	82037293	76.85%	0.02	96.75	91.96	50.09
DFM2DY4_1	118575510	111359682	16.7G	92421276	82.99%	0.02	96.51	91.79	56.43
DFM2DY4_2	107494708	101760644	15.26G	82733818	81.3%	0.02	96.7	92.14	49.84
DFM2DY4_3	142313048	138324398	20.75G	115239405	83.31%	0.02	96.52	91.55	53.96

Identification and characterisation of lncRNA

All transcripts were screened in five steps. Finally 3,554 lncRNAs were obtained, including 1,997 known lncRNAs and 1,557 novel lncRNAs (Figures 1A and B). Among the lncRNAs, lncRNA and antisense-lncRNA accounted for 89.3% and 10.7%, respectively (Figure 1C). In addition, we obtained 25,491 mRNAs. The lengths of lncRNA sequences were shorter than those of mRNAs. The lncRNA transcripts average length was 1417 bp, while the mRNAs was 2423 bp (Figure 1D). lncRNAs had fewer exons. The exon number of lncRNAs ranged from 1 to 10, whereas that of mRNAs from 1 to 30 (Figure 1E). The lncRNA have shorter ORF than mRNA. The ORF lengths of lncRNAs mostly ranged from 1 bp to 500 bp, whereas those of the mRNAs were mostly from 1 bp to 2000 bp (Figure 1F). The expression abundance of lncRNA were lower than mRNA (Figure 1G). The lncRNA were less conserved compared with the protein coding genes, as revealed through a phastCon analysis (Figure 1H). The structural feature comparative analysis of selected lncRNAs confirmed the accuracy of selection.

DE Analysis of lncRNAs and mRNAs

We use cufflinks to normalize transcript expression to FPKM value, and performed a differential expression transcript analysis between the Meishan and Duroc medium follicle samples. A total of 201 differentially expressed lncRNAs (DELs) [p value<0.05 and $|\log_{10}(\text{foldchange})| > 1.3$] between the two breeds were detected, in which 117 DELs were downregulated, 84 were upregulated in Meishan follicles (Figure 2 A). Moreover, 675 DE protein-coding genes were detected, and 265 DEGs were upregulated and 410 were downregulated in Meishan follicles (Figure 2 B). DE lncRNAs and mRNAs were widely distributed on chromosomes, with larger numbers on chromosomes 1, 6 and 13 (Figures 2 A and B). To verify the accuracy of sequencing, 6 DE-lncRNAs (ENSSSCT00000018610, ALDBSSCT0000001721, ALDBSSCT0000000051, LNC_000116, ALDBSSCT0000011300 and ALDBSSCT0000006152) and 6 DE-mRNAs (COL3A1, LRP8, ENSSSCT00000009222, SEPP1, COL5A2 and CYP19A1) were selected randomly, and their relative expression levels in the DFM2DAY4 and MFM2DAY4 groups were detected using qPCR. The expression analysis of the six lncRNA and six mRNA are displayed in the Figure 3, which are consistent with the RNA-Seq analysis results both in lncRNA and mRNA (Figure 3).

QTL Analysis of DE lncRNAs

To explore the preliminary DELs function, we mapped the differential expressions(DE) of lncRNAs transcri[pts to the QTL database, and performed QTL analysis. The result showed that a total of 1446 QTLs that 145 DELs were located in, the 127 DE lncRNAs are located in 119 reproduction trait-related loci (Figure 4A). By studying the distribution of QTLs on the chromosome, 119 QTLs related to reproduction deposition were found to be distributed in chromosomes 1, 2, 3, 4, 7, 8, 13, 15 and 18 (Figure 4B). The 127 lncRNAs associated with reproduction QTLs could affect corpus luteum number (53), litter size (18), androstenone (23), total number born alive (4), number of stillborn (3), FSH concentration (3) and number of viable embryos (2) (Figure 4C).

Prediction of lncRNA potential target genes

In order to further explore the regulatory functions of DE lncRNAs, we predicted their PTGs of lncRNAs in 2 ways-cis and trans[13, 14]. In our study, we found that lncRNA may regulate multiple coding genes. For PTGs regulated by lncRNAs in cis, we searched protein-coding transcripts located in 100 kb upstream or downstream of the DELs as its cis-regulatory target genes. We obtained a total of 320 co-localised

target genes of 118 DELs. In trans way analysis, we obtained 2930 PTGs of 127 DELs. We only showed the PTGs s of 24 DELs. The PTGs numbers for each DEL were varied. For example, the maximum number among the lncRNA is LNC_000179, and it had 77 PTGs, the second is LNC_000715 with 69 target genes. Some lncRNAs, such as LNC_000718 and LNC_000802, only had 2 target genes (Table 2).

Table 2
Differentially expressed lncRNAs (DELs) and their target genes (PTGs).

Up DELs	Numbers of PTGs			Down DELs	Number of PTGs		
	PTGs	Up Regulated	Down Regulated		PTGs	Up Regulated	Down Regulated
ALDBSSCT0000002624	4	4	0	LNC_000179	77	0	77
LNC_000811	15	13	2	LNC_000715	69	68	1
ALDBSSCT0000004325	21	18	3	LNC_001511	29	1	28
LNC_001167	48	48	0	LNC_000846	8	2	6
LNC_000252	36	35	1	LNC_001333	5	1	4
ALDBSSCT0000009309	26	26	0	LNC_000802	2	0	2
LNC_001490	26	10	16	ALDBSSCT0000010765	29	2	27
LNC_000718	2	2	0	ALDBSSCT0000010764	17	0	17
LNC_000460	20	19	1	LNC_000116	49	7	42
LNC_001307	17	10	7	ALDBSSCT0000001721	11	0	11

Function Enrichment Analysis for lncRNAs

To study the regulatory role of DE lncRNAs in Meishan and Duroc M2 follicles, we predicted the function of DE target genes by using GO and KEGG to speculate the lncRNA functions. The GO and KEGG analysis results revealed that PTGs participated in 1063 biological processes and 111 pathways, significantly. Many biological processes were involved in follicular development and ovulation, such as the regulation of progesterone biosynthetic, oestrogen metabolism, negative regulation of cell proliferation, ITGA3-ITGB1-THBS1 complex, cellular response to transforming growth factor beta stimulus, meiotic cell cycle and steroid catabolic process ($p < 0.05$) (Figure 5A;). The KEGG pathway analysis shown the PTGs were significantly involved in ovarian steroidogenesis, PI3K-Akt, MAPK, Wnt, BMP and TNF signalling pathway ($p < 0.05$) (Figure 5B).

Some PTGs that participated in oestrogen metabolic process and ovarian steroidogenesis signalling pathway are highlighted, such as CYP1A1, CYP19A1 and HSD3B1. A gene that participated in oestrogen metabolic process and ovarian steroidogenesis signalling pathway was CYP3A7. HSD17B8 participated in oestrogen metabolic process. ALOX5, LHCGR, IGF1, GNAS, CYP2J2 and CYP17A1 participated in ovarian steroidogenesis signalling pathway. In addition, one protein-coding transcripts may be regulated by multiple lncRNAs, such as DE-lncRNA ALDBSSCT0000001929, ALDBSSCT0000006256 and ALDBSSCT0000002430, which were correlated with their target gene CYP1A1 significantly ($p<0.05$). They were all downregulated in Meishan compared with Duroc sows. LNC_000644 and ALDBSSCT0000006057 were correlated with CYP3A7 significantly ($p<0.05$) and were unregulated in Meishan. ALDBSSCT0000000051 was correlated with its target gene HSD17B8 significantly ($p<0.05$). ALDBSSCT0000000051 was downregulated, whereas HSD17B8 was unregulated (Figure 5C). Therefore, we speculated that some DE-lncRNAs participate in the development of porcine follicles by positively or negatively regulating their target genes, which are related to hormone secretion and metabolism.

