



Elucidating the Genetic Regulation of Reproductive Prolificacy in Tibetan Sheep Through Ovarian Tissue Sequencing

Wu Sun^{1,2,3}, Xiayang Jin^{1,2,3}, Yuhong Ma^{1,2,3}, Shike Ma^{1,2,3}

10.18805/IJAR.BF-1787

ABSTRACT

Background: This study aims to unravel the genetic intricacies regulating the reproductive prolificacy of Tibetan sheep, an indigenous breed prominently found in the Qinghai-Tibet Plateau.

Methods: We employed meticulous sequencing and analytical strategies to discern the differential expressions of long non-coding RNAs (lncRNAs) and messenger RNAs (mRNAs) in the ovarian tissues of ewes with varying lambing numbers.

Result: In our study, we identified 817 genes with differential expression in twin-bearing ewes compared to those bearing single lambs. This group comprised 495 upregulated and 322 downregulated genes, featuring key candidates like COL1A1, COL1A2, BMPRII, GDF9, BMP15 and TGFB2. Furthermore, we discovered a notable difference in the expression of long non-coding RNAs (lncRNAs) in ewes birthing twins as opposed to singles, with 210 differentially expressed lncRNAs identified, including 91 upregulated and 119 downregulated. Key genes and lncRNAs associated with reproductive efficiency were identified, contributing valuable insights into the underlying molecular mechanisms and offering a robust foundation for enhancing breeding strategies and overall livestock productivity.

Key words: Genetic regulation, lncRNAs, mRNAs, Ovarian tissues, Reproductive prolificacy, Sequencings, Tibetan sheep.

INTRODUCTION

The reproductive traits of livestock, such as litter size, hold significant importance in agricultural economics. This is because improving the reproductive rate of livestock contributes to increasing the productivity of animal husbandry, meeting the growing human demand for meat and dairy products (Tian *et al.*, 2009; Abdoli *et al.*, 2016). Tibetan sheep, widely distributed in China's Qinghai region, are popular due to their adaptation to cold environments, excellent fur and meat quality. However, Tibetan sheep typically belong to a single-lamb breed, giving birth to only one lamb at each parturition. They are predominantly reared under extensive grazing systems, which take advantage of the vast natural grasslands available in high-altitude regions. This traditional rearing method supports the sustainability of the ecosystem but poses challenges for increasing reproductive rates and managing nutritional needs, particularly during harsh winter months. This undoubtedly limits their reproductive capacity, making it a priority to improve the litter size of Tibetan sheep.

Advancements in omics technologies, particularly transcriptome and lncRNA sequencing, have illuminated the genetic regulation of reproductive traits in mammals. They reveal dynamic, stage-specific gene expressions, essential for ovarian development and litter size determination (Yang *et al.*, 2018; Soygur *et al.*, 2021; Jones *et al.*, 2022). For instance, in mice, distinct expression patterns were observed amongst 15,454 and 16,646 transcriptionally active genes during various developmental stages, highlighting the functional relevance in follicle maturation (Pan *et al.*, 2014). In pig studies, notable discoveries include Zhang *et al.* (2015) identification of

¹Academy of Animal Science and Veterinary Medicine, Qinghai University, Xining, 810016, China.

²Key Laboratory of Livestock and Poultry Genetics and Breeding on the Qinghai-Tibet Plateau, Ministry of Agriculture and Rural Affairs, Xining, 810016, China.

³Plateau Livestock Genetic Resources Protection and Innovative Utilization Key Laboratory of Qinghai Province, Xining, 810016, China.

Corresponding Author: Shike Ma, Academy of Animal Science and Veterinary Medicine, Qinghai University, Xining, 810016, China. Email: 1064780694@qq.com

How to cite this article: Sun, W., Jin, X., Ma, Y. and Ma, S. (2024). Elucidating the Genetic Regulation of Reproductive Prolificacy in Tibetan Sheep Through Ovarian Tissue Sequencing. Indian Journal of Animal Research. 58(9): 1460-1473. doi: 10.18805/IJAR.BF-1787.

Submitted: 16-03-2024 **Accepted:** 05-06-2024 **Online:** 09-09-2024

1243 differentially expressed genes in Yorkshire pigs and a separate study revealing 2076 lncRNAs and 25,491 mRNAs in Duroc pigs, underscoring a complex gene interaction network in ovarian tissues (Zhang *et al.*, 2015; Liu *et al.*, 2018). In bovine research, Song *et al.* (2022) unveiled 3635 mRNAs, among which 1608 were significantly upregulated and 2027 were downregulated, spotlighting genes like ADCY3, ADCY4, ADCY9, BMP15, with differential expressions instrumental in ovarian and uterine tissue developments in Xiangxi yellow cattle. Sheep studies are also profound, identifying over 110 genes influencing Hu sheep ovarian activities and intricate expressions of 37,309 lncRNAs and 45,404 mRNAs, emphasizing genes like MSTRG.162061.1 in key developmental pathways (Chen *et al.*, 2012; Chen *et al.*, 2018;

Boruah *et al.*, 2021; Chang *et al.*, 2021). However, these comprehensive studies predominantly focus on high reproductive rate breeds, necessitating further research on breeds with naturally lower reproductive rates to bridge the existing knowledge gap. This concise summary retains pivotal numerical and gene-specific insights, enhancing the readability while maintaining academic rigor.

In this study, Tibetan sheep, a distinguished native breed prevalent in Qinghai, Tibet and Gansu, China, were utilized. Historically classified academically as a single-lamb breed, Tibetan sheep typically produce one lamb per litter. Through the comparative analysis of gene expression profiles between high and low lambing ovaries, our objective is to identify pivotal genes and regulatory pathways that influence lambing numbers. This investigation aims to enrich the theoretical foundation essential for enhancing the reproductive efficacy of Tibetan sheep, additionally offering valuable genetic molecular markers and insightful references for advancing the reproductive research of other livestock species.

MATERIALS AND METHODS

Sample collection and RNA extraction

In this study, we selected 20 healthy Tibetan ewes with comparable weights from the Plateau Ecological Animal Husbandry Park in Qinghai, split equally between multiple-birth (SL group: 43.94 ± 3.42 kg) and single-birth groups (ML group: 41.10 ± 4.88 kg). The ewes grazed freely, with extended foraging times to enhance feeding. Ovarian tissues were harvested using Xu *et al.* (2018) method, involving dissection on a cold surface and collecting 1 cm³ from different sections of the right ovary, followed by immediate freezing in liquid nitrogen and storage at -80°C for RNA extraction. All procedures complied with Qinghai University's animal care protocols. RNA was extracted using the TRIzol method, with purity and concentration verified by NanoDrop™ 2000 and integrity by Agilent 2100 Bioanalyzer.

Library construction and sequencing

Following the manufacturer's guidelines, we used the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® to establish the RNA libraries. The library construction began with the isolation of Poly (A)-containing mRNA using the NEBNext Poly (A) mRNA Magnetic Isolation Module. Subsequently, the kit's reagents facilitated library preparation, including fragmentation, cDNA synthesis, end repair, adapter ligation and PCR amplification. Upon completing the library construction, dual-end sequencing on the Illumina HiSeq × 10 platform yielded high-quality data. Quality control was then applied to the raw sequencing data, removing low-quality sequences, reads with over 5% unknown bases and adapter-contaminated sequences.

Ribosomal RNA filtering and reference genome alignment

The short-read alignment tool SOAP2 aligns clean reads to the ribosomal database, tolerating up to 5 mismatches

and removes reads aligning to ribosomes. The retained data are then used to predict and quantify lncRNAs. Reads filtered for rRNA are aligned to the sheep oar3.0 reference genome using Hisat2 software (Kim *et al.*, 2015).

Expression level calculation and transcripts analysis

We calculate the expression levels of both known and new genes and transcripts using the R package Ballgown. Ballgown employs the FPKM method (Pertea *et al.*, 2016) to represent gene expression levels by integrating results from StringTie. New transcripts are those without known annotation in the reference database. This means new transcripts may be either new splicing variants of known genes or entirely unknown genes' transcripts. We reconstructed the transcripts of each sample using StringTie (Pertea *et al.*, 2015). Subsequently, we integrated all samples' reconstruction data using StringTie merge. We then compared the integrated data with reference annotations using gffcompare (Pertea *et al.*, 2015), identifying and defining transcripts not in the known annotations as new.

