Pakistan J. Zool., vol. 56(4), pp 1743-1755, 2024 **DOI**: https://dx.doi.org/10.17582/journal.pjz/20220716100734

Transcriptomic Comparative Analysis of Two Breeds of Mongolian Sheep at 16-day Embryos

Hong Su[1,2], Lu Chen[1,3], Wenrui Guo[1], Daqing Wang[1,3], Aolei Dou[1,2], Jie Su[1], Yanyan Yang[4], Ying Tian[4], Tingyi He[4], Caiyun Wang[1], Chenguang Du[1], Haijun Li[1], Xihe Li[5], Guifang Cao[1]*, Yongli Song[5]* and Fuxiang Bao[3]*

[1]*College of Veterinary Medicine, Inner Mongolia Agricultural University, China*
[2]*Animal Embryo and Developmental Engineering Key Laboratory of Higher Education Institutions of Inner Mongolia Autonomous Region, China*
[3]*Inner Mongolia Autonomous Region Key Laboratory of Basic Veterinary Medicine, China China*
[4]*Institute of Animal Husbandry, Inner Mongolia Academy of Agriculture and Animal Husbandry*
[5]*College of Life Sciences, Inner Mongolia University, China*

Hong Su, Lu Chen, and Wenrui Guo contributed equally to this work.

Article Information
Received 16 July 2022
Revised 27 September 2022
Accepted 26 October 2022
Available online 01 June 2023
(early access)
Published 08 June 2024

Authors' Contribution
Conceptualization, HS. Methodology, HS. Software, HS, LC. Validation, HS, JS. Formal analysis, YY. Investigation, YT. Resources, GC. Data curation, HS, CD. Writing original draft preparation, HS, HL, TH. Writing review and editing, XL. Visualization, HS, YS. Project administration, DW, AD. Funding acquisition, CW.

Key words
16-day embryos, Hulunbuir short-tailed sheep, Ujumqin sheep, mRNA sequencing, Organogenesis

ABSTRACT

Early embryonic development determines fetal and adult development. However, little is known about the specific development of the early Mongolian sheep embryo. We aim to analyze the differentially expressed genes (DEGs) between 16-day embryos of two breeds of Mongolian sheep, namely, Hulunbuir short-tailed sheep and Ujumqin sheep, by mRNA sequencing. A total of 2152 DEGs were identified. Significantly DEGs include *WNT5A, WNT2, AXIN2, BMP5, TGFB2*, and *SMAD9*, which are involved in multiple signaling pathways, and genes involved in digestion and absorption and cellular differentiation, such as *DPP4, ACE2, PLPP1, CD22, CD44, TEC,* and *JAK1*. In the pathway analysis, the development of Hulunbuir short-tailed sheep 16-day embryo H(E16) highlighted pathways primarily associated with cellular interactions and organogenesis. These pathways are critical for transcriptional control during early embryonic development, especially organ formation. Genes that sophisticate organs and body mechanisms, and genes associated with protein and fat digestion, hematopoietic stem cell lineage and osteoblast differentiation, are enriched during the development of the Ujumqin sheep 16-day embryo U(E16). Our work provides important complementary reference data for the study of early embryonic development in sheep.

INTRODUCTION

The domestication of sheep (*Ovis aries*), one of the first domesticated animals, always accompanies human migration and social development (Chessa *et al*., 2009; Lv *et al*., 2015). The inner Mongolia region of China is the main breeding area for meat sheep in China, with about 17 taxa collectively known as Mongolian sheep. Among them, Ujumqin sheep and Hulunbuir short-tailed sheep are well-known and profitable breeds distributed in the Hulunbuir Grassland and Xilingol grassland, respectively (Li *et al*., 2018; Zhi *et al*., 2018). Ujumqin sheep is a multi-vertebral breed with a fat tail (Li *et al*., 2018). The tail of Hulunbuir short-tailed sheep is shorter than those of other breeds, and this trait is genetically stable. In addition, Hulunbuir short-tailed sheep have more intermuscular fat, which is evenly distributed, and it is a typical kind of short-tailed sheep. Due to the short-tailed phenotype, Hulunbuir short-tailed sheep have been applied to artificial and natural selection to increase meat production (Zhi *et al*., 2018). Li *et al*. (2018) showed by whole-genome resequencing (WGRS) that the multi-vertebral characteristic in Ujumqin sheep was associated with

* Corresponding author: guifangcao@126.com, songyongli625@163.com, baofuxiang@imau.edu.cn
0030-9923/2024/0004-1743 $ 9.00/0

mutations in *CAMK1D, MLLT6, FOXJ1, ADAMTSL2* and *GUSB*. Zhu *et al*. (2021a) used isobaric tags for relative and absolute quantitation (iTRAQ) to reveal candidate proteins of fat deposition in Chinese native sheep with different tails and identified *APOA2, GALK1, ADIPOQ* and *NDUFS4* genes. They also identified *SPAG17, Tbx15, VRTN, NPC2, BMP2*, and *PDGFD* as the most promising candidates for tail-type traits through a genome-wide association study of genes related to tail fat deposition in Chinese sheep (Zhu *et al*., 2021b). Most studies on sheep tail morphology focus on tail fat deposition. For example, Xu *et al*. (2017) conducted a genome-wide association analysis to determine the genetic basis of fat deposition in the tail of two types of Mongolian sheep, namely, small-tailed Han sheep and large-tailed Han sheep. Moradi *et al*. (2012) identified candidate regions associated with fat deposition by genome-wide selection scan and found that increased homozygosity of the region on chromosomes 5 and X favored fat-tailed sheep breeds and increased homozygosity of the region on chromosome 7 favored thin-tailed sheep. Luo *et al*. (2021) conducted whole-genome sequencing (WGS) for three types of Mongolian sheep and found that *GLIS1, LOC101117953, PDGFD* and *T* might be associated with fat deposition in the sheep tail and tail morphology. They also suggested that GLIS1 may play a key role in germ layer differentiation in sheep embryos. Few studies focus on the formation of sheep tail vertebrae. Zhi *et al*. (2018) identified nine candidate genes associated with the short-tailed phenotype by WGS of Hulunbuir short-tailed sheep. Among these genes, the *T* gene was the strongest candidate because it was associated with vertebrae development. Guillomot *et al*. (2004) concluded in Staging of ovine embryos and expression of the T-box genes Brachyury and Eomesodermin around gastrulation that the expression level of the T gene was the highest in 16-day sheep embryo (E16). *T* gene significantly affects germ layer development. Pennimpede *et al*. (2012) found that *in vivo* knockdown of *T* gene resulted in skeletal defects and tail degeneration. The studies on both Ujumqin sheep and Hulunbuir short-tailed sheep are expected to improve breeding technology and increase economic benefits. We found studies on embryo development in sheep are rare. We chose Hulunbuir short-tailed sheep 16-day embryo H(E16) and Ujumqin sheep 16-day embryo U (E16) for reference transcriptome sequencing, which will help us to analyze the different developmental states of the two Mongolian sheep embryos during the same developmental period for better understanding of embryo development and scientific regulation of breeding production. It is expected that this paper can help to improve sheep welfare and increase economic efficiency.

MATERIALS AND METHODS

Samples

Hulunbuir short-tailed sheep (Fig. 1A) were from Ewenki Autonomous Banner, Hulunbuir City, Inner Mongolia Autonomous Region, China. Ujumqin sheep (Fig. 1B) were from East Ujumqin Banner, Xilinhot City, Inner Mongolia Autonomous Region, China. The two regions were about 1000 km apart. To ensure that the embryos were obtained at the same time, we transported the two types of sheep to the same ranch (Wuchuan County Breeding Sheep Ranch, Hohhot City). After one month, when the ewes and rams adapted to the local climate, we started the experiment.