Screening of potential reproduction-related lncRNAs in PI3K-AKT signalling pathway in pig ovarian follicle

The PI3K-AKT signalling pathway is very important in porcine follicular development. In order to explore the regulatory roles of the DE-lncRNAs involved in ovarian follicle growth and the relationship between these lncRNAs and PI3K-AKT signalling pathway, we analysed the expression of lncRNAs and their PTGs in PI3K-AKT signalling pathway. We found 12 protein coding genes, namely, THBS1, ITGA3, ITGA6, ITGB1, ITGB4, ITGB7, PIK3C2B, AKT2, CREB3L3, IKBKG and TP53 in the PI3K-AKT signalling pathway that were regulated by 19 DE-

lncRNAs. Some DE-lncRNAs and their PTGs showed significant positive correlation ($p<0.05$), as follows: extracellular matrix (ECM) THBS1 and lncRNA-ALDBSSCT0000004244; ITGA3 and the novel lncRNAs LNC_000857 and LNC_001052; ITGA6 and ALDBSSCT0000010443; ITGB4 and LNC_000536; ITGB7 and ALDBSSCT0000004232; PIK3CA and ENSSSCT00000036444; AKT2 and LNC_000751, LNC_000819 and LNC_000058; CREB3L3 and ALDBSSCT0000008900 and ALDBSSCT000001184; IKBKG and ALDBSSCT0000004325 and LNC_001172; and TP53 and ALDBSSCT0000002268. Some DE-lncRNAs and their PTGS shown significant negative correlation ($p<0.05$), as follows: ITGA6 and ALDBSSCT0000009356; ITGB1 and LNC_001556; PIK3C2B and ALDBSSCT0000000805; CREB3L3 and ALDBSSCT0000011847; and TP53 and LNC_000076 (Figure 6).

Discussion

The follicle is a crucial tissue in mammal reproduction, and its development determines ovulation rate. Follicle groth and development is a complex and precise process, and long no coding RNA play an important role in thisfollicle development process. Some research evidence indicated the important roles of lncRNAs in animal reproduction with the development of gene chip and sequencing technologies. A. D. Macaulay used confocal transmission electron microscopy and RNA-Seq to determine that the cumulus cells around bovine oocytes transported a large number of nutrients and substances, e.g., mRNA and lncRNA on the oocytes of adult cows[15]. Sun detected 92 differentially expressed transcripts between Erhualian and large white pigs ovulatory follicles by using microarray[16]. Liu identified 2076 lncRNAs and 25491 mRNAs in Duroc's ovaries on days 0, 2, and 4 of follicular growth and development[17]. Meishan and Drouc sows show a huge difference in ovulation rate and ovarian follicle development and are good animal models for uncovering the molecular mechanisms underlying these differences. Therefore, we used RNA-Seq to identify the lncRNAs expressed in the M2 follicles of Meishan and Drouc sows. In our research, the newly identified lncRNAs in pig ovarian follicle have many obvious characteristics, it has shorter and fewer exons, longer exon length and lower abundance, and they are less conserved than mRNAs. The characteristics of these lncRNAs were compliance with those observed in other studies[18, 19]. This current research is a meaningful resource for further studies on the functional lncRNAs and mRNA in pig ovarian follicles.

We obtained 145 differentially expressed lncRNAs QTL analysis, and the result showed that 127 DE lncRNAs were located in known QTLs of corpus luteum number, litter size, androstenone, total number born alive, number of stillborn, FSH concentration and number of viable embryos. The previous studies showed that QTLs controlling ovulation rate are located in pig chromosome 8[20, 21]. QTLs of luteum and ovulation rate are in chromosome 3[22]. The QTLs of ovulation rate are in chromosomes 4, 7, 13 and 15. These QTLs are closely related to pig reproduction[23, 24]. In this study, we found that the QTLs enriched most of the differential lncRNAs, i.e., ovarian follicle lncRNAs, distributed in chromosomes 3, 7, 8, 13 and 15. The results indicated that the DE lncRNAs may be participate in the regulation of porcine reproductive traits and may be related to the follicular development in the two pig groups.

LncRNAs have a low expression abundance, and some lncRNAs are still unclearly. Thus, understanding the function of lncRNA is difficult. We obtained the target genes of DELs, and functional analysis found that these target genes participate in multiple biological processes and molecular signalling. CYP19A1 is the target gene of novel lncRNA-LNC_000757, it's expressed in granulosa cells(GCs) and is rate-limiting enzyme for oestrogen production. It can catalyse the conversion of androgens to oestrogens, thereby increasing the expression level of oestrogen, knockdown of CYP19A1 can result in the inability to ovulate and loss of corpus luteum in female mice[25]. CYP1A1, an important metabolic enzyme in the oestrogen metabolic process and ovarian steroidogenesis signalling pathway, can reduce the level of oestrogen and lead to the acceleration of follicular atresia[26, 27]. DE-lncRNA ALDBSSCT0000001929, ALDBSSCT0000006256 and ALDBSSCT0000002430 targeted and were positively correlated ($p<0.05$) with CYP1A1. The expression level of CYP1A1 was lower in Meishan, whereas that of CYP19A1 was higher in Meishan compared with Duroc. These results may indicate that the level of oestrogen in the M2 follicles of Meishan sows was higher than that in Duroc. In addition, IGF1 and HSD3B1 can stimulate and promote granulosa cells proliferation and promote follicular development[28, 29]. We also found HSD17B8, ALOX5, LHCGR, GNAS, CYP2J2 and CYP17A1 participate in ovarian steroidogenesis signalling pathway. These protein-coding genes may be candidate genes for porcine follicle development and are targeted by one or more lncRNAs (Figure 5C). Therefore, we speculated that lncRNAs may participate in porcine follicular development by regulating their target genes.

Studies have confirmed that the PI3K-AKT signalling pathway plays an major roles in the ovarian function regulation and follicular development[30-32]. PI3K-AKT signalling pathway regulates the primordial follicle maintenance and activation of s and promotes apoptosis of ovarian granulosa cells in humans and mice[33, 34]. KEGG analysis revealed that some target mRNAs of the DE lncRNAs were related to the PI3K-AKT signalling pathways. The THBS1-ITGA-ITGB complex members belong to the TGF-β family. TGF-β signalling can regulate the PI3K-Akt signalling pathway. The lncRNA-ALDBSSCT0000004 244 targets THBS1, which is an ECM. The expression of THBS1 is driven by oestradiol (E2), and its abnormal expression causes the apoptosis of granulosa cells and the acceleration of follicular atresia[35]. ITGA3 is the target gene of novel lncRNA LNC_000857 and LNC_001052. The expression of ITGA3 in M2 follicles of Meishan sows was higher than

in those of Duroc. However, at present, no reports of ITGA3 related to animal reproduction exist. ALDBSSCT0000010443 and ALDBSSCT0000009356 target ITGA6, which may be participate in the regulation of cumulus expansion and oocyte maturation[36]. ITGB1, ITGB4 and ITGB7 are the target genes of LNC_001556, LNC_000536 and ALDBSSCT0000004232, respectively. ITGB1 is associated with the apoptosis process[37] and is regulated by progesterone and oestrogen[38]. The expressions of PI3K and AKT is the golden standard for the activation of PI3K-AKT signalling pathway. We found that ENSSSCT000 00036444 and ALDBSSCT0000000805 target PIK3CA and PIK3C2B, and LNC_000751, LNC_000819 and LNC_000058 target AKT2. PIK3CA and PIK3C2B are the catalytic subunits of PI3K, and the knockout of PIK3CA and PIK3C2B causes early embryo death in mice[39]. PI3K can react with a variety of growth factors, thereby phosphorylating PIP2 to PIP3, activating PDK1, and then indirectly or directly activating Akt[40]. AKT2 is a subtype of AKT, and a downstream signalling core molecule of the PI3K-AKT classic signalling pathway[41, 42]. Activated AKT2 can phosphorylate its downstream signalling molecules and produces the cAMP that can activate CREB, TP53 and IKK protein for nuclear gene transcription and expression, thereby regulating cell proliferation[43, 44]. ALDBSSCT0000011847 and ALDBSSCT0000008900 target CREB3L3. ALDBSSCT0000004325 and LNC_001172 target the IKBKG, TP53 and ALDBSSCT0000002268 and were both significantly positively correlated in pairs ($p<0.05$), LNC_000076 and ALDBSSCT0000002268 target TP53. CREB3L3 is a transcription factor of CREB, and its function is similar to that of CREB. Its abnormal expression can attenuate the up-regulation of Egr1 by GnRH receptor activation, which in turn affects the expression level of LH-β and the growth of granular cells[45]. IKK can induce phosphorylation of IkB (inhibitory protein of NF-kB), dissociate the NF-kB/IkB dimer, activate NF-Kb, initiate the inhibition of apoptosis pathway and maintain porcine follicular development[46]. TP53 is a critical factor for cell survival, and suppression of p53 in oocytes can promote follicular growth and development [47]. In this study, we screened 19 DE-lncRNA candidates and their 12 target gene in the PI3K-AKT signalling pathway. We suggest that lncRNA may participate in follicular growth and development by regulating their target genes, but the specific functions still require systematic research.