Selection of candidate lncRNAs

In our study, we meticulously selected lncRNAs, transcripts over 200 nucleotides long that do not encode proteins, based on criteria reflecting their structural and functional characteristics. The initial selection included basic screening and coding potential analysis (Yan *et al.*, 2020). For basic screening, we used StringTie to merge transcripts from all samples, selecting those with at least 3× coverage and a length of 200bp or more. We identified new transcripts by comparing the merged results with existing lncRNA or mRNA annotations, when available. Coding potential was assessed using mainstream methods like CPC (Kong *et al.*, 2007), CNCI (Sun *et al.*, 2013) and pfam protein domain analyses (Mistry *et al.*, 2007). Transcripts deemed non-coding by at least two of these tools were classified as lncRNAs; the rest were categorized as mRNA. Our identification strategy, inspired by Yan *et al.* (2020) methodologies (Yan *et al.*, 2020), integrated contemporary coding potential assessment tools. This methodical screening process yielded a refined set of candidate lncRNAs, visually represented in a Venn diagram for comprehensive statistical interpretation. This approach ensures the rigorous and reliable selection of lncRNAs, providing robust data for further investigation in the study.

Screening of DEGs and lncRNA and functional analysis

Differential expression genes and lncRNAs were screened using DESeq2 software (Love *et al.*, 2014). The criteria set for differentially expressed genes and lncRNAs were $|\log_2(\text{FoldChange})| > 1$ and $\text{FDR} < 0.001$. KEGG Mapper tool (Kanehisa *et al.*, 2019) was utilized to perform Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis.

Protein-protein interaction network analysis

Proteins typically function by interacting and forming complexes. We primarily use the interaction relationships

in the STRING protein interaction database (<http://string-db.org/>) for the analysis of differential gene protein interaction networks (PPI) (Szkłarczyk *et al.*, 2023). Through PPI analysis, DEGs (Differentially Expressed Genes) that interact with each other usually have similar functions. We use the stringDB package in R's bioconductor (Damian *et al.*, 2021) to perform PPI analysis. A differential gene protein interaction network data file is provided, which can be directly imported into the Cytoscape software (Shannon *et al.*, 2003) for visual editing.

Validation of mRNA by qRT-PCR

Twelve candidate genes were selected for validation by qPCR. Total RNA was extracted using Trizol reagent (TransGen Biotech, Beijing, China) according to the manufacturer's instructions. After obtaining high-quality total RNA, mRNA was reverse transcribed using the RevertAid™ First Strand cDNA Synthesis Kit (TransGen Biotech, Beijing, China). Quantitative real-time PCR (qRT-PCR) analysis of mRNA was performed in the Bio-rad CFX96 real-time system using SYBR Green PCR Master Mix (ABI, USA, 4304437). β -actin was used as an internal control for each mRNA. All mRNA primer sequences were designed using primer 5

(Table 1). For mRNA quantification, the reaction conditions were: 95°C for 10 minutes, followed by 40 cycles, each cycle including 95°C for 10 seconds and 59°C for 50 seconds. The relative mRNA expression was assessed using the $2^{-\Delta\Delta CT}$ method (Livak *et al.*, 2001). Differences were considered significant when $P < 0.05$ (from one-way ANOVA). All statistical analyses were performed using the general linear model in R. Graphs could be visually displayed using the ggplot package in R (Gustavsson *et al.*, 2022).

RESULTS AND DISCUSSION

The Tibetan sheep is widely distributed on the Qinghai-Tibet Plateau and its adjacent areas. Using molecular markers of the FecB gene, a core population of highly prolific Tibetan sheep has been established, addressing the technical issue of low reproductive performance in ewes. However, there has been limited research on the reproductive performance of Tibetan sheep ewes. Contrary to this, our preliminary findings reveal the presence of the multi-lamb controlling FecB gene within the population, exerting a significant influence on lambing numbers (Qiao *et al.*, 2017; Qiao *et al.*, 2018; Sun *et al.*, 2023). Guided by

Table 1: mRNA primers for reverse-transcriptase quantitative PCR for ovarian tissues from different litter size.

Gene name	Accession numbers		Primer sequence (5'→3')	Product size (bp)	Annealing temperature (°C)
COLIA1	NM_000088.1	Forward	AGTGGTTTGGATGGTGCCAA	170	60
		Reverse	GCACCATCATTTCCACGAGC		
COLIA2	XM_004010473	Forward	TGGCAGTTTTGGTGGTATGC	292	60
		Reverse	TTCGTGAACAGCATCGGTGA		
TGFB2	XM_004013601.4	Forward	ACCCTCGGAAAATGCCATCC	224	60
		Reverse	CTCCTCCCCGTGTTTAGCAG		
IGFBP4	NM_001134302.1	Forward	AAGATGAAGGTCATCGGGGC	251	60
		Reverse	GGAAGCTTCACTCCCGTCTT		
		Reverse	CCATGTCTCGGTAGCACTGG		
CTSB	NM_001308587.1	Forward	GGCAACTCCTGGAACACTGA	118	60
		Reverse	AGTACTGATGAGTGACGGC		
PDGFC	XM_027956338.2	Forward	CACAGCCCCAAGGTTTCTCA	199	60
		Reverse	CACACCAGCGCCCTAAAATG		
SBSN	XM_027977407.2	Forward	TCAACAATGGCGTCGGACAT	289	60
		Reverse	TCACTTCCCACCTGGTTGAC		
AK5	XM_042249944.1	Forward	TGTTTGAAGCCTGAAGACCC	180	60
		Reverse	GTCCGTGTCACTTTCATTTCTT		
AK8	XM_042245507.1	Forward	CTTGCCAAACACCTGAGCAC	185	60
		Reverse	AGAATCCAGCCCCTCCTGAT		
IGFBP3	XM_027968353.2	Forward	CGACTCCAAATCCCACCTCG	86	60
		Reverse	AGTCAACCTTGAGCGCTGG		
AMT	XM_027957636.2	Forward	GCTCTGTTGAGAAACCCCGA	286	60
		Reverse	CCTCTGGACTCAGGATGGGA		
PDK4	XM_004007738.5	Forward	TGGTGTTCCTCTGAGAGTCA	109	60
		Reverse	GTAACCAAAACAGCCAGCG		
ACTB	NM_001009784	Forward	CCTGCGGCATTACAGAA	134	60
		Reverse	GCGGATGTCGACGTCACA		

these insights, ten single-lamb and ten twin-lamb Tibetan ewes were meticulously selected for the collection of ovarian tissues, subsequently subjected to transcriptome and lncRNA sequencing.

Overall assessment of sequencing data

Through high-throughput sequencing of 20 individuals, a total of 120 gigabytes of data were obtained, with an average data volume of approximately 6 gigabytes per individual (Table 2). The sequencing quality is good. Subsequently, these clean reads were aligned to the sheep reference genome (oar3.1, NCBI Accession: GCA_000298735.1) and the statistical results are shown in Table 3. From the table, it can be observed that the alignment efficiency ranged from 90.51% to 95.62%, indicating good sequencing quality with no data loss. The unique matching efficiency also ranged from 87.49% to 93.55%. In this study, a total of 20 Tibetan sheep samples' transcriptome and lncRNA data were obtained. After quality control, the alignment efficiency was all above 90%, indicating extremely high sequencing quality. This provides a reliable data foundation for the subsequent screening and analysis of differentially expressed genes and lncRNAs. In this sequencing, the assembly of transcripts was performed using StringTie software, followed by comparison with the genome annotation file using gffcompare software to identify new transcripts. In the end, 42,934 new transcripts were discovered. These 42,934 new transcripts will serve as a candidate dataset for the subsequent prediction of some lncRNAs.

lncRNA identification

Based on the results predicted by three software programs, the Venn diagram (Fig 1) intuitively displays the numbers of noncoding transcripts predicted by each method, including common and unique ones. In the end, a total of 7,041 lncRNAs predicted in this analysis will be used as a candidate lncRNA dataset for further analysis, while the remaining 35,893 transcripts belong to protein-coding genes.

Analysis of mRNA and lncRNA features

Through feature analysis, we found that mRNA and lncRNA in this sequencing have differences in Exon_Number, ORF_Length, Transcript_Length, Transcript_Number and Compare_Expression. The distribution of the number of exons is shown in Fig 2a. We can see that the number of exons in mRNA is significantly greater than in lncRNA. The number of mRNAs with 2-16 exons is all above 2000, while the number of lncRNAs with 2-5 exons is relatively low. By examining the expression level distribution plot (Fig 2b), it can be observed that mRNA has a higher expression level compared to lncRNA. This is because mRNA is responsible for encoding proteins, which are crucial for the normal metabolism and development of organisms, leading to generally high expression levels. In contrast, lncRNA does not encode proteins, resulting in relatively lower expression levels. Comparing the lengths of open reading frames (ORF) as shown in Fig 2c, we can observe significant differences between mRNA and lncRNA

Table 2: Overview of sequencing data.