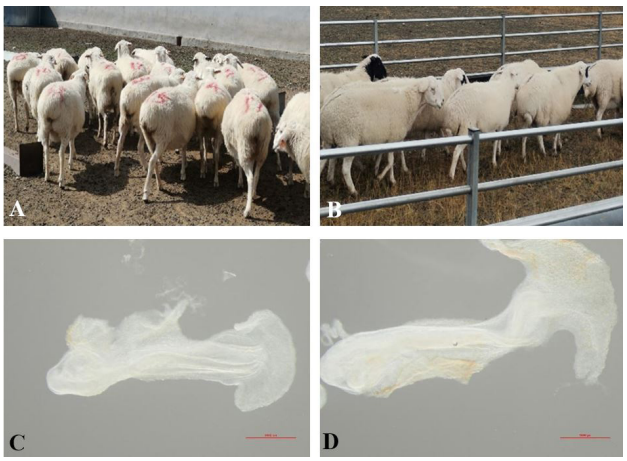

Fig. 1. A, Hulunbuir short-tailed sheep; B, Ujumqin sheep; C, Hulunbuir short-tailed sheep 16-day embryo, H(E16); D, Ujumqin sheep 16-day embryo, U(E16). Scale: 1000 μm.

Embryo sample collection

Twenty Ujumqin and 40 Hulunbuir healthy sheep, which were three years old, were treated with progesterone vaginal sponge embolization for 12 days. PMSG (330 IU/sheep) was intramuscularly injected before thrombectomy, and estrous ewes were mated after 36 hours. Embryos were collected at pregnancy 16 days (Fig. 1C, D). Embryos were washed three times using cold PBS and immediately transferred to tissue storage solution (Miltenyi,130-100-008) with 3% penicillin-streptomycin at room temperature.

mRNA library construction and sequencing

Total RNA was isolated and purified using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following manufacturer's procedures. RNA in each sample was

quantified, and the purity was evaluated using NanoDrop ND-1000 (NanoDrop, Wilmington, DE, USA). The RNA integrity was assessed by Bioanalyzer 2100 (Agilent, CA, USA) with RIN number >7.0 and confirmed by electrophoresis with denaturing agarose gel. Poly (A) RNA was purified from 1μg total RNA using Dynabeads Oligo (dT) 25-61005 (Thermo Fisher, CA, USA) by two rounds of purification. Then poly (A) RNA was fragmented into small pieces using Magnesium RNA fragmentation module (NEB, cat.e6150, USA) by being kept under 94°C for 5-7 min. Then the cleaved RNA fragments were reverse-transcribed to create cDNAs by Super Script™ II Reverse Transcriptase (Invitrogen, cat. 1896649, USA), which were subsequently used to synthesize U-labeled second-stranded DNAs with E. coli DNA polymerase I (NEB, cat.m0209, USA), RNase H (NEB, cat.m0297, USA) and dUTP Solution (Thermo Fisher, cat.R0133, USA). An A-base is then added to the blunt ends of each strand, preparing them for ligation to the indexed adapters. Each adapter contains a T-base overhang for ligating the adapter to the A-tailed fragmented DNA. Single- or dual-index adapters are ligated to the fragments, and size selection was performed with AMPureXP beads. After the heat-labile UDG enzyme (NEB, cat.m0280, USA) treatment of the U-labeled second-stranded DNAs, the ligated products were amplified with PCR under the following conditions: initial denaturation at 95°C for 3 min; eight cycles of denaturation at 98°C for 15 s, annealing at 60°C for 15 s, and extension at 72°C for 30 s; final extension at 72°C for 5 min. The average insert size for the final cDNA library was 300±50 bp. At last, we performed the 2×150bp paired-end sequencing (PE150) on an illumina novaseq™ 6000 (LC-bio technology Co., Ltd., Hangzhou, China) following vendor's recommended protocols.

Data analysis

The sequence data were deposited in the sequence read archive of the NCBI databases under bioproject number GSE186602 (https://www.ncbi.nlm.nih.gov/geo/info/linking.html, published on October 22, 2023).

Cutadapt software (https://cutadapt.readthedocs.io/en/stable/,version:cutadapt-1.9) was used to remove the reads that contained adaptor contamination (command line: ~ cutadapt -a ADAPT1 -A ADAPT2 -o out1.fastq -p out2.fastq in 1.fastq in 2 fastq -O 5 -m 100) (Beekman *et al*., 2018; Wang *et al*., 2012). After removing the low-quality bases and undetermined bases, we used HISAT2 software (https://daehwankimlab.github.io/hisat2/,version:hisat2-2.0.4) to map reads to the genome (for example: Homo sapiens Ensembl v96) (command line: ~hisat2 -1 R1.fastq.gz -2 R1.fastq.gz -S sample_mapped.sam) (Kim *et al*., 2015). The mapped reads of each sample were assembled using String Tie (http://ccb.jhu.edu/software/stringtie/,version:stringtie-1.3.4d.Linux_x86_64) with default parameters (command line: ~stringtie -p 4 -G genome.gtf -o output.gtf -l sample input.bam). Then, all transcriptomes from all samples were merged to reconstruct a comprehensive transcriptome using gff compare software (http://ccb.jhu.edu/software/stringtie/gffcompare.shtml,version:gffcompare-0.9.8.Linux_x86_64) (Pertea *et al*., 2015). After the final transcriptome was generated, StringTie and ballgown (http://www.bioconductor.org/packages/release/bioc/html/ballgown.html) were used to estimate the expression levels of all transcripts and expression level of mRNAs by calculating FPKM (FPKM = [total_exon_fragments / mapped_reads (millions) × exon_length(kB)]) (command line: ~stringtie -e -B -p 4 -G merged.gtf -o samples.gtf samples.bam). The differentially expressed mRNAs with fold change>2 or fold change < 0.5 and *P*-value<0.05 were selected by R package edgeR (https://bioconductor.org/packages/release/bioc/html/edgeR.html) or DESeq2 (http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html) (Robinson *et al*., 2010), and then GO ontology (Young *et al*., 2010) and KEGG enrichment analysis (Kanehisa *et al*., 2008) was performed for them.

qRT-PCR

RNA was isolated from Ujumqin and Hulunbuir sheep at E16 using RNA fast 200 (220011, Shanghai, China) according to manufacturer's directions. Reverse transcription was performed using PrimeScript RT reagent kit with gDNA Eraser (RR047A, Takara). Real-time PCR was performed using the ChamQ Universal SYBR qPCR master mix on a CFX96 PCR system (Bio-rad). Housekeeping gene *GAPDH* was used as internal control, and gene expression was calculated based on the ΔΔCT method. Statistical analysis was performed using one-way ANOVA and Student' s t-test (*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$). The primer sequences used are shown in Table I.

RESULTS

RNA-sequencing data

From the whole transcriptome of conceptus samples (n= 3 embryos of Hulunbuir short-tailed sheep and n = 3 embryos of Ujumqin sheep), an average of 52.41 million reads/sample was generated. After quality control, an average of 50.7 million reads/sample was kept. Moreover, an average of 87.36% of the reads were mapped against the sheep reference genome (*Ovis aries,* v.3.1, Ensembl 96).

DEGs between H(E16) and U(E16)

A total of 2152 DEGs between H(E16) and U(E16) were identified, with 922 genes up-regulated in H(E16) and 1230 genes up-regulated in U(E16) (Fig. 2, Table II).

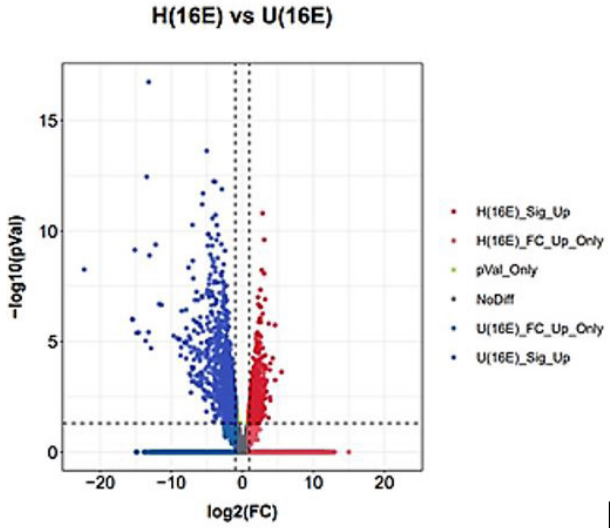

Fig. 2. The horizontal coordinate, log2(FC), represents the fold change of expression of genes in different samples. The vertical coordinate, -log10(pVal), represents the statistical significance of the difference in gene expression. Genes down-regulated in H(E16) are described as up-regulated in U(E16).

GO functional analyses

The GO enrichment analysis of the 992 upregulated genes in H(E16) revealed that these genes could be enriched to 877 GO terms, including 659 biological process (75.2%), 138 molecular function (15.7%) and 80 cellular component (9.1%) (Fig. 3A). The enriched GO terms highlighted several crucial factors, such as heart development, protein binding, axon guidance, outer ear morphogenesis, digestive tract morphogenesis, limb bud formation, and pharyngeal system development, for biological development.