Conclusions

We identified lncRNA and mRNA expressed in M2 follicles in Meishan and Duroc sows and found that some lncRNAs participate in the follicular development by regulating their target genes. These findings in transcriptome provide a valuable resource for follicular development and reproduction-related transcripts. The interactions between DE-lncRNAs and their PTGs and enriched pathways provide clues for further research on the role of follicular growth and development in pig. However, the lncRNAs function and molecular mechanism in follicular development remain unclear. Nevertheless, our study provides new insights into lncRNAs associated with follicular development and reproduction in pig.

Methods

Animals and sample collection

Three multiparous Meishan cyclic sows were maintained at the Agricultural Experimental Station of Yangzhou University. Three multiparous Duroc cyclic sows were raised at the Animal Experimental Station of the Shihezi University. Observe the sows every day and determine natural oestrous cycle, day 0 was the first oestrus day). Sows were injected with PGF2α according to the pigs' weight analogue (cloprostenol, Ningbo Second Hormone Factory) on the 14th day of the oestrous cycle to induce luteal regression and to synchronise the follicular growth phase. 4 days later, the sows were electric stunning and quick bleeding, then ovaries and medium follicles (follicles diameter: 5.0–6.9 mm, M2) were harvested , snap-frozen in LN2 and then stored at -80 ℃ refrigerator.

RNA-Seq preparation and sequencing analysis

Total RNAs of all the follicle samples were extracted using TRIzol reagent (Invitrogen). RNA degradation and contamination was monitored on 1% agarose gels. The total RNA concentration, integrity and purity were detected using qubit RNA assay kit in Qubit 2.0 flurometer (Life Technologies), RNA Nano 6000 assay kit by a bioanalyser 2100 (Agilent Technologies) and NanoPhotometer spectrophotometer (IMPLEN). The ribosomal RNAs of all samples were removed using the Ribo-zero rRNA Removal Kit (Epicentre). 6 strand-specific RNA-Seq pools for the M2 follicles of the six sows were constructed according to the manufacturer's instructions. Purify the samples library fragments using the AMPure XP system (Beckman Coulter, USA) to remove the preferred cDNA fragments of 150-200 bp. The blunt end cDNA fragments were augmented with A base and ligated to the sequencing adapter. The final products were purified (AMPure XP System), and the quality of library was assessed by the system Agilent Bioanalyser 2100. The samples pools were then analysed using one lane of 100-200nt paired-end HiSeq 4000 platform. Quality control (QC) of RNA-Seq reads was performed using Fast Q C.

Transcriptome assembly

We filtering reads with adapter and low-quality, poly-N reads from raw data through inhouse perl scripts, the clean reads were obtained. Clean data with high quality is the basis for all subsequent analyses. Reference genome files were downloaded from the Ensembl (Sus scrofa 10.2). Pigs reference genome annotation index was built using Bowtie v2.0.6[48]. Appropriate parameters were set using Tophat2 v 2.0.9[49]. The Scripture and Cufflinks [50] were used to assemble and splice the aligned sequences, which can be as small as possible. The transcript set, Cufflinks has specific parameters for the chain-specific library, and the directional information of the transcript chain were accurately provided.

lncRNA Identification

The following steps to identify lncRNAs from the nonredundant transcriptome. (1) Transcripts has single exon or two exons were filtered out. (2) Transcripts length with less than 200 bp were removed. (3) Any transcript with the FPKM (a fragment per kilobase of transcript per million mapped read) score lower than 0.5 in every pools was discarded. (4) The remaining transcripts were blasted in pig known annotation lncRNA database –ALDB database [51] using Cuffcompare. Only the transcript of lncRNAs whose splice sites were congruent between our results and those in ALDB were immediately brought in as a known lncRNA. (5) Transcripts of any known protein-coding were discarded, and transcripts that belonged to pre-miRNA, snRNA, rRNA, snoRNA and pseudogenes were removed. (6) The CNCI, CPC, and phyloCSF tool were used to calculate the transcripts that hascoding potential. Transcript with a CNCI score of <0[52], CPC score of <0[53], Pfam-scan E-value of <0.001[54] and phyloCSF Max-score of ≤100[55], as well as the intersection result of each tools were defined as novel found lncRNA transcript.

Differentially Expressed lncRNAs and mRNA Analysis.

Cufflinks package was used to conduct DE analysis between six follicle libraries of the Meishan and Drouc sows. The fold changes(FC) value were calculated in log2 (FPKM_MFM2/FPKM_DFM2) (FPKM_MFM2: FPKM of Meishan M2 follicle; FPKM_DFM2: FPKM of group Duroc M2 follicle). Transcripts with p-values of less than 0.05 would be identified as differentially expressed.

qRT-PCR Verification

Total RNAs were extracted using TRIzol (Invitrogen, CA, USA) and cDNA was synthesized using a RT-PCR kit (TaKaRa, Japan), qPCR reactions were performed using SYBR Green (TaKaRa Biotech, Dalian, China) according to the manufacturer's protocol. The reaction was conducted by combining 12.5 μl of 2× Real Master Mix (TaKaRa Biotechnology), 2 μL of cDNA, 1 μL each of the upstream and downstream primers, and 8.5 μL of RNase-free ddH2O water. Specific primers were designed using the Primer Premier 5.0 program (Table 3) and confirmed with BLAST. The expression levels of gene were normalized to linear GAPDH levels using $2^{(-\Delta\Delta Ct)}$ method[56] and the statistical difference was analysed using SPSS17.0. The correlation between the results of RNA-Seq and qPCR was calculated using a correlation test.

Table 3
The specific primers for qPCR.

Transcript type	Transcript name	Forward primer	Reverse primer	Product size (bp)
LncRNA	ENSSSCT00000018610	TGGTCTGCTCTAACCTGGACT	CTTCAGACAGGCTCAAGGGG	297
	ALDBSSCT0000001721	ACTCTTCAGTGGAGCTGACAA	TGGTCAAATTTTCCCTGGGATTG	81
	ALDBSSCT0000000051	AAGCAGAGGACGAAAAAG	CTACGCCACTCCAGAAAG	128
	LNC_000116	GCCTCTCTTGGTGCTTGGTT	TCGGTGGCTTCGGAGTTATT	134
	ALDBSSCT0000011300	CAAGGGGGTCAATTTTGCC	CACGCCTTGTGAATCGGTTT	122
	ALDBSSCT0000006152	AGTTCTCCAATGTCCCGTGTC	AGAAGACGCAGCCATCGGA	241
mRNA	COL3A1	ATCGCTGGTGTTGGAGGT	GAAGTCATAATCTTGTCGGTGT	100
	LRP8	CCAATCGCATCTACTGGTGTGAC	GGAGAGTGCAGCTGCTCATCAAT	115
	ENSSSCT00000009222	ATGCCTTCAATGGGACAACG	CAGTGGCTGGGGTAAGTCAA	262
	SEPP1	CCTTCATTGACCTCCACTAC	GTTGTCATACTTCTCATGGTTC	320
	COL5A2	GGGACATTTGGAAACCTGCC	GGGGAGTTATGGGGTCAGCA	114
	CYP19A1	CCAGCATTACCAGAAGCC	TGTGCCTCCATTACCGAG	92
	GAPDH	TTCCAGTATGATTCCACCCACG	TCGGCAGAAGGGGCAGAGAT	242

QTL analysis of Differentially Expressed lncRNAs

To predict the functions of DELs, QTLs analysis was performed. The location information of DELs was obtained from the transcriptome file and the QTL data of pig were downloaded from Animal QTLdb. Next, BEDTools and the “intersectBed” command were used for QTL analysis[57].

