Comparative Analysis of Long Non-Coding RNA and mRNA Expression Provides Insights Into Adaptation to Hypoxia in Tibetan Sheep

Fan Wang
College of Veterinary Medicine, Gansu Agricultural University
Jianbin Liu
Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences
Qiaoying Zeng (zengqy@gsau.edu.cn)
College of Veterinary Medicine, Gansu Agricultural University
Deqing Zhuoga
Institute of Livestock Research, Tibet Academy of Agriculture and Animal Science

Research Article

Keywords: Tibetan sheep, transcriptome, lncRNAs, mRNAs, hypoxia

Posted Date: July 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-77480/v2

Comparative analysis of long non-coding RNA and mRNA expression provides insights into adaptation to hypoxia in Tibetan Sheep

Fan Wang [a,b,], Jianbin Liu [c,e,], Qiaoying Zeng [a,*], and Deqing Zhuoga [d,*]

[a] College of Veterinary Medicine, Gansu Agricultural University, Lanzhou 730070, China; zengqy@gsau.edu.cn (Q.Z.)

[b] China Agricultural Veterinarian Biology Science and Technology Co. Ltd, Lanzhou 730046, wangfaneasy@126.com (F.W.)

[c] Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences, Lanzhou 730050, China; liujianbin@caas.cn (J.L.)

[d] Institute of Livestock Research, Tibet Academy of Agriculture and Animal Science, Lhasa 850000, China; 13989099966@163.com

[e] Sheep Breeding Engineering Technology Research Center, Chinese Academy of Agricultural Sciences, Lanzhou 730050, China

***** Correspondence: zengqy@gsau.edu.cn (Q.Z.); 13989099966@163.com(Q.D.); Tel.: +86-13893126975 (F.W.)

Numbers of Figures: 5

Numbers of Tables: 2

Numbers of Supplemental Figures: 4

Numbers of Supplemental Tables: 10

Numbers of Words: 9202

Figure and Table Legends of Supplementary information

Fig. S1: Identification of long non-coding RNAs (lncRNAs) and mRNAsI involved in high altitude hypoxia adaptation in liver tissue. Coding potentiality filter using Coding Potential Calculator, Pfam, Phylogenetic codon substitution frequency, and Coding-Non-Coding Index (A). The boxplot shows the expression levels (log10 (FPKM +1)) of lncRNAs and mRNAs (B). Transcript lengths distribution of mRNAs (C) and lncRNAs (D). Exon number distribution of mRNAs (E) and lncRNAs (F). Open reading frame (ORF) length distribution of mRNAs (G) and lncRNAs (H).

Fig. S2: The correlation coefficient of all liver and lung samples. The correlation coefficient tends toward 1 between samples from the same group indicating that it is reasonable to perform further data analyses

Fig. S3: The circus plot of the chromosomal distribution of DE lncRNAs and mRNAs in (A) liver and (B) lung tissue. The outermost ring represents different chromosomes. From outer ring to inner ring in order as mRNA (AW vs HS), lncRNA (AW vs HS), mRNA (HB vs HS), lncRNA (HB vs HS), mRNA (GJ vs HS), lncRNA (GJ vs HS), mRNA (WT vs HS), lncRNA (WT vs HS). 1< log2FC <5 (orange), 5< log2FC <Inf (yellow), log2FC =Inf (red), -5< log2FC <-1 (purple), -Inf< log2FC <5 (blue), log2FC =-Inf (green).

Fig. S4: Interaction network of DEGs in five sheep populations from different altitude. (A) mRNA–mRNA interaction network of DE mRNAs in liver tissue. (B) mRNA–mRNA interaction network of DE mRNAs in lung tissue. (C) lncRNA–mRNA interaction network in liver tissue. (D) lncRNA–mRNA interaction network in lung tissue. Circles and squares represent mRNAs and lncRNAs, respectively.

Table S1: The primers for mRNAs and lncRNAs.

Table S2: The determination results of the physiological and biochemical indices of five sheep populations.

Table S3: Statistics of clean reads of liver and lung tissue samples from five sheep populations.