GO annotation of the 1230 upregulated genes in U(E16) revealed that these genes could be annotated to 590 GO terms, including 403 biological process (68.3%), 121 molecular function (20.5%) and 66 cellular component (11.2%) (Fig. 3B). The enriched GO terms highlighted several crucial factors, such as extracellular space, apical plasma membrane, membrane, cell surface, platelet activation, and erythrocyte differentiation, for biological development. We used ggplot 2 to present GO enrichment analysis results in the form of bubble plots.

Table I. Primer sequences for qRT-PCR of sheep genes.

Gene name	Gene ID	Primer (5'-3')
GAPDH	443005	F CGGCACAGTCAAGGCAGAGAAC R CACGTACTCAGCACCAGCATCAC
ANXA1	101123269	F TGCTAAGGGTGACCGATCTGAGG R GCTCCTGGTGGTAAGAATGGTAGTG
AHSG	443392	F CCAGGTTCTGTCTCTGTGGAGTTTG R AGGTTGCACTTGGTTGGATCTACG
CDKN1A	100302083	F ACCACTTGGACCTGTCGCTGTC R GATCAGCCTGCGTTTGGAGTGG
FST	443323	F TCAAGTGGATGATTTTCAACGG R GTTCTTCTTGTTCATTCGGCAT
PRNP	493887	F GTGGCTACATGCTGGGAAGT R GGTGAAGTTCTCCCCCTTGG
YAP1	100913160	F AGAGGCCTGGCGCATCAAAA R CCAACAAGGTTCCAGCCCAT
TBXT	101114280	F CACCAAGAACGGCAGGAGGATG R CGTACTTCCAGCGGTGGTTGTC
NR2F1	101121774	F TTCGTCCGTTTGGTAGGTAAAA R GAGCACTGGATGGACATGTAAG
HOXB9	101114883	F CGAAGGAAGCGAGGACAAAGAGAG R CAGCGTCTGGTATTTGGTGTAGGG
SOX2	101110563	F GATCAGCATGTACCTCCCCG R ACATGTGAAGTCTGCTGGGG

KEGG functional analyses

KEGG enrichment analysis of the 922 up-regulated genes in H(E16) revealed that these genes were enriched in a total of 43 signaling pathways (Fig. 4A), including axon guidance, Hippo signaling pathway, signaling pathways regulating pluripotency of stem cells, TGF-beta signaling pathway, cardiac muscle contraction, MAPK signaling pathway, Wnt signaling pathway, etc.

KEGG enrichment analysis of the 1230 upregulated genes in U(E16) revealed that these genes were enriched in 68 signaling pathways (Fig. 4B), including protein digestion and absorption, fat digestion and absorption, PPAR signaling pathway, hematopoietic cell lineage, arachidonic acid metabolism, PI3K-Akt signaling pathway, etc.

qRT-PCR analysis

Ten genes were randomly selected from all mRNA sequencing data for qRT-PCR reaction to verify the accuracy of mRNA sequencing results (Fig. 5). The qRT-PCR results showed that the expression of mRNAs was different, and the expression trend was consistent with that obtained by sequencing, indicating that the mRNA sequencing results were reliable.

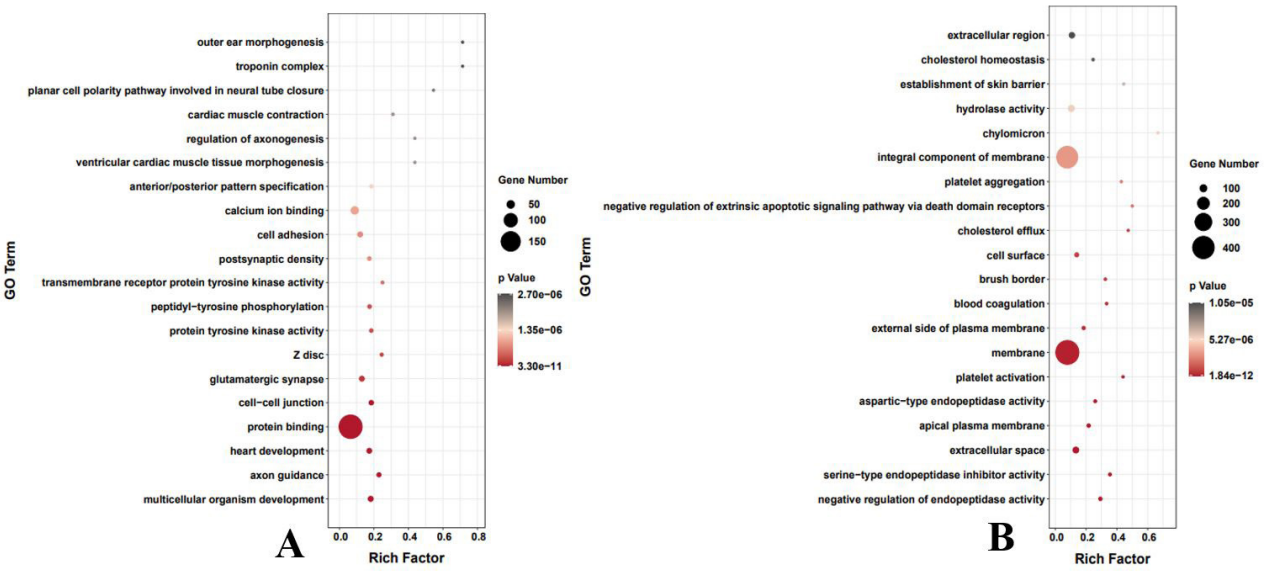

Fig. 3. Top 20 GO terms of H(E16) (A); U(E16) (B). The horizontal coordinate Rich Factor represents the number of differential genes in the GO/total number of genes in the GO; A larger Rich factor indicates a higher degree of GO enrichment. The vertical coordinate is GO term, i.e., GO functional annotation. In the scatter plot, the size of the dots represents the number of differential genes, and the color of the dots represents the *P*-value of the enrichment analysis, i.e., the significance of the enrichment, with a *P*-value smaller than or equal to 0.05 indicating significant enrichment. The GO enrichment analysis scatter plot is based on the significance (*P*-value) of the enrichment of the top 20 GO terms.

Table II. Top 10 upregulated genes in H(E16) and top 10 upregulated genes in U(E16).

Gene ID	Gene name	Description	log2 (FC)	pVal
H(E16)				
ENSOARG00000002632	*ENSOARG00000002632*		6.05	$6.9x10^{-22}$
ENSOARG00000011427	*ITGA7*	Integrin subunit alpha 7	2.85	$1.6x10^{-11}$
ENSOARG00000015275	*CCDC66*	Coiled-coil domain containing 66	3.08	$2.5x10^{-10}$
ENSOARG00000026938	*ENSOARG00000026938*		13.09	$1.5x10^{-9}$
ENSOARG00000006966	*UNC45B*	Unc-45 myosin chaperone B	2.73	$6.0x10^{-9}$
ENSOARG00000004153	*NKX2-5*	NK2 homeobox 5	3.05	$8.5x10^{-9}$
ENSOARG00000008405	*CASQ1*	Calsequestrin 1	2.54	$4.6x10^{-8}$
ENSOARG00000017490	*THBS4*	Thrombospondin 4	2.22	$1 x10^{-7}$
ENSOARG00000009361	*ENSOARG00000009361*		3.20	$1.2x10^{-7}$
ENSOARG00000012682	*HRC*	Histidine rich calcium binding protein	2.24	$2.1x10^{-7}$
U(E16)				
ENSOARG00000011224	*ENSOARG00000011224*	Meprin A subunit alpha	-13.20	$1.8X10^{-17}$
ENSOARG00000011165	*HEMGN*	Hemogen	-5.02	$2.4X10^{-14}$
ENSOARG00000001867	*ENSOARG00000001867*	Feline leukemia virus subgroup C receptor-related protein 2-like	-13.47	$3.6X10^{-13}$
ENSOARG00000010944	*ENSOARG00000010944*	Hemoglobin subunit alpha-1/2	-4.04	$5.8X10^{-13}$
ENSOARG00000006087	*ENSOARG00000006087*	Pancreatic trypsin inhibitor	-3.84	$6X10^{-13}$
ENSOARG00000015872	*ENSOARG00000015872*	Ubiquitin-conjugating enzyme E2 D3	-2.89	$1.3X10^{-12}$
ENSOARG00000013018	*S100G*	S100 calcium binding protein G	-5.54	$2X10^{-12}$
ENSOARG00000010469	*ATP6V0A4*	ATPase H+ transporting V0 subunit a4	-5.66	$6.3X10^{-12}$
ENSOARG00000026828	*ENSOARG00000026828*		-16.27	$7.7X10^{-12}$
ENSOARG00000007169	*KLF4*	Kruppel like factor 4	-3.79	$1.9X10^{-1}$