Prediction of PTGs of lncRNAs

Two methods were used to predicte the PTGs of lncRNAs. We identified the PTGs regulated by lncRNAs in cis, which were defined as protein-coding genes located at 100 Kb upstream and downstream of the lncRNA, by BEDTools 2.17.0[58]. The trans regulation of a lncRNA and its PTG was identified by the expression level correlation analysis of each pair of lncRNA and PTG. According to the Pearson’s correlation coefficients ($|r| > 0.95$), the PTGs were selected to construct a lncRNA–mRNA co-expression network.

GO and KEGG Pathway Enrichment Analysis

The prediction of the lncRNA and PTGs’ function such as gene enrichment and pathway analysis were performed using gene ontology (GO) (http://www.geneontology.org/)[59] and Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg)[60]. To further explore the main biological functions of differentially expressed genes, KOBAS software was used to detect the statistical enrichment of lncRNA PTGs or DEGs in KEGG pathways. The enrichment findings with a q-value of less than or equal to 0.05 was considered statistically significant.

Declarations

Ethics approval and consent to participate

All animal experiments in our study were carried out in accordance with the rules of Medical Ethics Committee, First Affiliated Hospital, Medical College, Shihezi University (2014-073-01, 5 March 2014).

Consent to publish

Not applicable.

Availability of data and materials

The datasets and supporting conclusions are included within this manuscript or its supporting files. The datasets generated during this study are available from the corresponding author upon request.

Competing interests

No potential conflicts of interest were disclosed.

Fundings

The work was supported by National Natural Science Foundation of China (NSFC: 31460586 and 31960645), The young and middle-aged leading talents of Xinjiang Production and Construction Corps Science and Technology Innovation Talents Program (2020CB018).

Authors' contributions

TH conceived and designed the experiments. MXL and YL analyzed main content of the data and wrote the paper with the help of TH. HB G and SX mainly collects follicle samples. ZCZ performed the RNA extraction. LPM and YSS performed the qPCR experiment and data analysis.

Acknowledgements

We are grateful for the generous help of TH and YL in data analysis. We are grateful to LPM, SX, HBG, YSS, ZCZ for the help of the experimental. In addition, we also grateful for the advice of doctoral student XL and the master student ZYH from Huazhong Agricultural University in RNA seq analysis.

References

1. Rosendo A, Druet T, Gogue J, Canario L, Bidanel JPJJoAS: Correlated responses for litter traits to six generations of selection for ovulation rate or prenatal survival in French Large White pigs. 2007(7):7.
2. Davis KL, Robison OW, Toelle VDJJoAS: Breed Differences in Uterine and Ovarian Measurements in Gestating Swine. 1987(3):3.
3. Lee CJ, Kim JKJV: Decreased retinoblastoma protein expression in γ-irradiated mouse ovarian follicles. 1999, 13(3):263-266.
4. Shimizu T, Miyahayashi Y, Yokoo M, Hoshino Y, Sasada H, Sato EJR: Molecular cloning of porcine growth differentiation factor 9 (GDF-9) cDNA and its role in early folliculogenesis: direct ovarian injection of GDF-9 gene fragments promotes early folliculogenesis. 2004, 128(5):537-543.
5. Yang S, Wang S, Luo A, Ding T, Reproduction SWJBo: Expression patterns and regulatory functions of microRNAs during the initiation of primordial follicle development in the neonatal mouse ovary. 2013, 89(5):126.
6. Jin Lu, Cuilian Zhang, Baoxia Gu, Shaodi Zhang, Jiaxuan Geng, Yuanhui Chen, Juanke Xie. MicroRNA-182 inhibits rat ovarian granulosa cell apoptosis by targeting Smad7 in polycystic ovarian syndrome. Int J Clin Exp Pathol 2017; 10(2):1380-1387
7. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q et al: Long noncoding RNAs with enhancer-like function in human cells. Cell 2010, 143(1):46-58.
8. Prickett AR, Oakey RJJMG, Genomics: A survey of tissue-specific genomic imprinting in mammals. 2012, 287(8):p.621-630.
9. Paralkar VR, Mishra T, Luan J, Yao Y, Kossenkov AV, Anderson SM, Dunagin M, Pimkin M, Gore M, Sun DJB: Lineage and species-specific long noncoding RNAs during erythro-megakaryocytic development. 2014, 123(12):1927-1937.
10. Zhao W, Mu Y, Ma L, Wang C, Tang Z, Yang S, Zhou R, Hu X, Li MH, Li KJSR: Systematic identification and characterization of long intergenic non-coding RNAs in fetal porcine skeletal muscle development. 2015, 5:8957.
11. Guttman, Mitchell, Garber, Manuel, Levin, Joshua, Donaghey, Julie, Biotechnology RJN: Corrigendum: Ab initio reconstruction of cell type–specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. 2010.

12. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG et al: The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012, 22(9):1775-1789.
13. Ghosal S, Das S, Cells JCJS, Development: Long Noncoding RNAs: New Players in the Molecular Mechanism for Maintenance and Differentiation of Pluripotent Stem Cells. 2013.
14. Rothstein JMJPT: The Journal: Past, Present, and Future. 1989, 69(10):393-401.
15. Macaulay AD, Gilbert I, Caballero J, Barreto R, Fournier E, Tossou P, Sirard MA, Clarke HJ, Khandjian EW, Richard FJJBoR: The Gametic Synapse: RNA Transfer to the Bovine Oocyte. 2014, 91(4):90-90.
16. Sun PR, Jia SZ, Lin H, Leng JH, Lang JH: Genome-wide profiling of long noncoding ribonucleic acid expression patterns in ovarian endometriosis by microarray. Fertil Steril 2014, 101(4):1038-1046 e1037.
17. Yi L, Li M, Bo X, Tao L, Ma L, Zhai T, Tao HJIJoMS: Systematic Analysis of Long Non-Coding RNAs and mRNAs in the Ovaries of Duroc Pigs During Different Follicular Stages Using RNA Sequencing. 2018, 19(6):1722-.
18. Yalan Y, Rong Z, Shiyun Z, et al. Systematic Identification and Molecular Characteristics of Long Noncoding RNAs in Pig Tissues[J]. Biomed Research International, 2017, 2017:6152582.
19. Tang Z, Wu Y, Yang Y, et al. Comprehensive analysis of long non-coding RNAs highlights their spatio-temporal expression patterns and evolutional conservation in Sus scrofa[J]. Sci Rep, 2017, 7: 43166.
20. Wilkie PJ, AA Paszek, et al.Scan of eight porcine chrosomes for growth, carcass and reproductive traits reveals two likely quantitative traits loci[J]. Anim, Genet,1996,27(Suppl.2):117-118.
21. Rathje TA, GA Rohrer, RK Johnson. Evidence for quantitative traits loci affecting ovulation rate in pigs[J]. Anim. Sci, 1997, 75:1486-1494.
22. Stao S, Atsuji K, et al. Identification of quantitative traits loci affecting lutea and number of teats in a Meishan x Duroc pigs F2 resource population[J], Anim Sci, 2006,84:2895-2901.
23. Cassady JP, RK Johnson, et al. Identification of quantitative traits loci affecting in pigs[J]. Anim. Sci, 2001, 79:623-633.
24. Holl JW, JP Cassady, et al. A genome scan for quantitative traits loci and imprinted regions affecting reproduction in pigs[J]. Anim. Sci, 2004, 82:3421-3429.
25. Charlotte Dubé, Bergeron F , Marie-Josée Vaillant, et al. The nuclear receptors SF1 and LRH1 are expressed in endometrial cancer cells and regulate steroidogenic gene transcription by cooperating with AP-1 factors[J]. Cancer Letters, 2009, 275(1):127-138.
26. Han W, Pentecost BT, Pietropaolo RL, Fasco MJ, Spivack SD: Estrogen receptor alpha increases basal and cigarette smoke extract-induced expression of CYP1A1 and CYP1B1, but not GSTP1, in normal human bronchial epithelial cells. 2010, 44(3):202-211.34.
27. Sowers MFR , Wilson AL , Kardia SR. CYP1A1 and CYP1B1 polymorphisms and their association with estradiol and estrogen metabolites in women who are premenopausal and perimenopausal[J]. American Journal of Medicine, 2006, 119(9-supp-S1):S44-S51.
28. Shiomi-Sugaya N , Komatsu K , Wang J , et al. Regulation of secondary follicle growth by theca cells and insulin-like growth factor 1[J]. Journal of Reproduction & Development, 2015, 61(3):161.
29. Sechman A, Antos P, Katarzyńska D, Grzegorzewska A, Wojtysiak D, Hrabia A. 2014. Effects of 2, 3, 7, 8-Tetrachlorodibenzo-p-Dioxin on Secretion of Steroids and STAR, HSD3B and CYP19A1 MRNA Expression in Chicken Ovarian Follicles. Toxicology Letters, 225(2): 264–74.
30. Reddy P, Zheng W, and Liu K. Mechanisms maintaining the dormancy and survival of mammalian primordial follicles [J]. Trends in Endocrinology & Metabolism, 2009: 1-8.
31. Reddy P, Liu L, Adhikari D, Jagarlamudi K, Rajareddy S, Shen Y, Du C, Tang W, Hamalainen T, Peng SL, Lan ZJ, Cooney AJ, Huhtaniemi I, and Liu K. Oocyte-specific deletion of Pten causes premature activation of the primordial follicle pool[J]. Science, 2008, 319(5863): 611-613.
32. Gonzalez-Robayna IJ, Falender AE, Ochsner S, Firestone GL, and Richards JS. Follicle-Stimulating hormone (FSH) stimulates phosphorylation and activation of protein kinase B (PKB/Akt) and serum and glucocorticoid-Induced kinase (Sgk): evidence for A kinase-independent signaling by FSH in granulosa cells[J]. Mol Endocrinol, 2000, 14(8): 1283-1300.
33. Zheng W, Nagaraju G, Liu Z, et al. Functional roles of the phosphatidylinositol 3-kinases (PI3Ks) signaling in the mammalian ovary[J]. molecular & cellular endocrinology, 2012, 356(1-2):24-30.
34. John GB, Shidler MJ, Besmer P. Kit signaling via PI3K promotes ovarian follicle maturation but is dispensable for primordial follicle activation[J]. Dev Biol, 2009, 331(2): 292-299.

35. Garside SA, Harlow CR, Hillier SG, Fraser HM, Thomas FHJE: Thrombospondin-1 Inhibits Angiogenesis and Promotes Follicular Atresia in a Novel in Vitro Angiogenesis Assay. 2010(3):3.
36. Yerushalmi G M , Salmon-Divon M , Yung Y. Characterization of the human cumulus cell transcriptome during final follicular maturation and ovulation[J]. Molecular Human Reproduction(8):719-735.
37. Shen M, Li T, Zhang G, Wu P, Chen F, Lou Q, Chen L, Yin X, Zhang T, Wang JJBG: Dynamic expression and functional analysis of circRNA in granulosa cells during follicular development in chicken. 20(1).
38. Yoshimura Y. Integrins□expression□modulation□and signaling in fertilization□embryogenesis and implantation[J]. Keio Journal of Medicine, 1997, 46(1): 16-24.
39. Stocker H, Andjelkovic M, Oldham S, Laffargue M, Wymann MP, Hemmings BA, and Hafen E. Living with lethal PIP3 levels: viability of flies lacking PTEN restored by a PH domain mutation in Akt/PKB [J]. Science, 2002, 295(5562): 2088-2091.
40. Zheng W, Nagaraju G, Liu Z, and Liu K. Functional roles of the phosphatidylinositol 3-kinases(PI3Ks) signaling in the mammalian ovary[J]. Molecular and cellular endocrinology, 2012, 356: 24-30.
41. Coffer PJ, Jin J, Woodgett JR. Protein kinase B (c-Akt): a multifunctional mediator of phosphatidylinositol 3-kinaseactivation[J]. Biochem Journal, 1998, 335: 1-13.
42. Alessi D R, Cohen P. Mechanism of activation and function of protein kinase B[J]. Current opinion in genetics &development, 1998, 8(1): 55-62.
43. Koseoglu S, Lu Z, Kumar C, et al. AKT1, AKT2 and AKT3-dependent cell survival is cell line-specific and knockdown of all three isoforms selectively induces apoptosis in 20 human tumor cell lines[J]. Cancer Biology & Therapy, 2007, 6(5):755-762.
44. Cantley LC, Neel BG. New insights into tumor suppression: PTEN suppresses tumor formation by restraining thephosphoinositide 3-kinase/AKT pathway[J]. Proceedings of the National Academy of Sciences, 1999, 96(8): 4240-4245.
45. Mayer SI, Thiel G. Calcium influx into MIN6 insulinoma cells induces expression of Egr-1 involving extracellular signalregulated protein kinase and the transcription factors Elk-1 and CREB, Eur J Cell Biol, 2009, 88(1): 19-33.
46. Mizuho N, Noboru M, Naoko I. Changes in the Expression of Tumor Necrosis Factor (TNF) α, TNFα Receptor (TNFR) 2, and TNFR-Associated Factor 2 in Granulosa Cells During Atresia in Pig Ovaries1[J]. Biology of Reproduction(2):2.
47. Zhang CX, Zhang Q, Xie YY, et al. Mouse Double Minute 2 Homolog Actively Suppresses p53 Activity in Oocytes during Mouse Folliculogenesis[J]. The American Journal of Pathology, 2016:S0002944016304655.
48. Langmead B, Salzberg SL: Fast gapped-read alignment with Bowtie 2. Nat Methods 2012, 9(4):357-359.
49. Kim D, Pertea G, Trapnell C, et al. Tophat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions[J]. Genome Biology, 2013, 14(4): R36.
50. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L: Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 2010, 28(5):511-515.
51. Li A, Zhang J, Zhou Z, Wang L, Liu Y, Liu YJPO: ALDB: A Domestic-Animal Long Noncoding RNA Database. 2015, 4(10).
52. Liang S, Luo H, Bu D, Zhao G, Yu K, Zhang C, Liu Y, Chen R, Yi ZJNAR: Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. 2013(17):17.
53. Lei Kong, Yong Zhang, Zhi-Qiang Ye, Xiao-Qiao Liu, Shu-Qi Zhao, Liping Wei, Ge Gao. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine[J]. Nucleic acids research, 2007, 35(Web Server issue):W345-9.
54. Bateman A, Coin L, Durbin R, et al. The Pfam Protein Families Database[J]. Nucleic Acids Research, 2008, 32(1):D138.
55. Lin MF, Jungreis I, Kellis M. Phylocsf: a Comparative Genomics Method to Distinguish Protein Coding and Non-coding Regions[J]. Bioinformatics, 2011, 27(13): i275-i282.
56. Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001, 25(4):402-408.
57. Yu H, Zhao X, Li QJSR: Genome-wide identification and characterization of long intergenic noncoding RNAs and their potential association with larval development in the Pacific oyster. 2016, 6:20796.
58. Aaron R, Quinlan,Ira M, Hall.BEDTools: a flexible suite of utilities for comparing genomic features[J]. Bioinformatics (Oxford, England), 2010, 26(6): 841-2.
59. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks[J]. Bioinformatics,2009,25(8):1091-1093.

60. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M: KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 2016, 44(D1):D457-462.