Sample ID	Clean reads (M) pair-end reads (M)	Clean bases (G)	Q20 (%)	Q30 (%)	GC (%)	Read length (bp)	Effective reads (M)
ML1	82.4536M	12.3680G	97.4	92.58	49.07	150	82.4533 (100.00%)
ML10	91.1177M	13.6676G	97.95	93.55	45.17	150	91.1176 (100.00%)
ML2	97.6820M	14.6523G	97.03	91.71	48.79	150	97.6813 (100.00%)
ML3	101.8451M	15.2768G	97.36	92.53	48.49	150	101.8449 (100.00%)
ML4	100.4161M	15.0624G	97.41	92.49	48.24	150	100.4160 (100.00%)
ML5	86.5301M	12.9795G	96.92	91.4	47.98	150	86.5294 (100.00%)
ML6	99.5202M	14.9280G	97.08	91.72	49.14	150	99.5199 (100.00%)
ML7	103.3107M	15.4966G	97.47	92.6	48.6	150	103.3104 (100.00%)
ML8	98.9434M	14.8415G	97.46	92.44	48.96	150	98.9430 (100.00%)
ML9	78.4072M	11.7611G	97.19	91.94	47.48	150	78.4071 (100.00%)
SL1	86.9242M	13.0386G	96.77	91.35	50.89	150	86.9240 (100.00%)
SL10	99.9592M	14.9939G	97.28	92.21	47.97	150	99.9589 (100.00%)
SL2	79.7140M	11.9571G	97.34	92.4	48.92	150	79.7139 (100.00%)
SL3	103.2831M	15.4925G	97.49	92.75	46.93	150	103.2829 (100.00%)
SL4	98.8611M	14.8292G	97.31	92.33	48.92	150	98.8609 (100.00%)
SL5	79.2243M	11.8836G	96.72	90.88	48.39	150	79.2241 (100.00%)
SL6	102.5764M	15.3865G	97.19	92.11	48.92	150	102.5762 (100.00%)
SL7	84.8939M	12.7341G	96.9	91.43	48.31	150	84.8938 (100.00%)
SL8	89.3595M	13.4039G	96.87	91.31	47.7	150	89.3594 (100.00%)
SL9	97.4390M	14.6159G	97.14	91.96	47.42	150	97.4389 (100.00%)

Table 3: Genome alignment statistics.

Sample	Total reads	Total mapped	Reads map To '+'	Reads map To '-'	Splice mapped	Non-splice mapped	Unique mapped	Multiple mapped
ML1	82453326	76542444 (92.83%)	36543906 (44.32%)	39995538 (48.51%)	15724266 (19.07%)	60818178 (73.76%)	74419790 (90.26%)	2122654 (2.57%)
ML10	91117608	87124080 (95.62%)	41198828 (45.22%)	45925252 (50.40%)	13146988 (14.43%)	73977092 (81.19%)	85242098 (93.55%)	1881982 (2.07%)
ML2	97681336	90458992 (92.61%)	43197068 (44.22%)	47261924 (48.38%)	15403071 (15.77%)	75055921 (76.84%)	88465472 (90.57%)	1993520 (2.04%)
ML3	101844948	93897128 (92.20%)	44978872 (44.16%)	48918256 (48.03%)	17181989 (16.87%)	76715139 (75.33%)	91565138 (89.91%)	2331990 (2.29%)
ML4	100416018	93698544 (93.31%)	43938886 (43.76%)	49759658 (49.55%)	15373802 (15.31%)	78324742 (78.00%)	91608446 (91.23%)	2090098 (2.08%)
ML5	86529368	80437900 (92.96%)	38166100 (44.11%)	42271800 (48.85%)	13303028 (15.37%)	67134872 (77.59%)	78699824 (90.95%)	1738076 (2.01%)
ML6	99519880	91936708 (92.38%)	42820454 (43.03%)	49116254 (49.35%)	16191921 (16.27%)	75744787 (76.11%)	89652120 (90.08%)	2284588 (2.30%)
ML7	103310404	95859354 (92.79%)	45312334 (43.86%)	50547020 (48.93%)	16131764 (15.61%)	79727590 (77.17%)	93715736 (90.71%)	2143618 (2.07%)
ML8	98943006	92710770 (93.70%)	43306822 (43.77%)	49403948 (49.93%)	18168789 (18.36%)	74541981 (75.34%)	90465454 (91.43%)	2245316 (2.27%)
ML9	78407116	73166618 (93.32%)	34992732 (44.63%)	38173886 (48.69%)	12313914 (15.71%)	60852704 (77.61%)	71589098 (91.30%)	1577520 (2.01%)
SL1	86924028	78671994 (90.51%)	36666578 (42.18%)	42005416 (48.32%)	20298300 (23.35%)	58373694 (67.15%)	76050426 (87.49%)	2621568 (3.02%)
SL10	99958916	92933804 (92.97%)	43994556 (44.01%)	48939248 (48.96%)	15222523 (15.23%)	77711281 (77.74%)	90794094 (90.83%)	2139710 (2.14%)
SL2	79713906	73763650 (92.54%)	30656796 (38.46%)	43106854 (54.08%)	17686351 (22.19%)	56077299 (70.35%)	70769370 (88.78%)	2994280 (3.76%)
SL3	103282898	96271936 (93.21%)	45900542 (44.44%)	50371394 (48.77%)	15980322 (15.47%)	80291614 (77.74%)	93887076 (90.90%)	2384860 (2.31%)
SL4	98860940	92228506 (93.29%)	43146934 (43.64%)	49081572 (49.65%)	17484646 (17.69%)	74743860 (75.61%)	89908532 (90.94%)	2319974 (2.35%)
SL5	79224148	73528570 (92.81%)	35051224 (44.24%)	38477346 (48.57%)	12648010 (15.96%)	60880560 (76.85%)	71880348 (90.73%)	1648222 (2.08%)
SL6	102576206	95303720 (92.91%)	44402380 (43.29%)	50901340 (49.62%)	18851171 (18.38%)	76452549 (74.53%)	92788354 (90.46%)	2515366 (2.45%)
SL7	84893806	78374444 (92.32%)	36968820 (43.55%)	41405624 (48.77%)	13037082 (15.36%)	65337362 (76.96%)	76616648 (90.25%)	1757796 (2.07%)
SL8	89359450	83216380 (93.13%)	39999720 (44.76%)	43216660 (48.36%)	14424554 (16.14%)	68791826 (76.98%)	81301910 (90.98%)	1914470 (2.14%)
SL9	97438884	90675584 (93.06%)	43333136 (44.47%)	47342448 (48.59%)	14818393 (15.21%)	75857191 (77.85%)	88546122 (90.87%)	2129462 (2.19%)

in ORF length. For both novel_lncRNA and annotated_lncRNA, there is almost no difference in their ORF lengths. In this study, the ORF lengths of mRNA are concentrated between 100-2000 nt, while the ORF lengths of lncRNA are relatively shorter, primarily concentrated between 100-500 nt. Transcript length comparison is shown in Fig 2d. We can observe certain differences between mRNA and lncRNA in terms of transcript length. mRNA lengths are mainly concentrated above 3000 nt, while lncRNA lengths, aside

from the distribution ≥ 3000 nt, are primarily concentrated in the range of 200-1200 nt. Transcript number comparison is shown in Fig 3. We can observe significant differences between mRNA and lncRNA in terms of transcript numbers. The number of mRNA transcripts, including known and newly identified mRNA, is greater than that of lncRNA transcripts, including known and newly predicted lncRNA.