Table S4: Identification of lncRNAs and mRNAs in sheep liver and lung tissue.

Table S5: The DE mRNAs and lncRNAs in all groups by pairwise comparison.

Table S6: The DE mRNAs and lncRNAs of the pairwise comparison between Tibetan sheep and Hu sheep (AW vs HS, HB vs HS, GJ vs HS, WT vs HS).

Table S7: Annotation of DE mRNAs and the targeted genes of DE lncRNAs.

Table S8: Significantly enriched GO terms and KEGG pathways of targeted genes of DE lncRNAs.

Table S9: The top 10 DE mRNAs with the highest degree and lncRNAs that interacted with more target genes.

Table S10: qRT-PCR verification of 7 DE lncRNAs, 13 target genes, and 20 DE mRNAs.

Author Contributions

Q.Z. and F. W. planned and designed the research. F.W. reviewed and revised the article. D.Z. and J.L. contributed to sample collection, performed experiments and data collection.

J.L. and F. W. analyzed the data, interpreted results, and wrote the manuscript. Q.Z., F.W. and D.Z. provided suggestions in drafting the manuscript and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Comparative analysis of long non-coding RNA and mRNA expression provides insights into adaptation to hypoxia in Tibetan Sheep

Abstract

Tibetan sheep have lived on the Qinghai-Tibetan Plateau for thousands of years and they have a good adaptability to the hypoxic environment and strong disease resistance. However, the molecular mechanism of the Tibetan sheep adapting to this extreme environment, especially the role of genetic regulation is still unknown. Emerging evidence suggests that long non-coding RNAs (lncRNAs) participate in the regulation of a diverse range of biological processes. To explore the potential lncRNAs involved in Tibetan sheep adapting to high altitude hypoxia environment, we analyzed the expression profile of lncRNAs and mRNAs in liver and lung tissue of sheep based on the comparative transcriptome analysis between four Tibetan sheep populations (high altitude) and Hu sheep (low altitude). The results showed a total of 7848 differentially expressed (DE) lncRNAs transcripts and 22971 DE mRNAs transcripts were detected by pairwise comparison. The expression patterns of selected mRNAs and lncRNAs were validated by qRT-PCR and the results correlated well with the transcriptome data. Moreover, the functional annotation analysis based on the Gene Ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) databases showed that DE mRNAs and the target genes of the lncRNAs were significantly enriched in organ morphogenesis, response to stimulus, heme binding, immune system, arginine and proline metabolism, and fatty acid biosynthesis. The prediction of mRNA-mRNA and lncRNA-mRNA interaction networks further revealed transcripts potentially involved in adaptation to high altitude hypoxia, the hub genes that *DDX24*, *PDCD11*, *EIF4A3*, *NDUFA11*, *SART1*, *PRPF8* and TCONS_00306477, TCONS_00306029, TCONS_00139593, TCONS_00293272, TCONS_00313398 were selected. Additionally, a set of target genes, *PIK3R1*, *IGF1R*, *FZD6*, *IFNB2*, *ATF3*, *MB*, *CYP2B4*, *PSMD13*, and *TGFB1* were also identified as candidate genes associated with high altitude hypoxia adaptation. In conclusion, **a** collection of aberrantly expressed lncRNA, a set of target genes and biological pathways known to be relevant for altitude adaptation were identified by comparative transcriptome analysis between Tibetan sheep and Hu sheep. Our results first identified the characterization and expression profile of lncRNAs between Tibetan sheep and Hu sheep

and provides insights into the genetic regulation mechanisms for Tibetan sheep adaptation to high altitude hypoxia environments.

Keywords: Tibetan sheep; transcriptome; lncRNAs; mRNAs; hypoxia

Abbreviations

LncRNA：Long non-coding RNA;
DEG: Differentially expressed gene;
GO: Gene Ontology;
KEGG: Kyoto Encyclopedia of Genes and Genomes;
FPKM: Fragments per kilo-base of exon per million fragments mapped;
qRT-PCR: Quantitative Real-Time PCR;
VEGF: Vascular endothelial growth factor;
HIF-1: Hypoxia inducible factor-1.