Figures

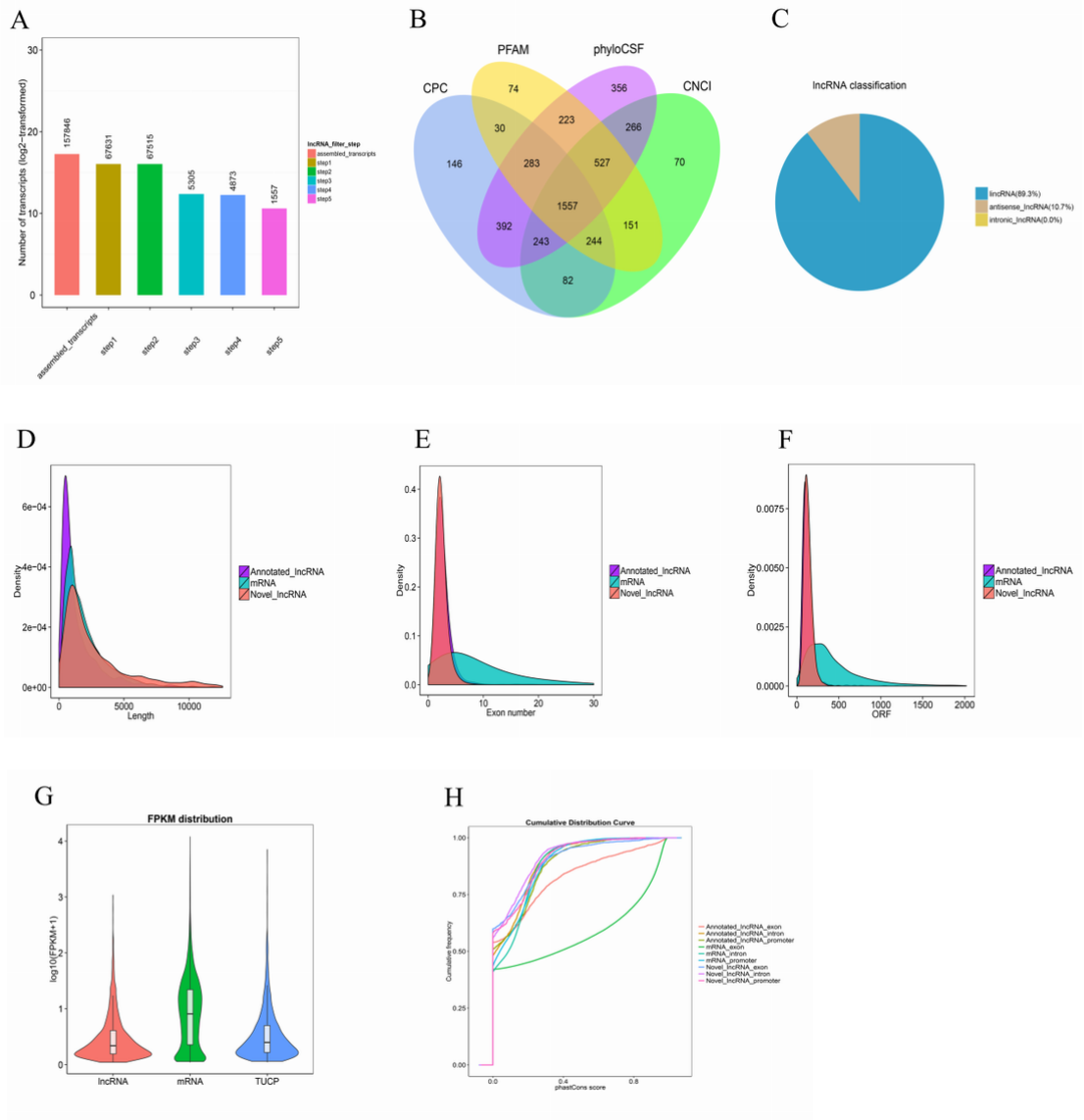


Figure 1

Identification and characterization lncRNAs and mRNAs in M2 follicles. (A) 157, 846 transcripts were assembled; (B) Identification of non-coding lncRNAs by using four tools-CPC, PFAM, phyloCSF and CNCI; (C)lncRNAs classification; (D) The length of lncRNAs and mRNAs; (E) The lncRNAs and mRNAs exon number; (F) open reading frame (ORF) length distribution of mRNAs and lncRNAs; (G) Violin plot of expression abundance (showed in log10 (FPKM + 1)) for mRNAs and lncRNAs. (H) Conservation compare with mRNAs and lncRNAs by using phasConsoftware.

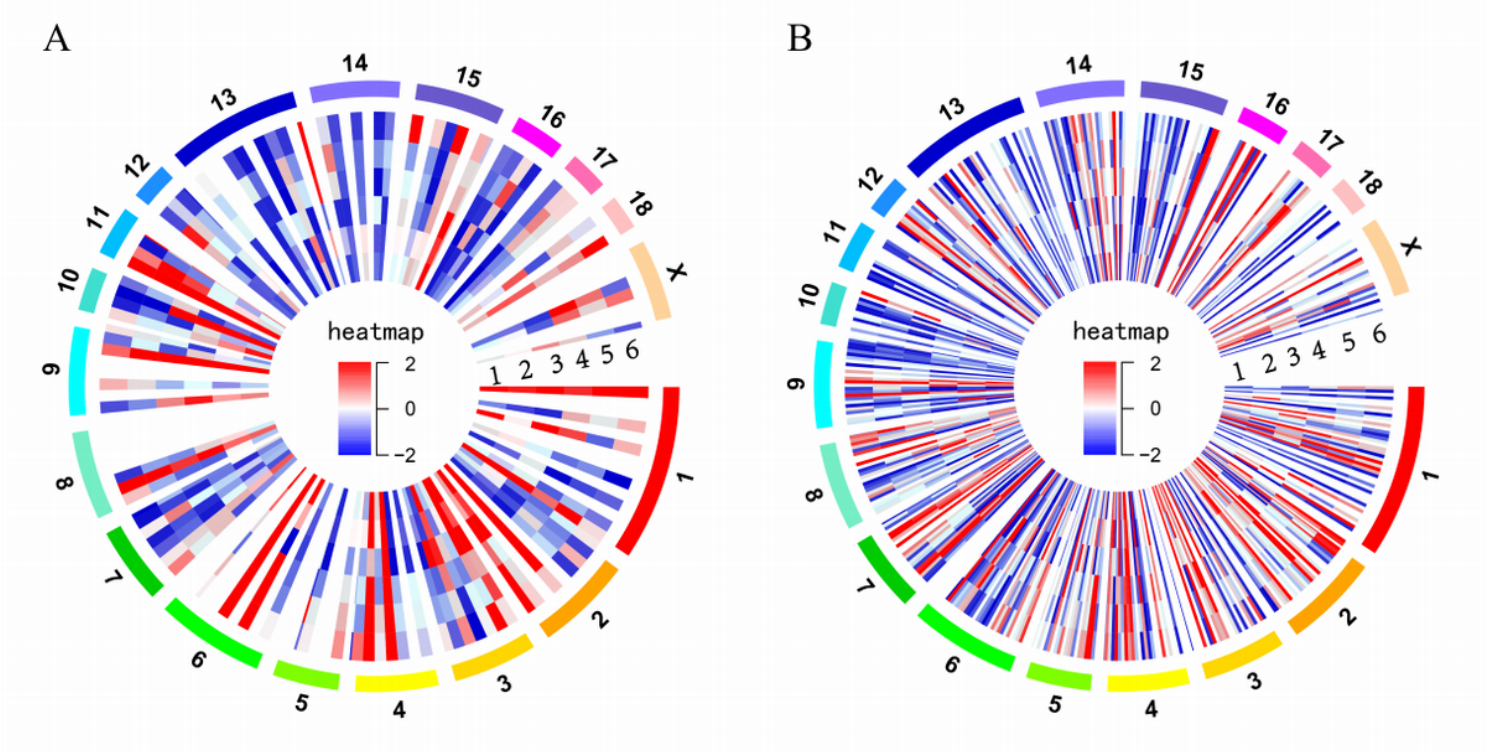
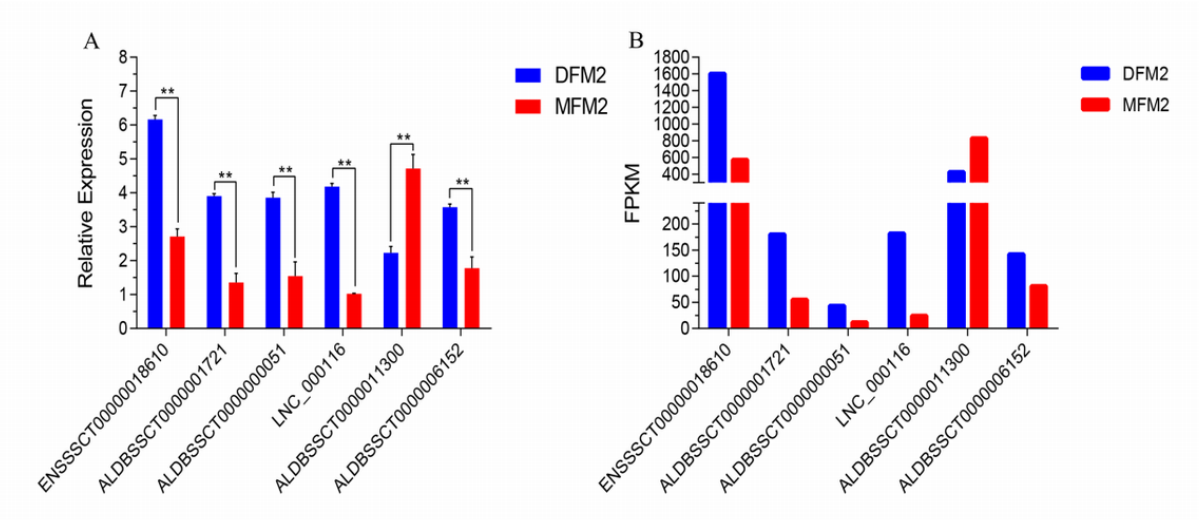