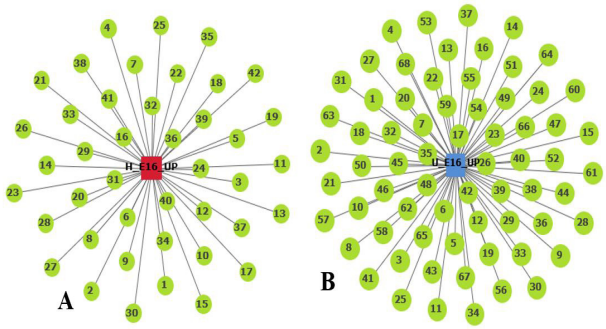

Fig. 4. A: The red node represents H(E16) 1, Dilated cardiomyopathy (DCM); 2, Axon guidance; 3, Hypertrophic cardiomyopathy (HCM); 4, Arrhythmogenic right ventricular cardiomyopathy (ARVC); 5, Cardiac muscle contraction; 6, Cell adhesion molecules (CAMs); 7, Adrenergic signaling in cardiomyocytes; 8, Glutamatergic synapse; 9, Breast cancer; 10, GABAergic synapse; 11, ECM-receptor interaction; 12, MAPK signaling pathway; 13, Basal cell carcinoma; 14, Hippo signaling pathway; 15, Morphine addiction; 16, Pathways in cancer; 17, Signaling pathways regulating pluripotency of stem cells; 18, cGMP-PKG signaling pathway; 19, Taste transduction; 20, TGF-beta signaling pathway; 21, Cholinergic synapse; 22, Nicotine addiction; 23, Rap1 signaling pathway; 24, Herpes simplex virus 1 infection; 25, Ras signaling pathway; 26, Wnt signaling pathway; 27, Cushing syndrome; 28, Central carbon metabolism in cancer; 29, Retrograde endocannabinoid signaling; 30, PI3K-Akt signaling pathway; 31, Protein digestion and absorption; 32, Oxytocin signaling pathway; 33, Vascular smooth muscle contraction; 34, Circadian entrainment; 35, cAMP signaling pathway; 36, Aldosterone synthesis and secretion; 37, Gastric cancer; 38, Calcium signaling pathway; 39, Apelin signaling pathway; 40, Melanogenesis; 41, Leukocyte transendothelial migration; 42, Focal adhesion; 43, Cortisol synthesis and secretion.
B: The blue node represents H(E16). 1, Protein digestion and absorption; 2, PPAR signaling pathway; 3, Cholesterol metabolism; 4, Platelet activation; 5, Complement and coagulation cascades; 6, Fc gamma R-mediated phagocytosis; 7, Hematopoietic cell lineage; 8, Ferroptosis; 9, Malaria; 10, Fat digestion and absorption; 11, Aldosterone-regulated sodium reabsorption; 12, p53 signaling pathway; 13, Osteoclast differentiation; 14, Epstein-Barr virus infection; 15, Proteoglycans in cancer; 16, Pantothenate and CoA biosynthesis; 17, Collecting duct acid secretion; 18, Chagas disease (American trypanosomiasis); 19, African trypanosomiasis; 20, Phagosome; 21, Chemical carcinogenesis; 22, Arachidonic acid metabolism; 23, Histidine metabolism; 24, B cell receptor signaling pathway; 25, Tyrosine metabolism; 26, Salmonella infection; 27, Lysosome; 28, Kaposi sarcoma-associated herpesvirus infection; 29, Phosphatidylinositol signaling system; 30, Regulation of actin cytoskeleton; 31, Human cytomegalovirus infection; 32, Phenylalanine metabolism; 33, Platinum drug resistance; 34, Mineral absorption; 35, NF-kappa B signaling pathway; 36, Human immunodeficiency virus 1 infection; 37, Fructose and mannose metabolism; 38, Riboflavin metabolism; 39, Drug metabolism - cytochrome P450; 40, Glutathione metabolism; 41, PI3K-Akt signaling pathway; 42, Adipocytokine signaling pathway; 43, Glioma; 44, Hepatitis B; 45, Transcriptional misregulation in cancer; 46, Glycolysis/ Gluconeogenesis; 47, AMPK signaling pathway; 48, Prolactin signaling pathway; 49, Tuberculosis; 50, Toll-like receptor signaling pathway; 51, Rheumatoid arthritis; 52, Porphyrin and chlorophyll metabolism; 53, Steroid hormone biosynthesis; 54, C-type lectin receptor signaling pathway; 55, Apoptosis; 56, Drug metabolism other enzymes; 57, Amoebiasis; 58, Jak-STAT signaling pathway; 59, Hepatitis C; 60, Endocrine resistance; 61, T cell receptor signaling pathway; 62, Bacterial invasion of epithelial cells; 63, Leukocyte transendothelial migration; 64, Viral carcinogenesis; 65, Amphetamine addiction; 66, Melanoma; 67, Natural killer cell mediated cytotoxicity; 68, Metabolism of xenobiotics by cytochrome P450. Green nodes represent specific pathways.

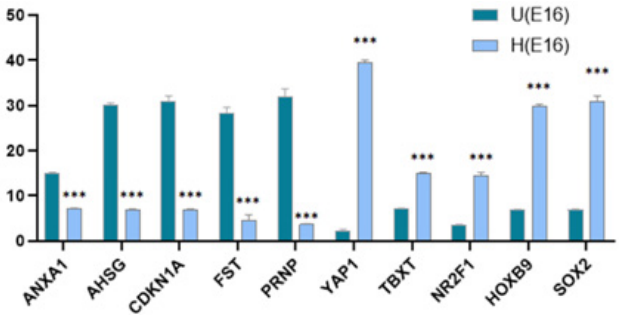

Fig. 5. mRNA sequencing results for randomly selected ten genes. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$

DISCUSSION

For ethical reasons, many questions about the human embryo cannot be directly investigated, and no animal model can truly reproduce human embryonic development. Pregnant sheep have been widely used to study maternal-fetal interactions, and a thorough understanding of the developmental events of the sheep embryo is necessary. In addition, studies of early embryonic development in sheep can help us to better regulate sheep gestation and reproduction, which will also contribute to economic development. Sanchez *et al*. (2021) showed that in 6-day sheep embryos produced in vivo, miRNAs regulate genes associated with the cell cycle and differentiation of stem cells and embryonic cells. The retrieved miRNAs also regulate the expression of genes involved in cellular

signaling pathways, such as MAPK, Wnt, TGF-beta, p53, and Toll-like receptors (Sanchez *et al*., 2021). These results are partially consistent with our data. By KEGG enrichment analysis of up-regulated genes in H(E16), we found that these genes could be enriched in signaling pathways regulating pluripotency of stem cells, TGF-beta signaling pathway, MAPK signaling pathway, and Wnt signaling pathway. In addition, we found that upregulated genes in H(E16) could also be enriched in axon guidance, Hippo signaling pathway, and cardiac muscle contraction, suggesting that the neurological and cardiac organ development of H(E16) is more prominent than that of 6-day sheep embryo.

GO analysis showed that H(E16) had more developing organs, such as heart, nerves, ear, digestive tract, limb buds, and pharynx, compared to U(E16) at the same age, which was mainly undergoing some physiological processes related to cell growth and hematopoietic cell differentiation. Based on the KEGG enrichment analysis of H(E16) and U(E16), we found that upregulated genes in U(E16) were enriched in protein digestion and absorption, fat digestion and absorption, PPAR signaling pathway, hematopoietic cell lineage, and PI3K-Akt signaling pathway, suggesting that U(E16) may have completed development of some organs while H(E16) was still in the state of organ development. This finding is consistent with the result of GO enrichment analysis, indicating that U(E16) developed faster than H(E16), which is a surprising finding. Clark *et al*. (2017) provided a sheep gene expression profiling dataset that extended the RNA-seq dataset to include data of adult sheep tissues such as the liver, spleen, ovary, testis, kidney, muscle, thymus, and left ventricle, and to include data during early development (blastocyst), early development (day 23), early development (day 35), early development (day 100), and maternal development (days 23, 35 and 100 days). Their work provides references for sheep research. However, until now, there have been no relevant data for 16-day sheep, and our study fills this research gap.