Introduction

Tibetan sheep is one of China's three primitive sheep populations, mainly distributed among the Qinghai-Tibetan Plateau, with Qinghai, Tibet and Gannan in Gansu[1,2]. The combined effects of cold and hypoxia in high altitude areas pose severe physiological challenges for Tibetan sheep. The decrease in oxygen partial pressure at high altitude reduces the oxygen supply to the cells of the respiratory tissue[3]. Decreased tissue oxygenation will severely limit aerobic metabolism and inhibit the ability to produce heat, thereby resisting the effects of the decrease in ambient temperature that accompanies an increase in altitude on the body[4,5]. The cold and hypoxic conditions at high altitudes have seriously affected the metabolic needs of Tibetan sheep[6]. Therefore, it is an interesting scientific question to analyze the adaptive evolution mechanism of Tibetan sheep in high cold and hypoxic environment. To elucidate the molecular mechanism of the Tibetan sheep adapting to high cold and hypoxia is a scientific goal of evolutionary genetics.

Oxygen is the key substance for the animal body to carry out metabolism and maintain life, and it is the first need for life activities[7]. The breathing oxygen is converted into available oxygen in the animal, called blood oxygen. The blood oxygen carried by the

blood inputs energy to the whole body, and the delivery amount of blood oxygen is closely related to the working state of the heart and brain. The stronger the heart's blood pumping ability, the higher the blood oxygen content; the strongly coronary blood transfusion ability of the heart, the higher the concentration of blood oxygen delivered to the heart and brain and whole body, the better operating state of important animal organs[8]. Medically, an area with an altitude of more than 3,000 meters is called a plateau. The impact of the plateau environment on animal organisms involves many aspects, such as atmospheric geography, geochemistry, ecology, and other factors. Factors such as low atmospheric pressure, low oxygen, low temperature, low humidity, solar radiation, airflow, rainfall, light, wind, and snow often act on animal organisms. Among them, the impact of plateau hypoxia has attracted wide attention to scholars because of its most significant impact[3,9].

In terms of molecular research on hypoxia adaptation, studies have shown that hypoxia inducible factor-1 (*HIF-1*) is one of the genes being selected by the high altitude and low oxygen environment[10,11]. The *HIF-1α* subunit is encoded by the *HIF-1A* gene, mainly regulated by oxygen concentration and determines the level of *HIF-1α* activity[12]. The expression level of *HIF-2α* mRNA in Tibetan sheep is higher in lung, liver, kidney, and myocardial tissues than that in low altitude sheep. It can quickly induce the expression of its downstream factors *VEGF* (Vascular endothelial growth factor), *EPO* (Erythropoietin), lung surfactant, strengthen the lipid oxidative phosphorylation reaction, and thereby enhance the lung ventilation capacity, heart pumping function, kidney *EPO* synthesis and liver metabolic reaction to meet the oxygen and energy required for body activities. The expression level of *STAT3* mRNA in Tibetan sheep is higher in lung, liver, and myocardial tissues than that in low altitude sheep[13]. It indicates that the hypoxia environment induces high expression of *STAT3* protein to promote the generation of activated *STAT3*, which can effectively regulate the transcription of *VEGF*, *COX-2*, *MnSOD*, *UCP1*, *HIF-1α* and its downstream target genes, thereby enhancing the ventilation capacity of Tibetan sheep lungs, hypoxic adaptability of heart pumping function and energy metabolism[14].

At present, a number of scientists have studied the mechanism of hypoxic adaptation in the Tibetan Plateau at the genetic level, but the molecular mechanism of Tibetan sheep

adapting to this extreme environment, especially the role of genetic regulation is still unknown. Long non-coding RNAs (lncRNAs) are transcripts longer than 200 nucleotides that can regulate mRNA expression of both posttranscriptional and transcriptional levels[15]. There is limited literature on the detailed functional roles of lncRNAs in high altitude adaptation of Tibet[16-18]. The role of lncRNA in Tibetan sheep's hypoxic adaptation has not been reported yet. The latest research progress of high-throughput sequence technology has opened up a new path to exploring the molecular genetic basis of adaptive physiological traits. Transcriptome sequencing (mRNA and lncRNA) can not only find important genes that cope with physiological challenges, but also provide an in-depth understanding of the role of transcription regulation in adapting to evolutionary variation.