Figure 2

Differential Expression(DE) Analysis of lncRNA and mRNA in two pig breeds. (A) DE lncRNAs between the Meishan and Duroc M2 follicle. Loop heat map from inside to outside, 1: MFM2DY4_1, 2: MFM2DY4_2, 3: MFM2DY4_3, 4: DFM2DY4_1, 2: DFM2DY4_2, 3: DFM2DY4_3. (B) DE protein-coding genes between the Meishan and Duroc follicle. Loop heat map from inside to outside, 1: MFM2DY4_1, 2: MFM2DY4_2, 3: MFM2DY4_3, 4: DFM2DY4_1, 2: DFM2DY4_2, 3: DFM2DY4_3.



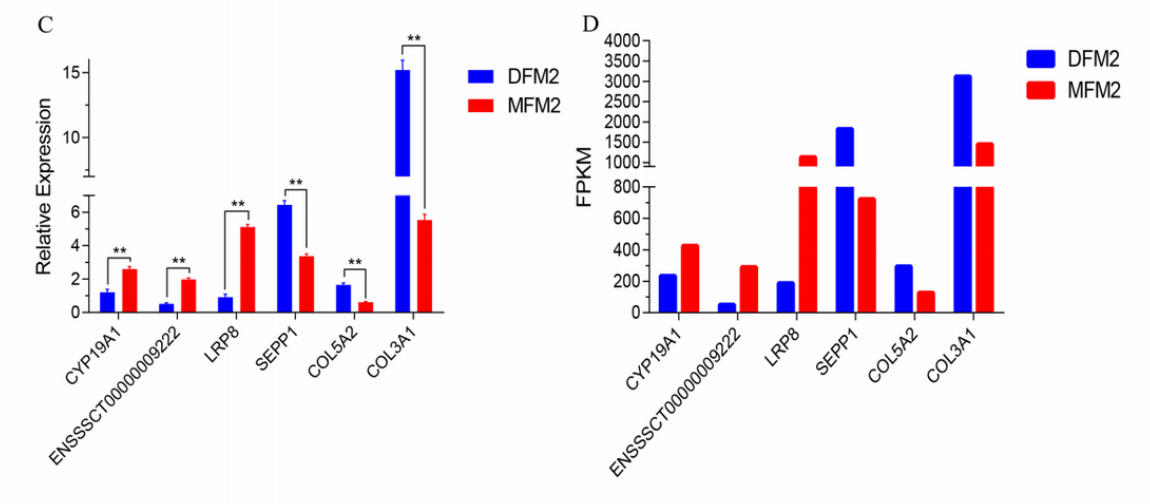


Figure 3

Validation of the expression levels of lncRNAs and mRNAs. (A, C) The qPCR verification of the 6 DE-mRNAs in DFM2DAY4 and MFM2DAY4. (B, D) The RNA-seq results between DFM2DAY4 and MFM2DAY4. The expression of transcripts was normalized by GAPDH. The results were expressed as mean±SE, * represents P<0.05, ** represents P<0.01.

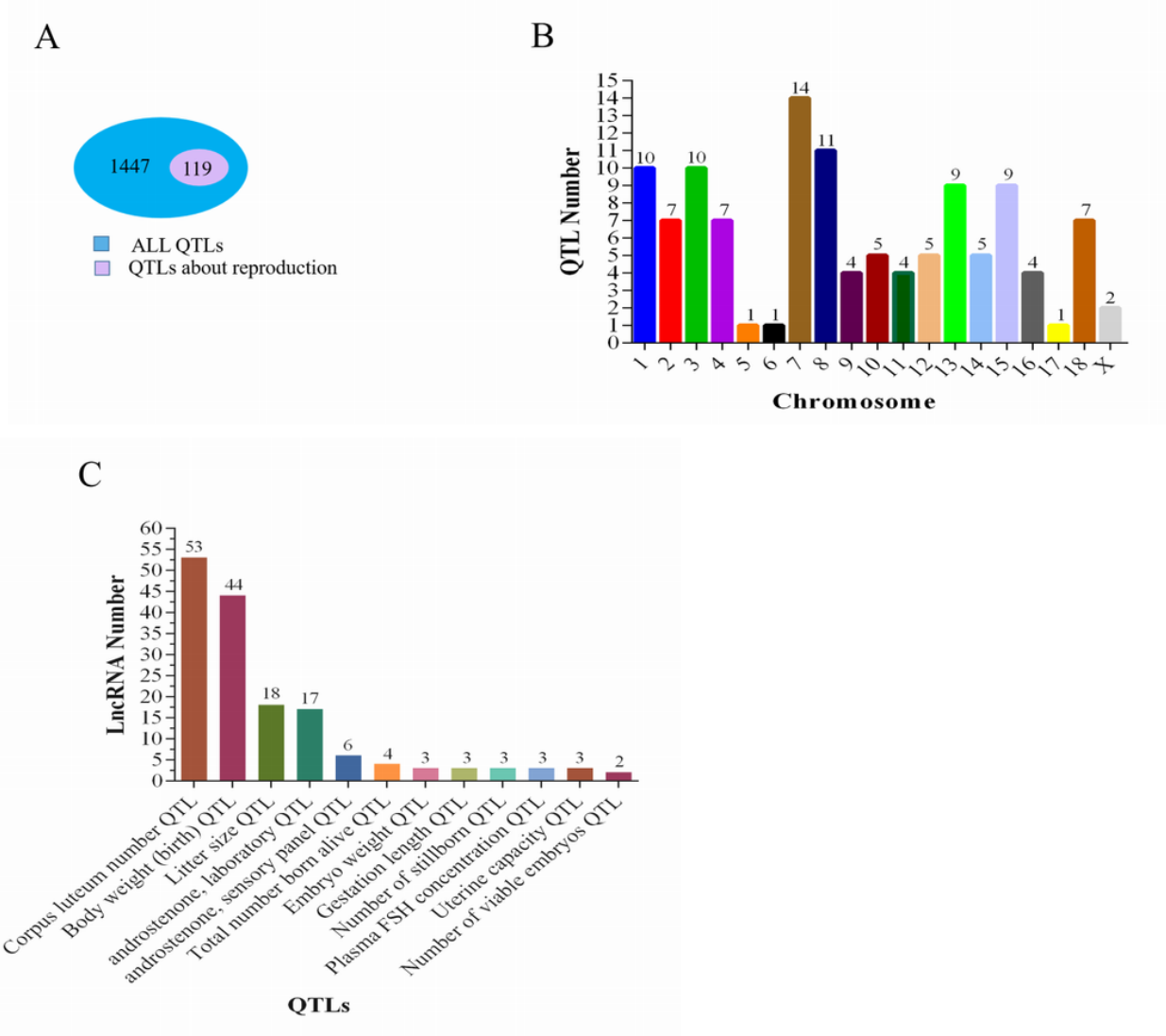


Figure 4

QTLs analysis of DELs. (A) The number distribution of QTLs associated with reproduction and all of the QTLs. (B) The chromosome distribution of QTLs associated with reproduction. (C) The lncRNA number of QTLs associated with reproduction QTLs.