This observation aligns cohesively with the findings from the feature analysis of mRNA and lncRNA in other species, such as mice, pigs, cattle, sheep and goats. Exon quantity and expression levels in mRNA are higher than in lncRNA, attributed to their fundamental biological roles. mRNAs, with 2 to 16 exons, are essential for protein coding, vital for organismal metabolism and development, necessitating higher expression levels. Conversely, lncRNAs, predominantly having 2 to 5 exons and lacking protein-coding function, exhibit lower expression levels (Guo *et al.*, 2017). This distinction underscores the diverse functional imperatives of mRNA and lncRNA in biological processes (Maialen *et al.*, 2021). Similar conclusions are echoed in studies involving other mammalian species such as mice (Sarropoulos *et al.*, 2019; Trovero *et al.*, 2020), pigs (Wang *et al.*, 2016; Yu *et al.*, 2017) and chickens (Liu *et al.*, 2017). For instance, Wang *et al.* (2016) utilized RNA-Seq to analyze lncRNA and mRNA expressions in the porcine endometrium, finding mRNA expression levels to be 2 to 3 times higher than lncRNAs, indicating a universal superiority of mRNA expression. This study highlights

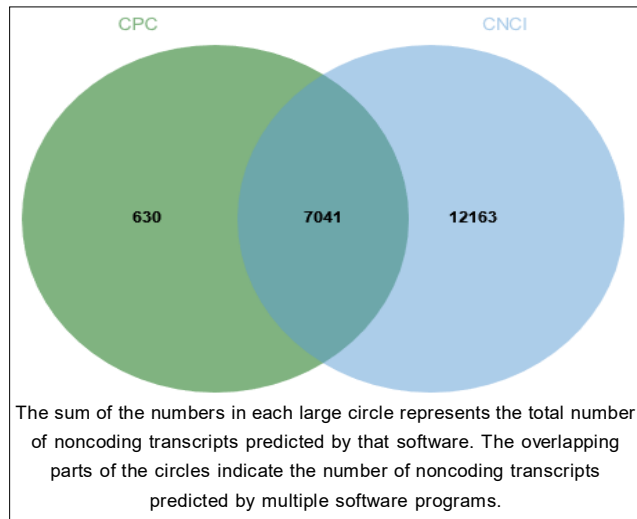


Fig 1: Venn diagram of the filtering results.

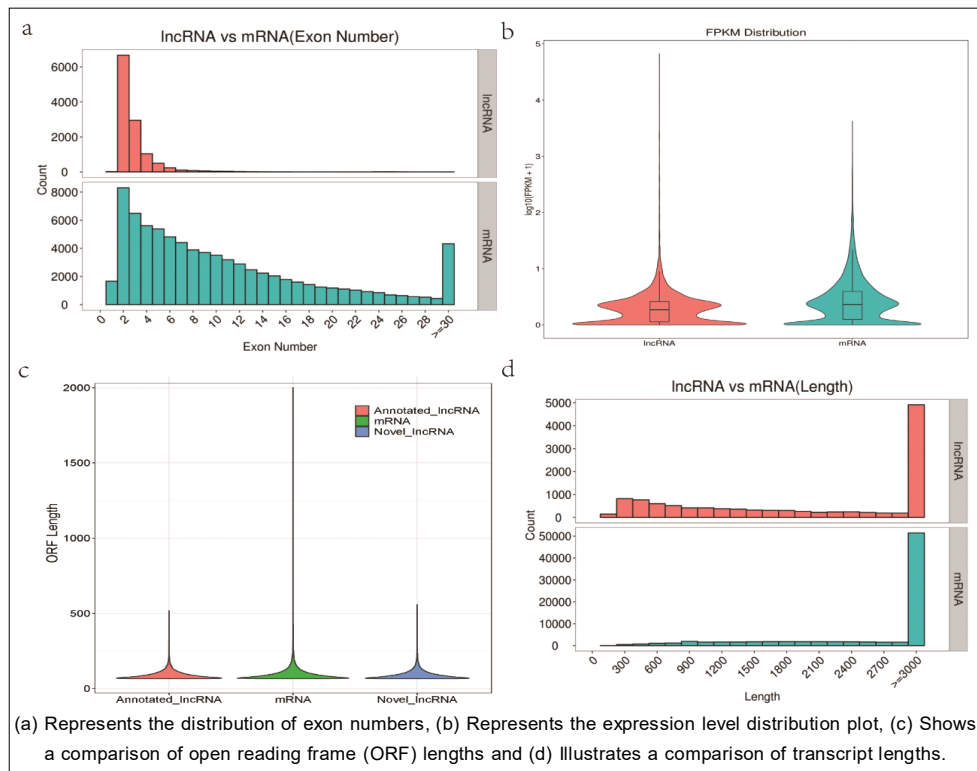


Fig 2: mRNA and lncRNA feature analysis.

notable disparities in ORF lengths between mRNA and lncRNA, with mRNA ORFs predominantly ranging from 100 to 2000 nt and lncRNA ORFs being considerably shorter. Notably, lncRNA ORFs in Tibetan sheep are longer compared to other species like mice and pigs (Xu *et al.*, 2023), suggesting species-specific roles of lncRNA in mammalian reproductive performance. Current research corroborates that lncRNA, despite shorter ORFs, can influence various molecular mechanisms such as RNA splicing and editing (Antonov *et al.*, 2019). Thus, a broader evaluation beyond ORF lengths is crucial for a comprehensive understanding of lncRNA's biological functions and mechanisms in diverse experimental contexts. Discrepancies in transcript lengths of mRNA and lncRNA have been observed, with mRNA predominantly exceeding 3000 nucleotides (nt) due to the necessity for ample exons and UTR sequences for functional translation and regulation. Conversely, lncRNA lengths are more concise, primarily ranging between 200-1200 nt, with exceptions occasionally exceeding 3000 nt (Chen *et al.*, 2018). Such patterns are consistent across various studies, including those involving ovarian tissues in fish (Hu *et al.*, 2023), pigs (Tang *et al.*, 2017). Notably, longer lncRNAs, albeit rare, appear to uniquely influence the reproductive performance of Tibetan sheep, underscoring their specialized roles in reproductive biology. A significant discrepancy exists between the quantities of mRNA and lncRNA transcripts, with mRNA being more abundant, constituting 80-90% of total transcripts in identical tissues of humans and mice, as per RNA-Seq analyses (Engreitz *et al.*, 2016). Conversely, lncRNA accounts for 10-20% of the total transcript count (Munsky *et al.*, 2012). This distribution aligns with findings from this study on Tibetan sheep. The differences in transcript quantities may be attributed to their distinct roles in gene regulation: mRNA primarily encodes proteins, while lncRNA modulates gene expression at the transcriptional level, suggesting that the observed variances are functionally consequential.

Differential gene expression analysis

In our differential gene expression analysis, we identified a total of 817 genes that were differentially expressed between twin-bearing and single-bearing ewes, with 495 genes being upregulated and 322 downregulated (Fig 4).

Some typical genes include COL1A1, COL1A2, COL14A1, COL5A3, COL6A3, COL13A1, LOC443512, LAMA4, EIF3E, FLII, CCL2, TMEM165, MKX, NLGN3, VMAC, CXCL13, LOC101111035, SBSN, BAALC, NR2C2AP, PDE4A, AK5, AK6, AK8, RLN3, SSR3, PDGFC, CCDC137, TGFB2, IGFBP3, IGFBP4, BMPRI3, GDF9, BMP15, DRB3, EPHB2, PDK4, DNAJB11, HKDC1, MOV10L1, CTSB, CTSS, TYROBP and more. The sequencing analysis revealed that, relative to the low-yield group, COL1A1, COL1A2, TGFB2, IGFBP4, CTSB and PDGFC were upregulated in the high-yield Tibetan sheep ovarian tissue, while SBSN, AK5, AK8, IGFBP3, AMT and PDK4 showed significant downregulation. To confirm the gene expression patterns identified by RNA sequencing, we selected a total of 12 genes, including 6 up-regulated genes (COL1A1, COL1A2, TGFB2, IGFBP4, CTSB, PDGFC) and 6 down-regulated genes (SBSN, AK5, AK8, IGFBP3, AMT, PDK4) and performed real-time quantitative PCR (qPCR) validation using ovarian samples from different litter sizes. The results showed that the expression patterns of these 12 genes were consistent with the RNA sequencing results (Fig 5 and Fig 6), indicating the high reliability of our RNA sequencing results and analysis methods.

Many candidate genes identified in this sequencing have also been found in other mammals and are implicated in reproductive efficiency, relating to embryo and follicle development. For instance, a study by (Zhang *et al.*, 2015) compared gene expressions in pigs with different fecundities, finding significant variations, with 152 differentially expressed genes, including 75 upregulated and 77 downregulated genes, especially in transcription factors related to reproductive biology. These findings enhance the understanding of molecular mechanisms in reproduction, offering valuable insights for breeding and farming strategies. Similarly, scholars conducted a differential gene expression analysis in ovarian tissues of pigs with varying fecundities using RNA sequencing (Fernandez *et al.*, 2011) and identified significant disparities in gene expression related to reproductive capacities, such as embryo and follicle development and hormone receptors. In a goats study, scholars (Ling *et al.*, 2014) identified 2201 differentially expressed genes (DEGs) in

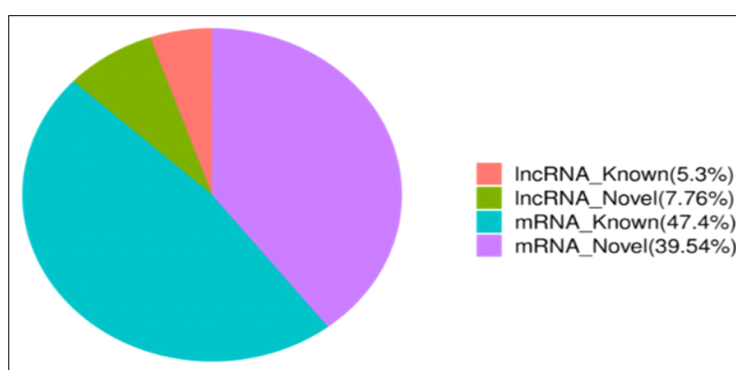


Fig 3: Relative proportions of different transcripts (mRNAs and lncRNAs) in the samples.