In this study, the genetic expression and regulation mechanisms of different populations and tissues in the Tibetan sheep under different altitude conditions were investigated based on the comparative transcriptome analysis. Functional genes associated with high altitude hypoxia adaptation in Tibetan sheep were found by the perspective of mRNA, lncRNA and their co-expression networks analysis.

Methods

Ethic statement

All the experimental procedures mentioned in the present study were approved by the Science Research Department (in charge of animal welfare issue) of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (IAS-CAAS). We have obtained written informed consent to use the animals in this study from the owners of the animals and all experiments on animals were conducted under a permit approved by the ethics committee of Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences (Approval No. SYXK-2014-0002). All animals were handled in strict accordance with the recommendations in the Regulations for the Administration of Affairs Concerning Experimental Animals of the State Council of the People’s Republic of China. This study was carried out in compliance with the ARRIVE guidelines.

Animal materials and sampling

Five sheep (*Ovis aries*) populations, including Huoba (HB, altitude 4468 m), Awang (AW, altitude 4452 m), Ganjia (GJ, altitude 3851 m) and Qilian (WT, altitude 3621 m) Tibetan sheep on the Qinghai-Tibet Plateau and Hu sheep (HS, altitude -67 m) were obtained from individual farmers in this study. All the experimental sheep were raised under an environment with natural light and free access to food and water. The adult individuals (rams, aged 3~4 years) within each population were randomly selected for this study. More than five sheep within each population were used for physiological and biochemical characteristics analyses. All experimental animals were anesthetized by pentobarbital sodium, and then were put to death after bloodletting. Three sheep within each population were selected to obtain the liver and lung samples for transcriptome analysis (Table 1), and all the samples were immediately snap-frozen in liquid nitrogen for total RNA extraction.

Physiological and biochemical measurement

To observe the changes of related blood physiological and biochemical parameters under hypoxia at high altitude in sheep, we sampled Tibetan sheep and Hu sheep from different altitudes. Jugular venous blood samples from 3~4 years adult sheep that lived at different altitudes as described in Table 1 were collected to measure the haematological parameters: red blood cell count (RBC), white blood cell count (WBC), haemoglobin contents (HGB) and others using a HC-3000 Auto Hematology Analyzer (Jinan Meiyyilin Electronic Instrument Co., Ltd, Jinan, China China). Moreover, the blood gas indices including partial pressure of carbon dioxide (PCO_2), oxygen partial pressure (PO_2), oxygen saturation (O_2S), standard base excess (SBE) and others were repeatedly measured on an IL 1302 pH/ Blood Gas Analyzer (Instrumentation Laboratories, MA). For the analysis of blood biochemical indices, the blood samples were centrifuged at 3000 *g* for 10 min, after which the supernatant was collected, the serum biochemical parameters: alanine aminotransferase (ALT), aspartate aminotransferase (AST), total protein (TP), plasma

cholinesterase (PChE) and others were analyzed with a Dirui CS-600 automatic biochemical analyzer (Dirui Industrial Co., Ltd; Changchun, China).

After euthanasia, lung organ was taken out immediately and excised into 7 pieces tissue from the same part of the left lung of each sheep. The lungs were washed with 0.9% saline, and then fixed with Bouin's solution containing 75 mL saturated picric acid solution, 25 mL formaldehyde, and 5 mL glacial acetic acid for 48 h at room temperature. Next, the tissues were embedded in paraffin and 4-μm-thick sections were made for hematoxylin-eosin (HE) staining. The morphology of lung tissues was observed under a virtual microscope (Olympus, BX51, Japan)[19].