Above all, for H(E16), we focus more on signal pathways related to organogenesis, such as axon guidance, Hippo signaling pathway, signaling pathways regulating pluripotency of stem cells, TGF-beta signaling pathway, cardiac muscle contraction, MAPK signaling pathway, and Wnt signaling pathway. For U(E16), we focus more on protein digestion and absorption, fat digestion and absorption, Hematopoietic cell lineage, osteoclast differentiation, etc.

H(E16)

As the nervous system develops, newly differentiated neurons need to extend their axons toward synaptic targets in order to form functional neural circuits (Stoeckli, 2018). During embryonic development, various kinds of guidance information influence the direction of axon extension. A highly dynamic and sensitive structure located at the end of the axon the growth cone is responsible for interpreting guidance information. The binding of specific guidance molecules to receptors at the growth cone activates intracellular signaling pathways, and the combined action of several extracellular signals regulates the activity of the cytoskeleton, allowing axons to extend in the correct direction (de Wit and Verhaagen, 2003). In our study, upregulated genes in H(E16) were enriched in axon guidance. These genes included members of the semaphorin family that can influence processes such as axon bundling and branching and synapse formation, namely, SEMA3A, SEMA3G, SEMA4F and SEMA6C and included plexins in the semaphorin receptor family, namely PLXNA2, PLXNA3 and PLXNB1. In the axon guidance pathway, the EPH receptor and ligand families are involved in the regulation of embryonic neural stem cells (Moskowitz and Lo, 2003). Our data suggested that *EPHB4, EPHA4, EPHB3, EPHB2* and *EPHA7* were enriched. WNT5A was also enriched in the axon guidance pathway. The involvement of WNT5A, as an important ligand in the Wnt signaling pathway, in axon attraction connects the axon guidance pathway and the Wnt signaling pathway (Salinas, 2012). The proteins secreted by the Wnt family bind to different receptors and mediate a variety of signaling pathways to regulate the biological functions of cells, including cell polarization, cell proliferation and differentiation (Wodarz and Nusse, 1998; van Amerongen and Nusse, 2009). In our study, WNT2, AXIN2, WIF1 and PORCN (Boonekamp *et al*., 2021; Onizuka *et al*., 2012; Tríbulo *et al*., 2018; Barrott *et al*., 2011) involved in the canonical Wnt signaling pathway and membrane receptors WNT5A and ROR2 (Yamada *et al*., 2010) involved in the non-canonical Wnt signaling pathway (PCP) were enriched. In fact, the WNT5A, WNT2, and AXIN2 were also enriched in the Hippo signaling pathway, which is an evolutionarily conserved signaling pathway that controls the size of organs of species from flies to humans (Zhao *et al*., 2011). We also identified enriched DLG4, the repressor in the Hippo signaling pathway, and enriched transcription factors ID1 and SOX2. The hippo signaling pathway, as an important pathway regulating organogenesis during embryonic development, is not only associated with the Wnt signaling pathway but also interacts with the TGF-beta signaling pathway. BMP5 and TGFB2, as important ligands in the TGF-beta signaling pathway, affect osteogenic differentiation and apoptosis through the BMP signaling pathway and canonical TGF-beta signaling pathway, respectively (Hata and Chen,

2016; Lochab and Extavour, 2017). In our study, enriched transcriptional factor SMAD9 in the canonical TGF-beta signaling pathway (Hata and Chen, 2016) and the BMP inhibitor, i.e., CHRD (Hyvönen, 2003) were identified. In addition, the TGF-beta signaling pathway activated the MAPK-ERK signaling pathway and phosphorylated ERK upregulates the expression of ubiquitin ligase SMURF1 to inhibit the transportation of SMADS protein in the BMP signaling pathway into the nucleus, thereby affecting osteogenic differentiation (Sun *et al*., 2018). Our results showed that some factors of the CACN ligand family in the canonical MAPK signaling pathway, such as CACNA1A, CACNA1C, CACNB2, CACNB1 and CACNA1D and partial receptors and ligands of the MAPK-RAS pathway, such as FGF8, KIT, FGFR1, VEGFC, and NTRK1 (Faivre *et al*., 2006) were enriched. The upregulation of ERK was not found. In the analysis of the above signaling pathways, we were surprised to find that upregulated genes in H(E16) were not only enriched in axon guidance, Hippo signaling pathway, TGF-beta signaling pathway, MAPK signaling pathway and Wnt signaling pathway but also enriched in signaling pathways regulating pluripotency of stem cells, and cardiac muscle contraction, such as WNT2, SMAD9, SOX2, AXIN2, WNT5A, ID1, FGFR1, CACNA1C, CACNB2, CACNA1D, and CACNB1. Despite previous reports on Hulunbuir short-tailed sheep (Zhi *et al*., 2018) and studies on sheep embryos (Sanchez *et al*., 2021), the enrichment of these genes that we reported demonstrates that H(E16) is at the stage of organogenesis, which is a novel finding.

U(E16)

Protein is an essential dietary component for the nutritional balance of the human body as it is for sheep, which are ruminants. In U(E16), some of the upregulated genes are significantly enriched in protein digestion and absorption. Although most of these genes remain unexplored, there are still some genes that are well known, such as DPP4, also known as CD26. It is a type II transmembrane protein released from the cell membrane by a non-classical secretion mechanism. This exopeptidase selectively degrades a variety of substrates, including insulinotropic hormones, growth factors and cytokines (Nargis and Chakrabarti, 2018). XPNPEP2, a member of the aminopeptidase family, can perform amino acid excision and hydrolyze proteins to amino acids (Duan *et al*., 2005). MME is a member of the matrix metalloproteinase (MMP) family, also known as MMP-12. Unlike other MMP members, MME breaks down plasminogen to produce angiostatin, which has an inhibitory effect on vascular endothelial cell proliferation (Gorrin-Rivas *et al*., 2000). ACE2 is an important amino acid transport factor (Khodadoost *et al*., 2020) and regulates SLC7A7, SLC7A8, SLC7A9 and SLC15A1 of the solute carrier family (SLC) for the transmembrane transport of many important endogenous and exogenous substances, including nutrients and drugs (Peitzsch *et al*., 2014).

The digestion of lipids begins in the mouth, where the process of being broken down to triglycerides starts under the action of lingual lipase secreted by glands in the tongue. Lipid emulsion enters the duodenum as fine lipid droplets, which then mix with the bile and pancreatic juice and undergo significant changes in chemical and physical forms. Emulsification continues in the duodenum and is accompanied by hydrolysis and micellization, which set conditions for absorption through the intestinal wall (Iqbal and Hussain, 2009). In our study, some genes associated with fat digestion and absorption were significantly enriched. PLPP1 is a lipid phosphatase (Tang *et al*., 2015). MTTP is an endoplasmic reticulum resident protein (Wetterau *et al*., 1997) that primarily transfers triglycerides to promote optimal folding of nascent APOB and shuttles among other lipid classes such as cholesteryl esters, free cholesterol, phospholipids, ceramides, and sphingolipids to further promote lipoprotein formation (Iqbal *et al*., 2015; Peng *et al*., 2021). Fatty acid-binding protein FABP1 is a liver-specific FABP that plays an important role in intracellular lipid metabolism in the liver (Pi *et al*., 2019). A previous study has shown that FABP1 overexpression significantly increases fatty acid uptake by hepatocytes (Wu *et al*., 2016). APOB is a large amphiphilic glycoprotein that plays a central role in lipoprotein metabolism. Two forms of APOB (APOB-48 and APOB-100) are produced from the APOB gene through a unique post-transcriptional editing process. APOB-48 is necessary for chylomicron production in the small intestine, and APOB-100 is a ligand for LDL receptor-mediated endocytosis of LDL particles (Whitfield *et al*., 2004).