Anhui White Goats, with 1583 upregulated and 618 downregulated, associating many with cellular processes and components. Twelve genes were linked to high fecundity, enriching 11 KEGG pathways like ECM-receptor interaction, laying a foundation for understanding goat reproductive biology and improving breeding efficiency. In a Xinjiang sheep study, scholars (Chen *et al.*, 2015) found 303 DEGs related to reproductive processes, immunity and metabolism, providing genetic markers for reproductive performance, enhancing understanding of sheep biology and breeding and farming efficiency. These studies elucidate genetic influences on reproductive capacities in sheep and goats, offering significant references for enhancing breeding strategies and livestock

productivity. In a Hu sheep study (Feng *et al.*, 2018), scholars explored mRNAs and lncRNAs related to Hu sheep prolificacy through ovarian genome-wide analysis, comparing high prolific (HP) and low prolific (LP) groups during the follicular phase. They found higher plasma luteinizing hormone levels in the HP group and identified 76 differentially expressed mRNAs (DE-mRNAs) and 5 differentially expressed lncRNAs (DE-lncRNAs). We identified several genes, including EPHB2, PDK4, BMP1B, GDF9, BMP15 and IGFBP3, with potential roles in reproduction. EPHB2 is associated with testicular development and sperm formation, influencing cellular activities such as proliferation and migration in the male reproductive system. PDK4, essential in cellular energy

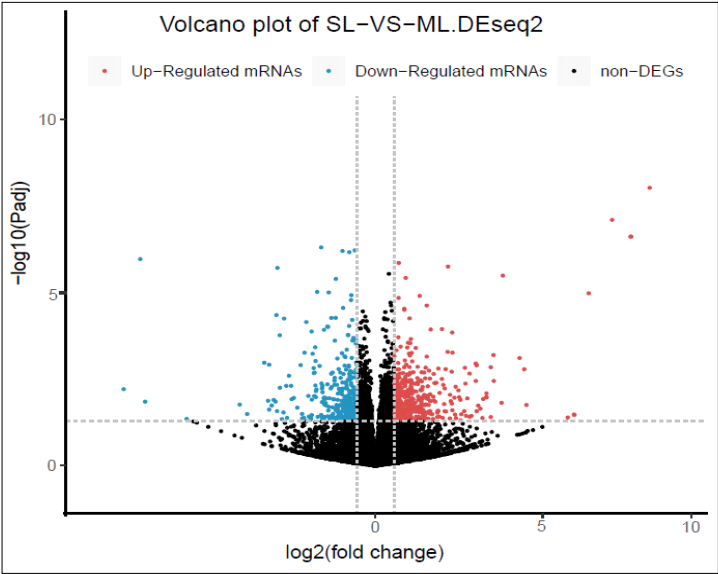


Fig 4: Volcano plot of differential expression mRNAs.

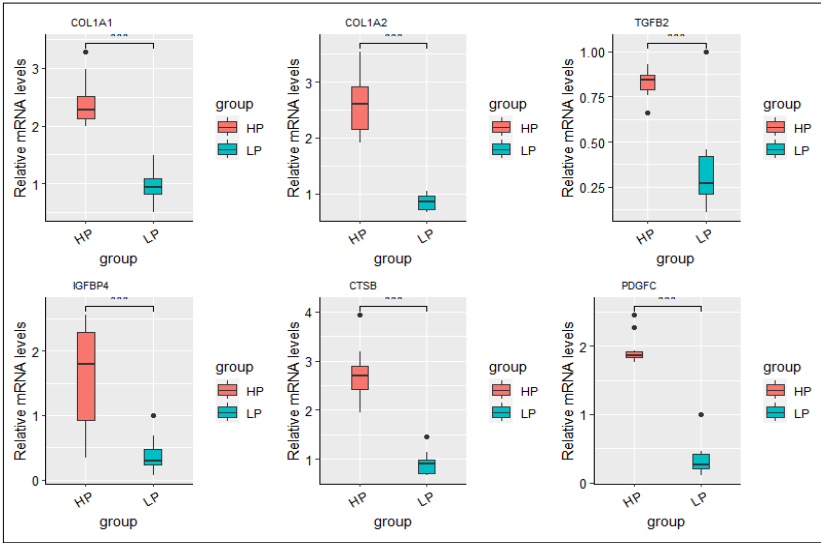


Fig 5: Quantitative expression profile of Up-regulated genes.

metabolism, indirectly influences reproductive processes by regulating energy consumption essential for reproductive functions like fertilization and pregnancy. Although direct evidence linking PDK4 to reproduction is lacking, its role in metabolic pathways suggests a potential impact on reproductive success and health. IGFBP3, interacting with insulin-like growth factors (IGFs), plays a crucial role in cell growth and survival and is involved in various reproductive processes such as gonadal development, follicular development and embryo growth, indicating its significance in fertility and reproductive health. In high-altitude environments, species' adaptability is closely linked to their physiological mechanisms. Studies have shown significant impacts of high-altitude environments on ovarian gene expression. For example, in a study on high-altitude Tibetan sheep (Zheng *et al.*, 2019), scholars used transcriptome sequencing and discovered significant differences in the expression of the *FecB* gene between multi-lamb and single-lamb individuals, which affects their reproductive capacity. In another study (Miao *et al.*, 2016) on large sheep breeds in high-altitude areas, scholars used RNA-seq technology to reveal significant differences in genes, such as *BMP15* and *GDF9*, between individuals with different reproductive capacities. This research highlighted the critical roles these genes play in regulating ovarian function and fertility, providing a molecular basis for improving reproductive performance in livestock in high-altitude areas. Similar to the results of the above two studies, our experimental results also found differences in the *FecB* (*BMP15*), *BMP15* and *GDF9* genes, which fully demonstrates that these three genes play important roles in the reproductive system of Tibetan sheep.

Analysis of differentially expressed lncRNAs

In the analysis of differentially expressed lncRNAs, 210 differentially expressed lncRNAs, including 91 up-regulated and 119 down-regulated, were identified in ewes producing twin lambs compared to those producing single lambs (Fig 7). This study aimed to elucidate the roles of lncRNAs in the genetic regulation of sheep prolificacy by comparing sequencing results of Tibetan sheep with varying lambing numbers to other mammals. Miao *et al.* (Miao *et al.*, 2016) conducted a genome-wide analysis, revealing significant correlated expression patterns of lncRNAs and mRNAs, emphasizing lncRNAs' regulatory roles in sheep prolificacy. Zheng *et al.* (2019) analyzed the transcriptome of high and low prolificacy sheep, identifying 57 differentially expressed (DE) lncRNAs and 298 DE mRNAs, enriched in hormone-related pathways and other reproductive functions, constructing a co-expression network involving genes like *SMAD2*, *NMB* and *EFNB3*. Boruah *et al.* (2021) identified significant differences in the expression of genes in Hu sheep at various developmental stages, uncovering 6,716 mRNAs and 1,972 lncRNAs enriched in pathways like TGF- β and PI3K-Akt, regulating key ovarian development genes such as *CTNNB1*, *CCNA2* and *CDK2*. Another 2,903 mRNAs and 636 lncRNAs were identified in different age comparisons, involved in pathways like PI3K-Akt and hormone-mediated oocyte maturation, regulating genes like *FGF7*, *PRLR* and *AMH*. These findings underscore the pivotal roles of these lncRNAs and mRNAs in understanding the molecular mechanisms of Tibetan sheep prolificacy.