RNA isolation, library construction and sequencing

Total RNA was extracted using TRIzol® Reagent (Life Technologies) according to the manufacturer's protocol. RNA degradation and contamination, especially DNA contamination, was monitored on 1.5% agarose gels. RNA concentration and purity were measured using the NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 System (Agilent Technologies, CA, USA).

A total amount of 1.5 μg RNA per sample was used as input material for rRNA removal using the Ribo-Zero rRNA Removal Kit (Epicentre, Madison, WI, USA). Sequencing libraries were generated using NEBNext R UltraTM Directional RNA Library Prep Kit for Illumina R (NEB, USA) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. Briefly, Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). First strand cDNA was synthesized using random hexamer primer and reverse transcriptase. Second-strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure was ligated to prepare hybridization. In order to select insert fragments of preferentially 150~200 bp in length, the library fragments were

purified with AMPure XP Beads (Beckman Coulter, Beverly, USA). Then 3 μL USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37 °C for 15 min before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index(X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 and qPCR[20]. The libraries were sequenced on an Illumina Hiseq 2500 platform (Illumina Inc., San Diego, CA, USA) according to the manufacturer's instructions and 125 bp paired-end reads were generated.

Quality control and RNA-seq data analysis

Raw data (raw reads) of fastq format were firstly processed through in-house perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing poly-N and low-quality reads from raw data. At the same time, Q20, Q30, GC-content and sequence duplication level of the clean data were calculated. All the downstream analyses were based on clean data with high quality[21,22]. The clean reads were aligned to the genome using TopHat version 2.0.9. Mapped reads from TopHat for each sample were assembled using Cufflinks vision 2.1.1. The multiple assembled transcript files (GTF format) for different samples were then merged together to produce a unique transcriptome set using the Cuffmerge utility provided by Cufflinks package.

LncRNA analysis and quantification of gene expression levels

The transcriptome was assembled using the Cufflinks and Scripture based on the reads mapped to the reference genome. The assembled transcripts were annotated using the Cuffcompare program from the Cufflinks package. The unknown transcripts were used to screen for putative lncRNAs. Four computational approaches including CPC/ CNCI/ CPAT/ Pfam were combined to sort non-protein-coding RNA candidates from putative protein-coding RNAs in the unknown transcripts. Putative protein-coding RNAs were filtered out using a minimum length and exon number threshold. Transcripts with lengths more than 200 nt and have more than two exons were selected as lncRNA candidates and further screened using CPC/ CNCI/ CPAT/ Pfam that has the power to distinguish the

protein-coding genes from the non-coding genes. As well as the different types of lncRNAs including lincRNA, intronic lncRNA, anti-sense lncRNA were selected using cuffcompare.

Cuffdiff (v2.1.1) was used to calculate fragments per kilo-base of exon per million fragments mapped (FPKMs) of both lncRNAs and coding genes in each sample[23]. Gene FPKMs were computed by summing the FPKMs of transcripts in each gene group based on the length of the fragments and reads count mapped to this fragment.

Differential expression analysis

Differential expression analysis of five groups was performed using the DESeq R package (1.10.1). DESeq provides statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting *P*-values were adjusted using Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted *P*-value <0.01 and absolute value of log2 (Fold change) >1 found by DESeq were assigned as differentially expressed.

Prediction of potential target genes of lncRNAs

To explore the roles of lncRNA in Tibetan sheep adaption to hypoxia at high altitude, lncRNA-targeted genes were predicted based on cis and trans principle. Cis action indicated a positional relationship, representing a gene located within the range of 100 kb from the lncRNA. Trans action meant that there were complementary sequences between the mRNA and lncRNA; the sequences of mRNAs that overlapped with lncRNAs were predicted by LncTar software[24]. These two interaction mechanisms were considered preferential for the prediction of lncRNA-targeted genes.