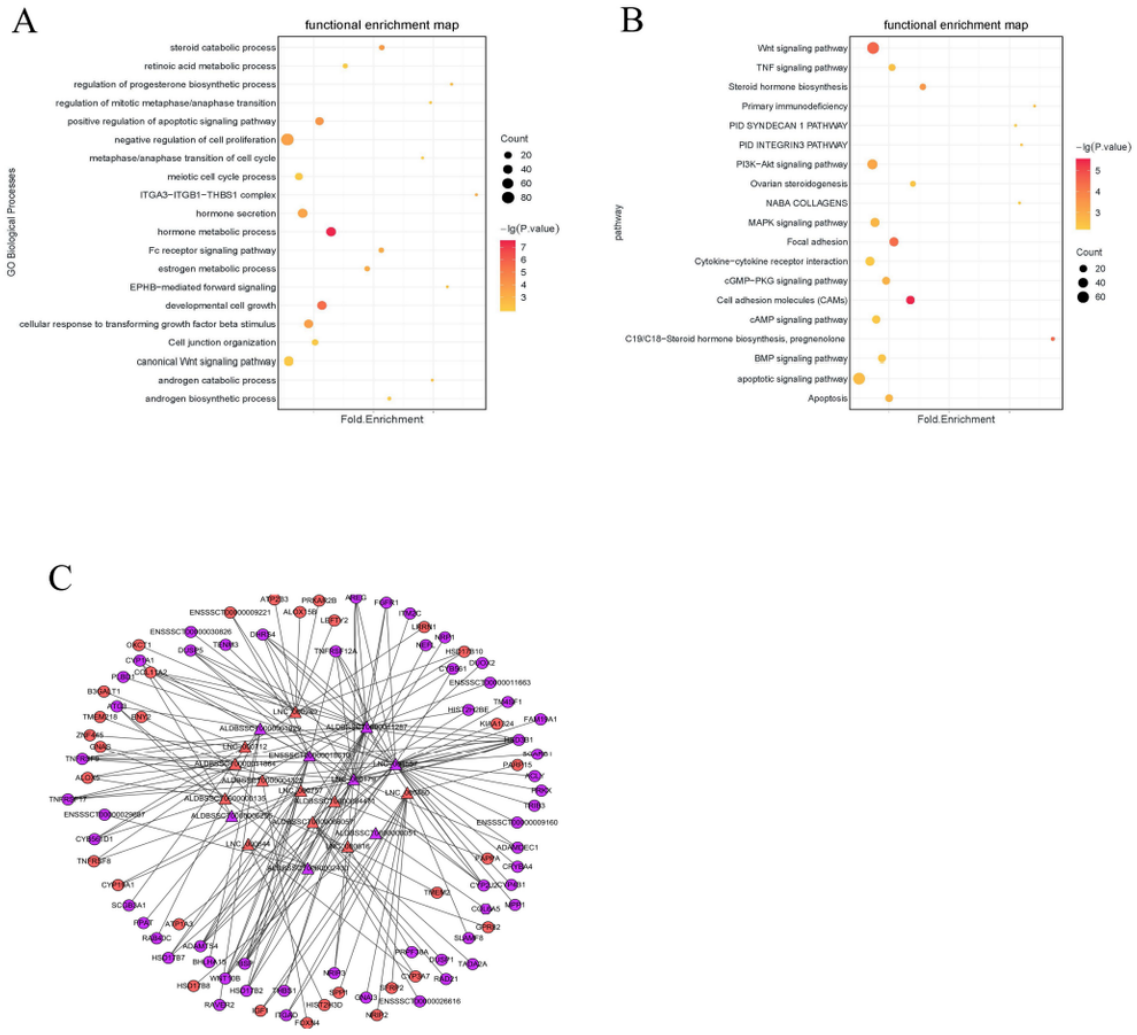


Figure 5

GO and KEGG pathway analysis of the PTGs of differentially expressed lncRNAs (DELs) (A) Biological processes of PTGs of DELs. (B) Pathways of PTGs of DELs. (C) The interaction analysis of PTGs and DELs in estrogen metabolic process and ovarian steroidogenesis signaling pathway. Triangles represent lncRNAs, circles represent mRNA, and red is up-regulated genes, purple is down-regulated genes.

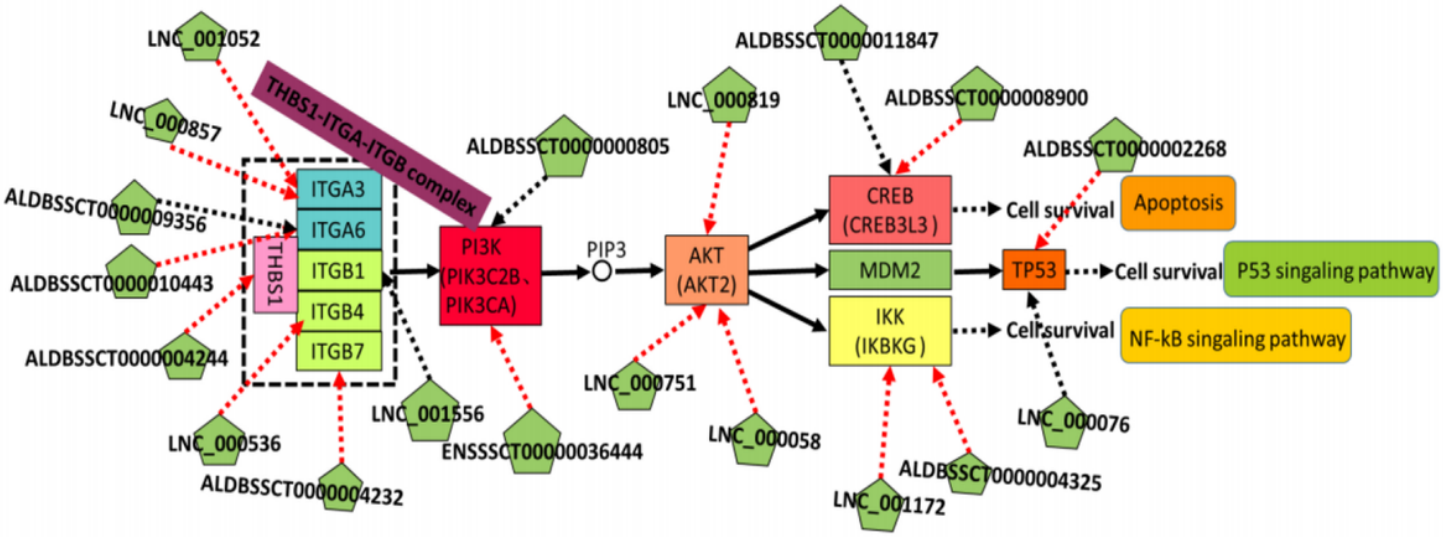


Figure 6

The regulatory network analysis and differentially esxpressed transcripts involved in PI3K-AKT signaling pathway in our study(RNA-Seq data). The red arrow represents positive correlation, the black arrow represents negative correlation.