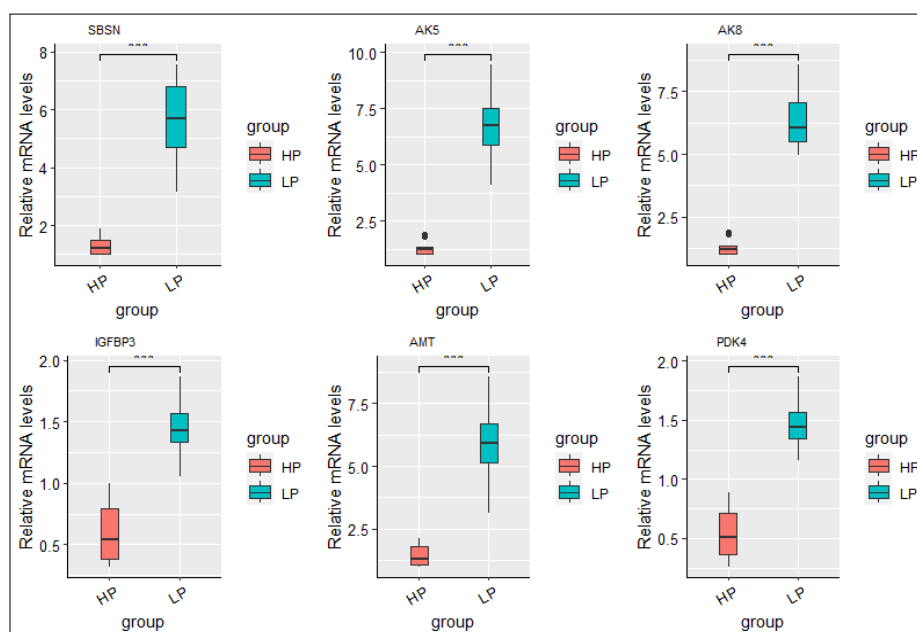


Fig 6: Quantitative expression profile of down-regulated genes.

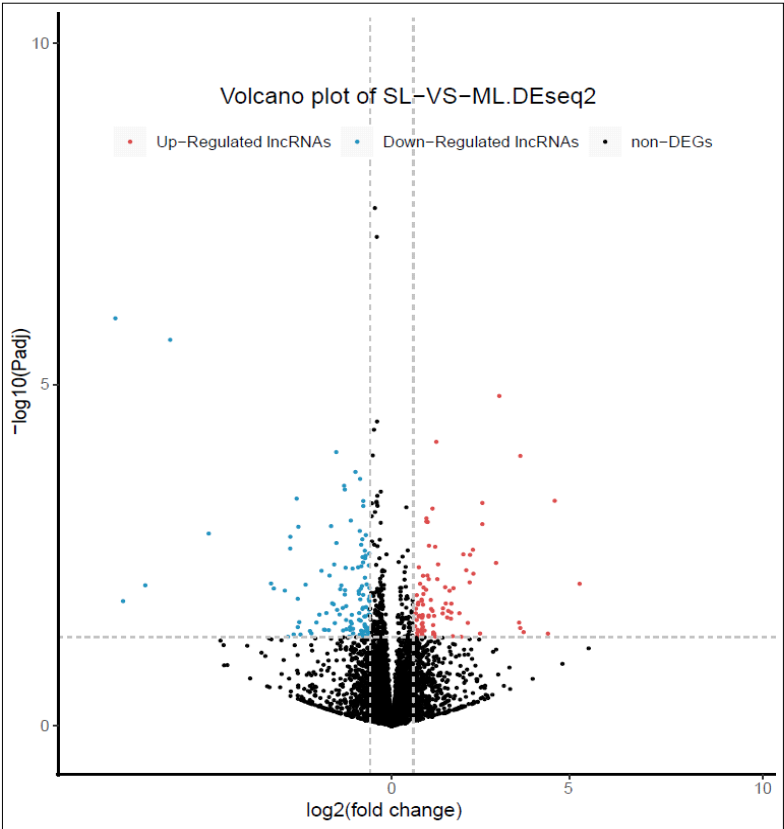


Fig 7: Volcano plot of differentially expressed lncRNAs.

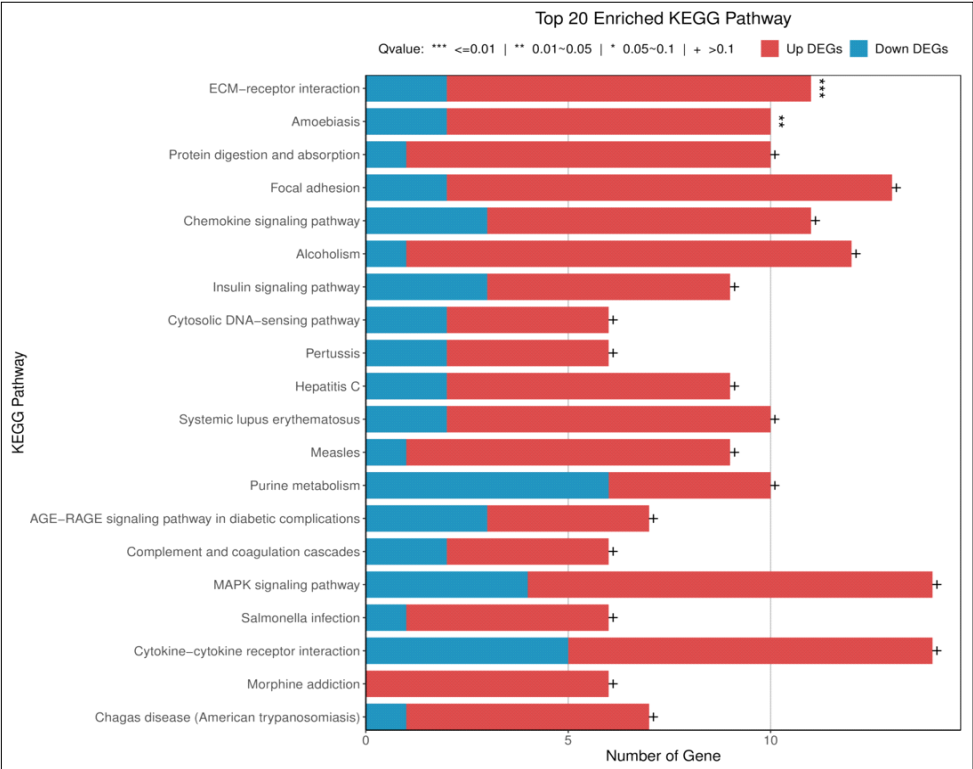


Fig 8: Pathway enrichment analysis of differentially expressed mRNAs.

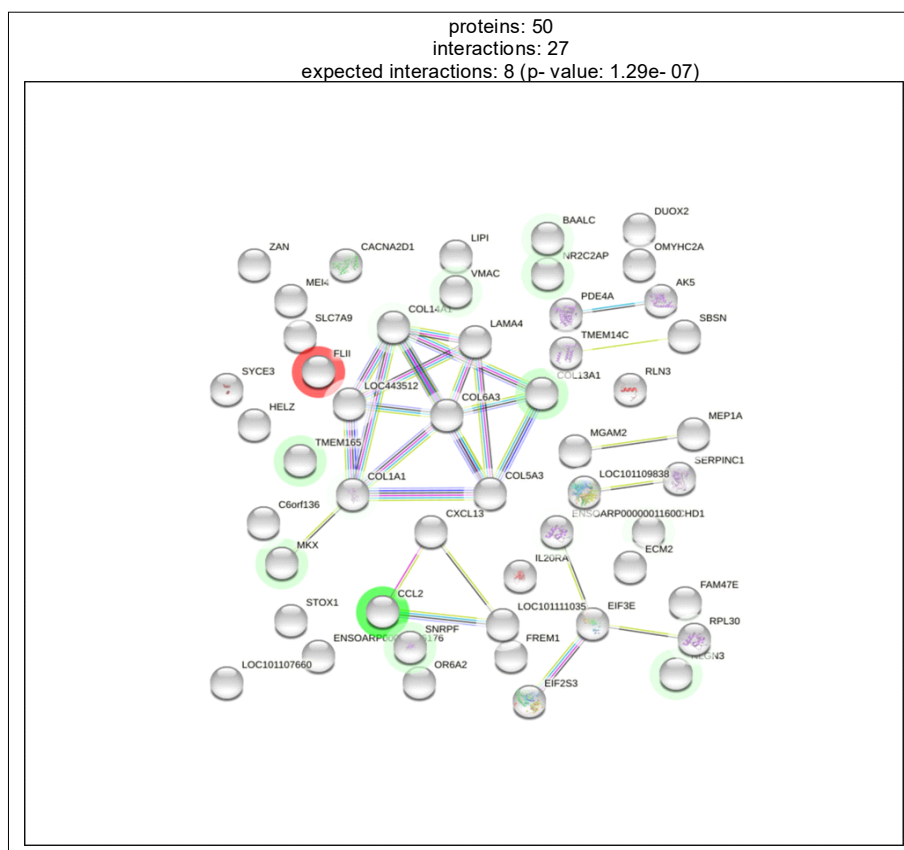


Fig 9: Protein-protein interaction network of differentially expressed genes.