Functional enrichment analysis

Gene functional enrichment analysis included Nr (NCBI non redundant protein sequences), COG (clusters of orthologous groups of proteins), Swiss-Prot (a manually annotated and reviewed protein sequence database), GO (Gene Ontology; http://www.geneontology.org/), and KEGG (Kyoto Encyclopedia of Genes and Genomes[25],

which were utilized for screening genes related to highland hypoxia. GO enrichment analysis of the differentially expressed genes was implemented by the topGO R packages. KOBAS software[26] was used to test the statistical enrichment of differentially expressed genes (DEGs) in KEGG pathways.

Interaction Network Construction

Protein-protein interaction analysis of DEGs was based on the commonly used STRING database[27]. Briefly, the sequences of differentially expressed lncRNA targets were blasted (blastx) to *Ovis aries* genome to predict protein-protein interactions using the STRING database. Then, the mRNA-mRNA and lncRNA–mRNA interaction networks were visualized by using Cytoscape version 3.5.1[28].

Quantitative Real-Time PCR (qRT-PCR) Validation

Several DE lncRNAs, target genes and mRNAs putatively associated with hypoxia responses at high altitudes were selected and confirmed by qRT-PCR with GAPDH used as an internal reference. The primers used for qRT-PCR are listed in Table S1. The qRT-PCR was carried out with a Roche LightCycler® 96 using iTaq™ Universal SYBR® Green Supermix (Bio-Rad, United States). Each real-time RT-PCR reaction (in 25 μ L) involved 12.5 μL 2× SYBR Green Realtime PCR Master Mix (TaKaRa, Dalian), 1 μL of each primer, 2 μL cDNA and 8.5 μL H_2O. The amplification procedures were 95 ℃ for 5 min initially, followed by 45 cycles of 95 ℃ for 15 s and 60 ℃ for 1 min. Quantification of mRNAs and lncRNAs was performed using the standard curve method with average cycle thresholds (Ct). The qRT-PCR data were generated from three independent samples per group. The correlation between the results of RNA-seq and qPCR was calculated using a correlation test.

Statistical analysis

All data was presented as the mean ± SD. The Statistical Program for Social Sciences (SPSS) 20.0 software was employed to perform all of the statistical analyses. The

multi-group comparisons of the means were analyzed using one-way analysis of variance. Statistically significant differences were determined at p-value < 0.05.

Results

The physiological and biochemical indices of five sheep populations from different altitude region

The measurement results of physiological and biochemical indices are shown in Table S2 and Fig. 1. The results show that he lung of Tibetan sheep at high altitude regions has typical characteristics in tissue structure. Compared with Hu sheep from low-altitude regions, the terminal bronchioles of Tibetan sheep showed larger diameter, the thickness of the alveolar septum increased, and the number of alveoli and blood vessels has increased per unit area (P<0.05). Furthermore, the measurement results from blood gas indices of Tibetan sheep at different altitudes found that the PO_2, O_2S, and SBE decreased significantly as the altitude increased (P<0.05).

In addition, the results showed that the HGB, RBC, white WBC, and other indices of Tibetan sheep at high altitude were significantly higher than those of low-altitude Hu sheep (P<0.05), and the red blood cell distribution width coefficient of variation (RDW-CV) became smaller. The measurement results from serum biochemical indices showed that AST, TP, and others increased with altitude (P<0.05), while ALT and PChE decreased.

Identification of lncRNAs and mRNAs in sheep liver and lung tissue

The Illumina sequencing of 30 sample's cDNA libraries derived from five sheep populations yielded a total of 187.28 Gb and 190.63 Gb clean data for liver and lung tissue, respectively. At least 78.05% of the clean reads from each sample was mapped to the sheep reference genome (Oar_v3.1, ftp://ftp.ensembl.org/pub/release-78/fasta/ovis_aries/), of which more than 78.52% were uniquely mapped (Table S3).