Blood cells are derived from hematopoietic stem cells (HSCs), which can self-renew or differentiate into multi-lineage committed progenitors, such as common lymphoid progenitors (CLPs) or common myeloid progenitors (CMPs). CLP produces leukocytes or the lymphoid lineage of leukocytes natural killer (NK) cells and T and B lymphocytes. CMP produces the bone marrow lineage, which includes the remaining leukocytes, erythrocytes, and megakaryocytes that produce platelets important for clotting. Cells undergoing these differentiation processes express surface markers corresponding to specific particular stages and lineages. Thus, the cell stage is determined by the specific expression patterns of these genes (Pouzolles *et al*., 2016). MME (CD10) is a marker of the early stage of the differentiation of HSC into B cells.

CD22 is expressed during B cell formation and on the surface of the formed B cell 54,55. A recent study suggests that MME is also a pro-adipogenic, specific cell surface marker for adipose stem cells (ASCs) (Chakraborty *et al*., 2021). IL1R2 (CD121) and CSF1R (CD115) are marker genes for the early stage of neutrophil formation (Maekawa *et al*., 2019). CSF1R(CD115) and CD14 are markers for the early stage of macrophage formation (Auffray *et al*., 2009). IL6R (CD126) is a marker in the middle and late stages of neutrophil formation and a maker for the whole process of platelet formation. ITGA1, ITGA2 and ITGA6 are expressed on the membrane of platelets (Kobayashi *et al*., 1994). CD44 is specifically expressed in the early phase of T cell formation and is also a surface marker of erythrocytes (An and Chen, 2018; Föger *et al*., 2000).

Osteoclasts are multinucleated cells derived from the hematopoietic monocyte-macrophage lineage and are responsible for bone resorption. Mice lacking FOS (encoding C-FOS) develop osteosclerosis due to a block to early differentiation in the osteoblast lineage. C-FOS is a component of dimeric transcription factor-activating protein-1 (AP-1), consisting mainly of FOS (C-FOS, FOSB, FOSL1) and JUN proteins (Matsuo *et al*., 2000). CSF1R is abundantly expressed in the early stage of osteoclast differentiation and is used to track osteoclast precursors (Yahara *et al*., 2020). The high expression level of APC5 implies the formation of osteoclasts as it is the marker for osteoclasts (Muise *et al*., 2016). Osteoclasts can be formed through the indirect action of JAK1 and STAT3 in inflammatory cells and other helper cells (Sims, 2020). Tyrosine kinase TEC can mediate osteoclast formation through a complex of LCP2 and BLNK (Shinohara *et al*., 2008).

CONCLUSIONS

This study compares the transcriptional changes of E16 of two types of Mongolian sheep, namely, Hulunbuir short-tailed and Ujumqin sheep. Upregulated genes in H (E16) associated with axon guidance, Hippo signaling pathway, signaling pathways regulating pluripotency of stem cells, TGF-beta signaling pathway, cardiac muscle contraction, MAPK signaling pathway, and Wnt signaling pathway prove transcriptional regulation related to cellular interactions and organogenesis mainly occurs in early prenatal development. The biological processes, including protein digestion and absorption, fat digestion and absorption, Hematopoietic cell lineage, and osteoclast differentiation, are predominantly evidenced by upregulated genes in U(E16), reflecting the more complex development of organs and body structures during the U(E16) phase.

ACKNOWLEDGMENTS

The author are grateful to Wuchuan County Breeding Sheep Ranch Hohhot City, China for the support in completing this study.

Funding

This project is supported by the National Natural Science Foundation of China (No. 31860689); Inner Mongolia Autonomous Region Science and Technology Plan of China (No. 2019GG241) (No. 2020ZD0007); Inner Mongolia Autonomous Region Science and Technology Achievements Transformation Project (No.2020CG0078); Key Technology Project of Inner Mongolia Autonomous Region (No. 2021GG0062).

IRB approval

The study and animals experiment in the study were reviewed by the Ethics Committee of Inner Mongolia Agricultural University

Ethics statement

All sheep experimental procedures and protocols in this study were approved and authorized by the animal care and use committee of Inner Mongolia Agricultural University (License NO. SYXK, Inner Mongolia, 2016-0017).

Statement of conflict of interest

The authors have declared no conflict of interest.

REFERENCES

An, X. and Chen, L., 2018. Flow cytometry (FCM) analysis and fluorescence activated cell sorting (FACS) of erythroid cells. *Methods Mol. Biol.*, **1698**: 153-174. https://doi.org/10.1007/978-1-4939-7428-3_9

Auffray, C., Sieweke, M.H. and Geissmann, F., 2009. Blood monocytes: Development, heterogeneity, and relationship with dendritic cells. *Annls Rev. Immunol.*, **27**: 669-692. https://doi.org/10.1146/annurev.immunol.021908.132557

Barrott, J.J., Cash, G.M., Smith, A.P., Barrow, J.R. and Murtaugh, L.C., 2011. Deletion of mouse Porcn blocks Wnt ligand secretion and reveals an ectodermal etiology of human focal dermal hypoplasia/ Goltz syndrome. *Proc. natl. Acad. Sci. U.S.A.*, **108**: 12752-12757. https://doi.org/10.1073/pnas.1006437108

Beekman, R., Chapaprieta, V., Russiñol, N., Vilarrasa-Blasi, R., Verdaguer-Dot, N., Martens, J.H.A.,

Duran-Ferrer, M., Kulis, M., Serra, F., Javierre, B.M., Wingett, S.W., Clot, G., Queirós, A.C., Castellano, G., Blanc, J., Gut, M., Merkel, A., Heath, S., Vlasova, A., Ullrich, S., Palumbo, E., Enjuanes, A., Martín-García, D., Beà, S., Pinyol, M., Aymerich, M., Royo, R., Puiggros, M., Torrents, D., Datta, A., Lowy, E., Kostadima, M., Roller, M., Clarke, L., Flicek, P., Agirre, X., Prosper, F., Baumann, T., Delgado, J., López-Guillermo, A., Fraser, P., Yaspo, M.L., Guigó, R., Siebert, R., Martí-Renom, M.A., Puente, X.S., López-Otín, C., Gut, I., Stunnenberg, H.G., Campo, E. and Martin-Subero, J.I., 2018. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat. Med.*, **24**: 868-880. https://doi.org/10.1038/s41591-018-0028-4

Boonekamp, K.E., Heo, I., Artegiani, B., Asra, P., van Son, G., de Ligt, J. and Clevers, H., 2021. Identification of novel human Wnt target genes using adult endodermal tissue-derived organoids. *Dev. Biol.*, **474**: 37-47. https://doi.org/10.1016/j.ydbio.2021.01.009

Chakraborty, S., Ong, W.K., Yau, W.W.Y., Zhou, Z., Bhanu Prakash, K.N., Toh, S.A., Han, W., Yen, P.M. and Sugii, S., 2021. CD10 marks non-canonical PPARγ-independent adipocyte maturation and browning potential of adipose-derived stem cells. *Stem Cell Res. Ther.*, **12**: 109. https://doi.org/10.1186/s13287-021-02179-y

Chessa, B., Pereira, F., Arnaud, F., Amorim, A., Goyache, F., Mainland, I., Kao, R.R., Pemberton, J.M., Beraldi, D., Stear, M.J., Alberti, A., Pittau, M., Iannuzzi, L., Banabazi, M.H., Kazwala, R.R., Zhang, Y.P., Arranz, J.J., Ali, B.A., Wang, Z., Uzun, M., Dione, M.M., Olsaker, I., Holm, L.E., Saarma, U., Ahmad, S., Marzanov, N., Eythorsdottir, E., Holland, M.J., Ajmone-Marsan, P., Bruford, M.W., Kantanen, J., Spencer, T.E. and Palmarini, M., 2009. Revealing the history of sheep domestication using retrovirus integrations. *Science*, **324**: 532-536. https://doi.org/10.1126/science.1170587

Clark, E.L., Bush, S.J., McCulloch, M.E.B., Farquhar, I.L., Young, R., Lefevre, L., Pridans, C., Tsang, H.G., Wu, C., Afrasiabi, C., Watson, M., Whitelaw, C.B., Freeman, T.C., Summers, K.M., Archibald, A.L. and Hume, D.A., 2017. A high resolution atlas of gene expression in the domestic sheep (*Ovis aries*). *PLoS Genet.,* **13**: e1006997.