KEGG functional enrichment analysis

As shown in Fig 8, KEGG enrichment analysis results indicate that the candidate genes are enriched in pathways such as the estrogen pathway, MAPK pathway and cholesterol synthesis pathway. By comparing the Tibetan sheep sequencing results with those of other mammals, we found that these enriched items and pathways have also been reported in other species. For example, in a study on goats (Xu *et al.*, 2019), KEGG enrichment analysis revealed that these genes also play crucial roles in the estrogen signaling pathway. In cattle (Xu *et al.*, 2019), the MAPK signaling pathway has been shown to be essential for regulating cell proliferation and differentiation. These findings suggest that the gene expression characteristics of Tibetan sheep share conserved key biological pathways with other mammals.

Protein-protein interaction network analysis

Based on the stringDB package within the R Bioconductor and Cytoscape software, we constructed the protein-protein interaction network for this experiment. As shown in Fig 9, we identified eight core genes, including COL1A1, COL14A1, COL5A3, COL6A3, COL13A1, LOC443512, LAMA4 and EIF3E. Additionally, we found related genes such as FLII, CCL2, TMEM165, MKX, NLGN3, VMAc, CXCL13, LOC10111035, SBSN, BAALC, NR2C2AP, PDE4A, AK5 and RLN3. This study focuses on several core

genes implicated in the regulation of litter size in Tibetan sheep. The COL1A1 gene, encoding collagen type I alpha 1 chain, is fundamental in animal development, influencing structures like bone and skin (Chen *et al.*, 2021). The eIF3e gene, part of the eukaryotic translation initiation factor 3 complex, is essential for embryonic development and cell proliferation (Daichi *et al.*, 2018). The CCL2 gene, encoding a cytokine, is involved in various processes such as immune responses and inflammation (Feng *et al.*, 2017; Nakamura *et al.*, 2020). The NLGN3 gene, associated with synaptic connections between neurons, influences brain development and cognitive abilities, with potential implications in hormone regulation and reproductive tract diseases, although its specific role in reproduction requires further exploration (Jamain *et al.*, 2003; El-Fishawy *et al.*, 2010). Lastly, the MKX gene is crucial in mammalian embryonic development (Otabe *et al.*, 2015). The identification of these genes in the experiment suggests their potential regulatory roles in the prolificacy of Tibetan sheep.

CONCLUSION

This study, through transcriptome sequencing of Tibetan sheep ovarian tissues, unveils potential genetic regulators of litter size, identifying 42,934 new transcripts, including 35,893 protein-coding and 7,041 long non-coding RNAs (lncRNAs). An average of 817 differentially expressed genes

(DEGs) were detected, encompassing key reproduction-related genes like COL1A1, BMP1B, GDF9, BMP15 and TGFB2. Additionally, 210 differentially expressed lncRNAs were identified, potentially influencing litter size by regulating associated genes. Functional analysis suggests these DEGs are involved in cellular functions, connections and metabolic pathways such as the estrogen pathway and MAPK, indicating a multifaceted genetic regulation of litter size in Tibetan sheep. These findings not only provide preliminary insights into the molecular mechanisms influencing litter size in Tibetan sheep but also offer valuable genetic markers and references for reproductive research in other livestock species. Furthermore, these discoveries hold significant importance for optimizing existing farming systems, aiding in the improvement of breeding efficiency and meeting the growing demand for meat and dairy products, thereby promoting sustainable development in animal husbandry.

Funding

This work was supported by China Agriculture Research System of MOF and MARA(CARS-39-35), the Youth Fund of Qinghai Provincial Science and Technology Department (2021-ZJ-978Q) and the Open Project of State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University(2023-ZZ-11).

Author contribution

Wu Sun designed the experiments and drafted the manuscript. Sequencing data processing and visualization was done by Wu Sun and Xiayang Jin. Shike Ma and Yuhong Ma carried out quantitative experiments.

Informed consent

Informed consent has been obtained from all individuals included in this study.

Data availability statement

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

Conflict of interest

Authors state no conflict of interest.

REFERENCES

- Abdoli, R., Zamani, P., Mirhoseini, S. *et al.* (2016). A review on prolificacy genes in sheep. *Reproduction in Domestic Animals*. 51: 631-637.
- Antonov, I.V., Mazurov, E., Borodovsky, M. *et al.* (2019). Prediction of lncRNAs and their interactions with nucleic acids: Benchmarking bioinformatics tools. *Brief Bioinform.* 20: 551-564.
- Boruah, P., Shabbir, S., Kulyar, F. *et al.* (2021). Genome-wide transcriptome profiling uncovers differential miRNAs and lncRNAs in ovaries of Hu sheep at different developmental stages. *Scientific Reports*. 11(1): 5865. doi: 10.1038/s41598-021-85245-y.
- Chang, W.H., Cui, Z.L., Wang, J.H. *et al.* (2021). Identification of potential disease biomarkers in the ovaries of dolang sheep from xinjiang using transcriptomics and bioinformatics approaches. *Indian Journal of Animal Research*. 55: 112-125. doi: 10.18805/ijar.B-1265.
- Chen, H.Y., Shen, H., Jia, B. *et al.* (2015). Differential gene expression in ovaries of qira black sheep and hetian sheep using RNA-Seq technique. *Plos One*. 10. e0120170. doi: 10.1371/journal.pone.0120170.
- Chen, L., Li, W.W., Li, D.Q. *et al.* (2018). Screening and analysis of differentially expressed genes of ovary tissue in Hu sheep at the estrous stage. *Heilongjiang Animal Husbandry and Veterinary Medicine*. 259: 37-41.
- Chen, L., Liu, K., Zhao, Z. *et al.* (2012). Identification of sheep ovary genes potentially associated with off-season reproduction. *J. Genet Genomics*. 39: 181-190.
- Chen, L., Zhang, Y.H., Pan, X. *et al.* (2018). Tissue expression difference between mRNAs and lncRNAs. *Int J. Mol Sci*. 19: 3416. doi: 10.3390/ijms19113416.
- Chen, Y., Yang, S., Lovisa, S. *et al.* (2021). Type-I collagen produced by distinct fibroblast lineages reveals specific function during embryogenesis and osteogenesis imperfecta. *Nature Communications*. 12. 7199. doi: 10.1038/s41467-021-27563-3.
- Daichi, S., Tomio, O., Saki, G.S. *et al.* (2018). Eukaryotic translation initiation factor 3 (eIF3) subunit e is essential for embryonic development and cell proliferation. *Febs Open Bio*. 8: 1188-1201.
- Damian, S., Gable, A.L., Nastou, K.C. *et al.* (2021). The string database in 2021: Customizable protein-protein networks and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*. 49: 605-612.
- El-Fishawy, P., State, M.W. (2010). The genetics of autism: Key issues, Recent findings and clinical implications. *Psychiatr Clin. North. Am.* 33: 83-105.
- Engreitz, J.M., Haines, J.E., Perez, E.M. *et al.* (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature*. 539: 452-455.
- Feng, C.Y., Wang, X., Liu, T.J. *et al.* (2017). Expression of CCL2 and its receptor in activation and migration of microglia and monocytes induced by photoreceptor apoptosis. *Molecular Vision*. 23: 765-777.
- Feng, X., Li, F., Wang, F. *et al.* (2018). Genome-wide differential expression profiling of mRNAs and lncRNAs associated with prolificacy in Hu sheep. *Biosci. Rep.* 38: 20171350. doi: 10.1042/BSR20171350.
- Fernandez, A., Munoz, M., Fernandez, A. *et al.* (2011). Differential gene expression in ovaries of pregnant pigs with high and low prolificacy levels and identification of candidate genes for litter size. *Biol. Reprod.* 84: 299-307.
- Guo, C.J., Xiao, X., Sheng, L. *et al.* (2017). RNA sequencing and bioinformatics analysis implicate the regulatory role of a long noncoding RNA-mRNA network in hepatic stellate cell activation. *Cellular Physiology and Biochemistry*. 42: 2030-2042.
- Gustavsson, E.K., Zhang, D. (2022). ggtranscript: An R package for the visualization and interpretation of transcript isoforms using ggplot 2. *Bioinformatics*. 38: 3844-3846.