In liver samples, a total of 6,249 lncRNAs was identified after coding potential filters using CNCI, CPC, CPAT and Pfam-scan software, 20,078 mRNAs and 2,728 novel transcripts were identified. In addition, a total of 1,798 lncRNAs, 13,319 mRNAs and 750

novel transcripts were identified from lung samples (Table S4). A majority of lncRNAs comprised two or three exons, whereas mRNAs contained a broad range of exon numbers from two to thirty. The transcript length and ORF length of lncRNAs were significantly shorter than those of mRNAs. Moreover, the expression levels of mRNAs and lncRNAs were further analyzed using FPKM. The mRNA transcript levels were higher than those of lncRNAs in both liver (Fig. S1) and lung (Fig. 2) tissue.

DE lncRNA and mRNA between Hu sheep and Tibetan sheep

The correlation coefficient tends to 1 between samples of the same group, thus indicating that it is reasonable to perform further data analyses (Fig. S2). Based on a fold change of ≥2 and false discovery rate of <0.05, in total, 7,848 DE lncRNA transcripts and 22,971 DE mRNA transcripts were detected in all groups by pairwise comparison (Table S5).

Furthermore, we focused on the analysis of the pairwise comparison between Tibetan sheep and Hu sheep (AW vs HS, HB vs HS, GJ vs HS, WT vs HS). We found that most of the DE genes were enriched on chromosomes 1, 2, 3, and chromosome X (Fig. 3, Fig. S3 and Table S6). The analysis of common DE genes between liver and lung by pairwise comparison found 3 common DE lncRNAs and 316 common DE mRNAs (Fig. 4A and B). Further analysis of common differentially expressed genes (DEGs) among Tibetan sheep and Hu sheep groups, 2 common DE lncRNAs and 99 common DE mRNAs in liver were detected by pairwise comparison (Fig. 4C and D). Additionally, 1 common DE lncRNA transcripts and 63 common DE mRNA transcripts were detected in lung (Fig. 4E and F).

Functional Annotation of DE mRNAs and the target genes of DE lncRNAs

The overall functional annotation of DE mRNAs in pairwise comparison (AW vs HS, HB vs HS, GJ vs HS, WT vs HS) is described in Table S7. As lncRNAs could exert effects on cis-acting or trans-acting target genes, we predicted 1,739 target genes of DE lncRNAs that might be potentially involved in the adaptation to high altitudes based on the cis and trans RNA-RNA interaction principle. In total, 218 and 197 target genes were differentially expressed and annotated from liver and lung tissue by pairwise comparison (Table S7).

To elucidate the functions of DE mRNAs, GO enrichment analysis was performed using topGO to search the most significant GO term of DE mRNAs. The most enriched GO terms of the DE mRNAs in each comparison are shown in Table S8. All of the DE mRNAs were assigned to GO terms of biological processes, cellular components, and molecular function. Moreover, response to stimulus, tissue development, skeletal system development, response to stress, regulation of signal transduction, organ development, antioxidant activity, calcium ion binding, heme binding, and lipid antigen binding were identified as significantly enriched GO terms. According to the KEGG analysis, the most enriched pathway to DE mRNAs (P<0.05) is predominantly associated with amino acid metabolism, lipid metabolism and immune system, such as arginine and proline metabolism, hematopoietic cell lineage, fatty acid biosynthesis, glutathione metabolism, and biosynthesis of unsaturated fatty acids.

LncRNA-mRNA co-expression networks construction

In order to understand the effects of lncRNAs on the regulation of the target genes, we constructed mRNA-lncRNA regulatory network. Firstly, we selected the target genes of DE lncRNAs to construct functional networks by referring to the STRING database, with each gene corresponding to a node. Two genes were connected by an edge, indicating a correlation between mRNA-mRNA. Next, DE lncRNAs and their corresponding target genes were used to construct the lncRNA-mRNAs interaction network. Then, the mRNA-mRNA and lncRNA-mRNAs network were merged (Fig. S4). Within the network analysis, we focused on the top 10 DE mRNAs with the highest degree and lncRNAs that interacted with more target genes (Table S9), constituting probably the core of the network. The results implied that DE mRNAs *DDX24*, *PDCD11*, *EIF4A3*, *NDUFA11*, *SART1*, *PRPF8* and DE lncRNAs TCONS_00306477, TCONS_00306029, TCONS_00139593, TCONS_00293272, and TCONS_00313398 might play key roles in adaptation to hypoxia. Although these lncRNAs requires further experimental validation, this information might be helpful to explore the potential mechanisms involved in the adaptation to hypoxia at high altitudes.