De Wit, J. and Verhaagen, J., 2003. Role of semaphorins in the adult nervous system. *Prog. Neurobiol.*, **71**: 249-267. https://doi.org/10.1016/j.pneurobio.2003.06.001

Duan, Q.L., Nikpoor, B., Dube, M.P., Molinaro, G., Meijer, I.A., Dion, P., Rochefort, D., Saint-Onge, J., Flury, L., Brown, N.J., Gainer, J.V., Rouleau, J.L., Agostoni, A., Cugno, M., Simon, P., Clavel, P., Potier, J., Wehbe, B., Benarbia, S., Marc-Aurele, J., Chanard, J., Foroud, T., Adam, A. and Rouleau, G.A., 2005. A variant in XPNPEP2 is associated with angioedema induced by angiotensin I-converting enzyme inhibitors. *Am. J. Hum. Genet.*, **77**: 617-626. https://doi.org/10.1086/496899

Faivre, S., Djelloul, S. and Raymond, E., 2006. New paradigms in anticancer therapy: Targeting multiple signaling pathways with kinase inhibitors. *Semin. Oncol.*, **33**: 407-420. https://doi.org/10.1053/j.seminoncol.2006.04.005

Föger, N., Marhaba, R. and Zöller, M., 2000. CD44 supports T cell proliferation and apoptosis by apposition of protein kinases. *Eur. J. Immunol.*, **30**: 2888-2899. https://doi.org/10.1002/1521-4141(200010)30:10<2888::AID-IMMU2888>3.0.CO;2-4

Gorrin-Rivas, M.J., Arii, S., Furutani, M., Mizumoto, M., Mori, A., Hanaki, K., Maeda, M., Furuyama, H., Kondo, Y. and Imamura, M., 2000. Mouse macrophage metalloelastase gene transfer into a murine melanoma suppresses primary tumor growth by halting angiogenesis. *Clin. Cancer Res.*, **6**: 1647-1654.

Guillomot, M., Turbe, A., Hue, I. and Renard, J.P. 2004. Staging of ovine embryos and expression of the *T*-box genes Brachyury and Eomesodermin around gastrulation. *Reproduction*, **127**: 491-501. https://doi.org/10.1530/rep.1.00057

Hata, A. and Chen, Y.G. 2016. TGF-β signaling from receptors to SMADS. *Cold Spring Harb. Perspect. Biol.*, **8**: https://doi.org/10.1101/cshperspect.a022061

Hyvönen, M., 2003. CHRD, a novel domain in the BMP inhibitor chordin, is also found in microbial proteins. *Trends Biochem. Sci.*, **28**: 470-473. https://doi.org/10.1016/S0968-0004(03)00171-3

Iqbal, J. and Hussain, M.M., 2009. Intestinal lipid absorption. *Am. J. Physiol. Endocrinol. Metab.*, **296**: E1183-94. https://doi.org/10.1152/ajpendo.90899.2008

Iqbal, J., Walsh, M.T., Hammad, S.M., Cuchel, M., Tarugi, P., Hegele, R.A., Davidson, N.O., Rader, D.J., Klein, R.L. and Hussain, M.M., 2015. Microsomal triglyceride transfers protein transfers and determines plasma concentrations of ceramide and sphingomyelin but not glycosylceramide. *J. biol. Chem.*, **290**: 25863-25875. https://doi.

org/10.1074/jbc.M115.659110

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. and Yamanishi, Y. 2008. KEGG for linking genomes to life and the environment. *Nucl. Acids Res.*, **36**: D480-484. https://doi.org/10.1093/nar/gkm882

Khodadoost, M., Niknam, Z., Farahani, M., Razzaghi, M. and Norouzinia, M., 2020. Investigating the human protein-host protein interactome of SARS-CoV-2 infection in the small intestine. *Gastroenterol. Hepatol. Bed. Bench*, **13**: 374-387.

Kim, D., Langmead, B. and Salzberg, S.L., 2015. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods*, **12**: 357-360. https://doi.org/10.1038/nmeth.3317

Kobayashi, S., Teramura, M., Oshimi, K. and Mizoguchi, H. 1994. Interleukin-11. *Leuk. Lymphoma*, **15**: 45-49. https://doi.org/10.3109/10428199409051676

Li, S., Luo, R., Lai, D., Ma, M., Hao, F., Qi, X., Liu, X. and Liu, D., 2018. Whole-genome resequencing of Ujumqin sheep to investigate the determinants of the multi-vertebral trait. *Genome*, **61**: 653-661. https://doi.org/10.1139/gen-2017-0267

Lochab, A.K. and Extavour, C.G., 2017. Bone morphogenetic protein (BMP) signaling in animal reproductive system development and function. *Dev. Biol.*, **427**: 258-269. https://doi.org/10.1016/j.ydbio.2017.03.002

Luo, R., Zhang, X., Wang, L., Zhang, L., Li, G. and Zheng, Z., 2021. GLIS1, a potential candidate gene affect fat deposition in sheep tail. *Mol. Biol. Rep.*, **48**: 4925-4931. https://doi.org/10.1007/s11033-021-06468-w

Lv, F.H., Peng, W.F., Yang, J., Zhao, Y.X., Li, W.R., Liu, M.J., Ma, Y.H., Zhao, Q.J., Yang, G.L., Wang, F., Li, J.Q., Liu, Y.G., Shen, Z.Q., Zhao, S.G., Hehua, E., Gorkhali, N.A., Farhad Vahidi, S.M., Muladno, M., Naqvi, A.N., Tabell, J., Iso-Touru, T., Bruford, M.W., Kantanen, J., Han, J.L. and Li, M.H., 2015. Mitogenomic meta-analysis identifies two phases of migration in the history of eastern Eurasian Sheep. *Mol. Biol. Evol.*, **32**: 2515-2533. https://doi.org/10.1093/molbev/msv139

Maekawa, S., Wang, P.C. and Chen, S.C., 2019. Differential expression of immune-related genes in head kidney and spleen of cobia (Rachycentron canadum) having Streptococcus dysgalactiae infection. *Fish Shellfish Immunol.*, **92**: 842-850. https://doi.org/10.1016/j.fsi.2019.07.009

Matsuo, K., Owens, J.M., Tonko, M., Elliott, C., Chambers, T.J. and Wagner, E.F., 2000. Fosl1 is a transcriptional target of c-Fos during osteoclast differentiation. *Nat. Genet.*, **24**: 184-187. https://doi.org/10.1038/72855

Moradi, M.H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K.G., McEwan, J.C., 2012. Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genet*, **13**: 10. https://doi.org/10.1186/1471-2156-13-10

Moskowitz, M.A. and Lo, E.H., 2003. Neurogenesis and apoptotic cell death. *Stroke*, **34**: 324-326. https://doi.org/10.1161/01.STR.0000054047.14853.AD

Muise, E.S., Podtelezhnikov, A.A., Pickarski, M., Loboda, A., Tan, Y., Hu, G., Thomspon, J.R. and Duong le, T., 2016. Effects of long-term odanacatib treatment on bone gene expression in ovariectomized adult rhesus monkeys: Differentiation from alendronate. *J. Bone Miner. Res.*, **31**: 839-851. https://doi.org/10.1002/jbmr.2752

Nargis, T. and Chakrabarti, P., 2018. Significance of circulatory DPP4 activity in metabolic diseases. *IUBMB Life*, **70**: 112-119. https://doi.org/10.1002/iub.1709

Onizuka, T., Yuasa, S., Kusumoto, D., Shimoji, K., Egashira, T., Ohno, Y., Kageyama, T., Tanaka, T., Hattori, F., Fujita, J., Ieda, M., Kimura, K., Makino, S., Sano, M., Kudo, A. and Fukuda, K., 2012. Wnt2 accelerates cardiac myocyte differentiation from ES-cell derived mesodermal cells via non-canonical pathway. *J. Mol. Cell Cardiol.*, **52**: 650-659. https://doi.org/10.1016/j.yjmcc.2011.11.010

Peitzsch, C., Perrin, R., Hill, R.P., Dubrovska, A. and Kurth, I., 2014. Hypoxia as a biomarker for radioresistant cancer stem cells. *Int. J. Radiat. Biol.*, **90**: 636-652. https://doi.org/10.3109/09553002.2014.916841