- Hu, Q., Xia, X., Lian, Z. *et al.* (2023). Regulatory mechanism of LncRNAs in gonadal differentiation of hermaphroditic fish, *Monopterus albus*. *Bid Sex Differ.* 14: 74. doi: 10.1186/s13293-023-00559-y.
- Jamain, S., Quach, H., Betancur, C. *et al.* (2003). Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nature Genetics.* 34: 27-29.
- Jones, H.E., Wilson, P.B. (2022). Progress and opportunities through use of genomics in animal production. *Trends in Genetics.* 38: 1228-1252.
- Kanehisa, M., Sato, Y., Furumichi, M. *et al.* (2019). New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* 47: D590-d595.
- Kim, D., Langmead, B., Salzberg, S.L. *et al.* (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods.* 12: 357-360.
- Kong, L., Yong, Z., Ye, Z.Q. *et al.* (2007). CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Research.* 35: 345-349.
- Ling, Y.H., Xiang, H., Li, Y.S. *et al.* (2014). Exploring differentially expressed genes in the ovaries of uniparous and multiparous goats using the RNA-Seq (Quantification) method. *Gene.* 550: 148-153.
- Liu, Y., Li, M., Bo, X. *et al.* (2018). Systematic analysis of long non-coding RNAs and mRNAs in the ovaries of duroc pigs during different follicular stages using RNA sequencing. *Int. J. Mol. Sci.* 19: 1722. doi: 10.3390/ijms19061722.
- Liu, Y., Sun, Y., Li, Y. *et al.* (2017). Analyses of long non-coding RNA and mRNA profiling using RNA sequencing in chicken testis with extreme sperm motility. *Scientific Reports.* 7: 9055. doi: 10.1038/s41598-017-08738-9.
- Livak, K.J., Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻[Delta Delta C(T)] method. *Methods.* 25: 402-408.
- Love, M.I., Huber, W. and Anders, S. *et al.* (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology.* 15: 550. doi: 10.1186/s13059-014-0550-8.
- Maialen, S.D., Itziar, G.M., Ane, O.G. *et al.* (2021). The role of lncRNAs in gene expression regulation through mRNA stabilization. *Noncoding RNA.* 7: 3. doi: 10.3390/ncrna7010003.
- Miao, X., Luo, Q., Zhao, H. *et al.* (2016). Ovarian transcriptomic study reveals the differential regulation of miRNAs and lncRNAs related to fecundity in different sheep. *Scientific Reports.* 6: 35299. doi: 10.1038/srep35299.
- Mistry, J., Bateman, A., Finn, R. *et al.* (2007). Predicting active site residue annotations in the pfam database. *BMC Bioinformatics.* 8: 298. <https://doi.org/10.1186/1471-2105-8-298>.
- Munsky, B., Neuert, G., van Oudenaarden, A. *et al.* (2012). Using gene expression noise to understand gene regulation. *Science.* 336: 183-187.
- Nakamura, Y., Aihara, R., Iwata, H. *et al.* (2020). IL1B triggers inflammatory cytokine production in bovine oviduct epithelial cells and induces neutrophil accumulation via CCL2. *American Journal of Reproductive Immunology.* 85: e13365. <https://doi.org/10.1111/aji.13365>.
- Otake, K., Nakahara, H., Hasegawa, A. *et al.* (2015). Transcription factor Mohawk controls tenogenic differentiation of bone marrow mesenchymal stem cells *in vitro* and *in vivo*. *Journal of Orthopaedic Research.* 33: 1-8.
- Pan, L., Gong, W., Zhou, Y. *et al.* (2014). A comprehensive transcriptomic analysis of infant and adult mouse ovary. *Genomics, Proteomics and Bioinformatics.* 12: 239-248.
- Pertea, M., Kim, D., Pertea, G.M. *et al.* (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, stringTie and ballgown. *Nature Protocols.* 11: 1650-1667.
- Pertea, M., Pertea, G.M., Antonescu, C.M. *et al.* (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology.* 33: 290-295.
- Qiao, G.Y., Qiao, H.S., Ma, S.K. *et al.* (2017). Multiparous reproductive performance analysis of plateau Tibetan sheep. *Chinese Journal of Animal Husbandry and Veterinary Medicine.* 17-18.
- Qiao, H.S., Yang, B.H., Yue, Y.J. *et al.* (2018). Detection and analysis of BMPR-IB gene polymorphism in plateau Tibetan sheep ram population. *Heilongjiang Animal Science and Veterinary Medicine.* 1: 56-58.
- Sarropoulos, I., Marin, R., Cardoso-Moreira, M. *et al.* (2019). Developmental dynamics of lncRNAs across mammalian organs and species. *Nature.* 571: 1-5.
- Shannon, P. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research.* 13: 2498-2504.
- Song, W., Li, J., Sun, A. *et al.* (2022). Transcriptomic analysis of differentially expressed genes between uterus and ovary in Xiangxi cattle. *Pakistan Journal of Agricultural Sciences.* 59: 898-909.
- Soygur, B., Laird, D.J. (2021). Ovary development: Insights from a three-dimensional imaging revolution. *Front. Cell. Dev. Biol.* 9, 698315. doi: 10.3389/fcell.2021.698315.
- Sun, L., Luo, H., Bu, D. *et al.* (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Research.* 41: e166. doi: 10.1093/nar/gkt646.
- Sun, W., Ma, S.K., Ma, Y.H. *et al.* (2023). Single-nucleotide polymorphism scanning of bone morphogenetic protein receptor gene and its correlation with the prolificacy of plateau Tibetan sheep. *Indian Journal of Animal Research.* 57: 556-571. doi: 10.18805/IJAR.BF-1616.
- Szklarczyk, D., Kirsch, R., Koutrouli, M. *et al.* (2023). The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 51: D638-d646.
- Tang, Z., Wu, Y., Yang, Y. *et al.* (2017). Comprehensive analysis of long non-coding RNAs highlights their spatio-temporal expression patterns and evolutionary conservation in *Sus scrofa*. *Sci. Rep.* 7: 43166. doi: 10.1038/srep43166.
- Tian, X., Sun, H.X., Wang, Y.J. *et al.* (2009). Genetic polymorphism of BMPR-IB gene in three sheep populations and its effects on litter size. *Journal of Northwest A and F University (Natural Science Edition).* 2: 11-24.
- Trovero, M.F., Rodríguez-Casuriaga, R., Romeo, C. *et al.* (2020). Revealing stage-specific expression patterns of long noncoding RNAs along mouse spermatogenesis. *Rna Biology.* 17: 350-365.

- Wang, Y., Xue, S., Liu, X. *et al.* (2016). Analyses of long non-coding RNA and mRNA profiling using RNA sequencing during the pre-implantation phases in pig endometrium. *Sci. Rep.* 6: 20238. doi: 10.1038/srep20238.
- Xu, Q., Luo, Y., Chao, Z., *et al.* (2023). Integrated analysis of transcriptome expression profiles reveals miRNA-326-NKX3.2-regulated porcine chondrocyte differentiation. *Int. J. Mol. Sci.* 24: 7257. doi: 10.3390/ijms24087257.
- Xu, X., Ji, H., Jin, X. *et al.* (2019). Using pan RNA-seq analysis to reveal the ubiquitous existence of 5' and 3' end small RNAs. *Frontiers in Genetics.* 10: 105. doi: 10.3389/fgene.2019.00105.
- Yan, X., Ma, L., Yang, M. *et al.* (2020). Identification and characterization of long non-coding RNA (lncRNA) in the developing seeds of *Jatropha curcas*. *Sci. Rep.* 10: 10395. doi: 10.1038/s41598-020-67410-x.
- Yang, F., Wang, M., Zhang, B. *et al.* (2018). Identification of new progesterone-associated networks in mammalian ovulation using bioinformatics. *BMC Systems Biology.* 12: 36. doi: 10.1186/s12918-018-0577-7.
- Yu, L., Tai, L., Zhang, L. *et al.* (2017). Comparative analyses of long non-coding RNA in lean and obese pig. *Oncotarget.* 8: 41440-41450.
- Zhang, X., Huang, L., Wu, T. *et al.* (2015). Transcriptomic analysis of ovaries from pigs with high and low litter size. *PLoS One.* 10: e0139514. <https://doi.org/10.1371/journal.pone.0139514>.
- Zheng, J., Wang, Z., Yang, H. *et al.* (2019). Pituitary transcriptomic study reveals the differential regulation of lncRNAs and mRNAs related to prolificacy in different FecB genotyping sheep. *Genes.* 10: 145-158.