Peng, H., Chiu, T.Y., Liang, Y.J., Lee, C.J., Liu, C.S., Suen, C.S., Yen, J.J., Chen, H.T., Hwang, M.J., Hussain, M.M., Yang, H.C. and Yang-Yen, H.F., 2021. PRAP1 is a novel lipid-binding protein that promotes lipid absorption by facilitating MTTP-mediated lipid transport. *J. biol. Chem.*, **296**: 100052. https://doi.org/10.1074/jbc.RA120.015002

Pennimpede, T., Proske, J., König, A., Vidigal, J.A., Morkel, M., Bramsen, J.B., Herrmann, B.G. and Wittler, L., 2012. *In vivo* knockdown of Brachyury results in skeletal defects and urorectal malformations resembling caudal regression syndrome. *Dev. Biol.*, **372**: 55-67. https://doi.org/10.1016/j.ydbio.2012.09.003

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T. and Salzberg, S.L., 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.*, **33**: 290-295. https://doi.org/10.1038/nbt.3122

Pi, H., Liu, M., Xi, Y., Chen, M., Tian, L., Xie, J., Chen, M., Wang, Z., Yang, M., Yu, Z., Zhou, Z. and Gao, F., 2019. Long-term exercise prevents hepatic steatosis: A novel role of FABP1 in regulation of autophagy-lysosomal machinery. *FASEB J.*, **33**: 11870-11883. https://doi.org/10.1096/fj.201900812R

Pouzolles, M., Oburoglu, L., Taylor, N. and Zimmermann, V.S., 2016. Hematopoietic stem cell lineage specification. *Curr. Opin. Hematol.*, **23**: 311-317. https://doi.org/10.1097/MOH.0000000000000260

Robinson, M.D., McCarthy, D.J. and Smyth, G.K., 2010. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**: 139-140. https://doi.org/10.1093/bioinformatics/btp616

Salinas, P.C., 2012. Wnt signaling in the vertebrate central nervous system: from axon guidance to synaptic function. *Cold Spring Harb. Perspect. Biol.*, **4**: 5. https://doi.org/10.1101/cshperspect.a008003

Sanchez, D.J.D., Vasconcelos, F.R., Teles-Filho, A.C.A., Viana, A.G.A., Martins, A.M.A., Sousa, M.V., Castro, M.S., Ricart, C.A., Fontes, W., Bertolini, M., Bustamante-Filho, I.C. and Moura, A.A., 2021. Proteomic profile of pre-implantational ovine embryos produced *in vivo*. *Reprod. Domest. Anim.*, **56**: 586-603. https://doi.org/10.1111/rda.13897

Shinohara, M., Koga, T., Okamoto, K., Sakaguchi, S., Arai, K., Yasuda, H., Takai, T., Kodama, T., Morio, T., Geha, R.S., Kitamura, D., Kurosaki, T., Ellmeier, W. and Takayanagi, H., 2008. Tyrosine kinases Btk and Tec regulate osteoclast differentiation by linking RANK and ITAM signals. *Cell*, **132**: 794-806. https://doi.org/10.1016/j.cell.2007.12.037

Sims, N.A., 2020. The JAK1/STAT3/SOCS3 axis in bone development, physiology, and pathology. *Exp. mol. Med.*, **52**: 1185-1197. https://doi.org/10.1038/s12276-020-0445-6

Stoeckli, E.T., 2018. Understanding axon guidance are we nearly there yet? *Development*, **145**: 10. https://doi.org/10.1242/dev.151415

Sun, X., Xie, Z., Ma, Y., Pan, X., Wang, J., Chen, Z. and Shi, P., 2018. TGF-β inhibits osteogenesis by upregulating the expression of ubiquitin ligase SMURF1 via MAPK-ERK signaling. *J. Cell Physiol.*, **233**: 596-606. https://doi.org/10.1002/jcp.25920

Tang, X., Benesch, M.G. and Brindley, D.N., 2015. Lipid phosphate phosphatases and their roles in mammalian physiology and pathology. *J. Lipid Res.*, **56**: 2048-2060. https://doi.org/10.1194/jlr.R058362

Tríbulo, P., Siqueira, L.G.B., Oliveira, L.J., Scheffler, T. and Hansen, P.J., 2018. Identification of potential embryokines in the bovine reproductive tract. *J. Dairy Sci.*, **101**: 690-704. https://doi.org/10.3168/jds.2017-13221

van Amerongen, R. and Nusse, R., 2009. Towards an integrated view of Wnt signaling in development. *Development*, **136**: 3205-3214. https://doi.org/10.1242/dev.033910

Wang, L., Wang, S. and Li, W., 2012. RSeQC: Quality control of RNA-seq experiments. *Bioinformatics*, **28**: 2184-2185. https://doi.org/10.1093/bioinformatics/bts356

Wetterau, J.R., Lin, M.C. and Jamil, H., 1997. Microsomal triglyceride transfer protein. *Biochim. biophys. Acta*, **1345**: 136-150. https://doi.org/10.1016/S0005-2760(96)00168-3

Whitfield, A.J., Barrett, P.H., van Bockxmeer, F.M. and Burnett, J.R., 2004. Lipid disorders and mutations in the APOB gene. *Clin. Chem.*, **50**: 1725-1732. https://doi.org/10.1373/clinchem.2004.038026

Wodarz, A. and Nusse, R. 1998. Mechanisms of Wnt signaling in development. *Annls Rev. Cell Dev. Biol.*, **14**: 59-88. https://doi.org/10.1146/annurev.cellbio.14.1.59

Wu, Y.L., Peng, X.E., Zhu, Y.B., Yan, X.L., Chen, W.N. and Lin, X., 2016. Hepatitis B virus X protein induces hepatic steatosis by enhancing the expression of liver fatty acid binding protein. *J. Virol.*, **90**: 1729-1740. https://doi.org/10.1128/JVI.02604-15

Xu, S.S., Ren, X., Yang, G.L., Xie, X.L., Zhao, Y.X., Zhang, M., Shen, Z.Q., Ren, Y.L., Gao, L., Shen, M., Kantanen, J. and Li, M.H., 2017. Genome-wide association analysis identifies the genetic basis of fat deposition in the tails of sheep (*Ovis aries*). *Anim. Genet.*, **48**: 560-569. https://doi.org/10.1111/age.12572

Yahara, Y., Barrientos, T., Tang, Y.J., Puviindran, V., Nadesan, P., Zhang, H., Gibson, J.R., Gregory, S.G., Diao, Y., Xiang, Y., Qadri, Y.J., Souma, T., Shinohara, M.L. and Alman, B.A., 2020. Erythromyeloid progenitors give rise to a population of osteoclasts that contribute to bone homeostasis and repair. *Nat. Cell Biol.*, **22**: 49-59. https://doi.org/10.1038/s41556-019-0437-8

Yamada, M., Udagawa, J., Matsumoto, A., Hashimoto, R., Hatta, T., Nishita, M., Minami, Y. and Otani, H., 2010. Ror2 is required for midgut elongation during mouse development. *Dev. Dyn.*, **239**: 941-953. https://doi.org/10.1002/dvdy.22212

Young, M.D., Wakefield, M.J., Smyth, G.K. and Oshlack, A., 2010. Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.*, **11**: R14. https://doi.org/10.1186/gb-2010-11-2-r14

Zhao, B., Tumaneng, K. and Guan, K.L., 2011. The Hippo pathway in organ size control, tissue regeneration and stem cell self-renewal. *Nat. Cell Biol.*, **13**: 877-883. https://doi.org/10.1038/ncb2303

Zhi, D., Da, L., Liu, M., Cheng, C., Zhang, Y., Wang, X., Li, X., Tian, Z., Yang, Y., He, T., Long, X., Wei, W. and Cao, G., 2018. Whole genome sequencing of hulunbuir short-tailed sheep for identifying candidate genes related to the short-tail phenotype. *G3* (Bethesda), **8**: 377-383. https://doi.org/10.1534/g3.117.300307

Zhu, C., Cheng, H., Li, N., Liu, T. and Ma, Y., 2021a. Isobaric tags for relative and absolute quantification-based proteomics reveals candidate proteins of fat deposition in Chinese indigenous sheep with morphologically different tails. *Front. Genet.*, **12**: 710449. https://doi.org/10.3389/fgene.2021.710449

Zhu, C., Li, N., Cheng, H. and Ma, Y., 2021b. Genome wide association study for the identification of genes associated with tail fat deposition in Chinese sheep breeds. *Biol. Open*, **10**: 5. https://doi.org/10.1242/bio.054932