The candidate genes screen for the adaptation to hypoxia at high altitudes

A set of candidate genes that were putatively associated with hypoxia responses to high altitudes were screened from the target genes of DE lncRNAs. For a detailed description of the relevant functions and pathway to candidate genes see Table 2. Among them, three (*PIK3R1*, *IGF1R* and *PDK1*) were located in the classical HIF-1 pathway; five were found in the corresponding downstream vascular endothelial growth factor (*VEGF*, *FZD4* and *IFNB2*) and glycolysis/gluconeogenesis pathways (*ATF3*, *PPCK1* and *PFKFB2*). The dysregulation of genes in these relevant pathways indicated that hypoxia-induced factors, angiogenesis, and glycolysis metabolism are the most important factors that allowed sheep to manage extreme hypoxic environmental pressure. We also found six genes (*MB*, *CYP2C31*, *CYP2B4*, *CYP2B5*, *CYP1A1*, and *CYCS*) were functionally involved in oxygen binding, oxygen transport, and heme binding.

Moreover, we found the GO terms of the target genes were primarily associated with metabolic processes (amino acid, lipid, and fatty acid), regulation of immune system process, immune response, biological regulation, response to stress and response to stimulus (Table S7). These functional terms are biologically relevant to the plateau adaptations because they are involved in energy metabolism, immune function, oxidation reaction, and stress response, which are important regulating factors of Tibetan sheep response to extreme hypoxic environments.

Validation of DE mRNAs and lncRNAs by qRT-PCR

To evaluate the reliability of RNA sequencing, 7 DE lncRNAs, 13 target genes, and 20 DE mRNAs were selected and validated by qRT-PCR from different sheep groups. All levels of DE lncRNAs and mRNAs were consistent with the RNA-seq results, indicating that the RNA-seq data were reliable (Table S10). The result of qRT-PCR and RNA-seq for the common DE lncRNA TCONS_00139593 and TCONS_00332125 in liver and TCONS_00377466 in lung are shown in Fig. 5. Additionally, the confirmed results of 3 DE mRNAs that were enriched in heme binging term (*PTGS2* and *LOC101107056*) and HIF-1 pathway (*TFF3*) are also shown in Fig. 5.

Discussion

Comparison of physiological and biochemical indices

At present, the research on the adaptability of Tibetan sheep to hypoxia at high altitude has made important progress in histology, morphology, physiology, and anatomy[29-32]. Studies suggest that compared to Tan sheep living at low altitude areas, Tibetan sheep have developed a cerebral arterial system, the main arteries are thicker in diameter, and the collateral branches in the cerebral arteries are developed and stretched longer. There are many small arteries, and this feature is conducive to effective blood supply to the brain tissue and the regulation of cerebral arterial blood pressure[33]. This might be one of the anatomical characteristics of Tibetan sheep aiding in adapting to high altitude hypoxia environment. Anatomical studies on the vascular system of other tissues and organs of Tibetan sheep have similar results. For example, compared to small-tailed Han sheep, Tibetan sheep have more capillaries in the alveolar septum, and they are mostly open, which also increases with altitude. The alveolar septum is thick, indicating that the alveolar septum is rich in capillaries and elastic fibers. These structural features are conducive to increasing alveolar ventilation, increasing pulmonary blood flow, accelerating blood oxygen transport, and improving lung gas exchange rate in a hypoxic environment. Compared to low-altitude sheep, Tibetan sheep have more red blood cells and higher hematocrit and hemoglobin content. Under low oxygen environment, Tibetan sheep mainly adapt to the low oxygen environment by increasing the hemoglobin content of the blood[34].

In this study, we first examined the haematological changes and serum biochemical parameters in four Tibetan sheep and one Hu sheep population. In agreement with previous reports, the haematological parameters, serum biochemical parameters, the blood gas indices and the morphology of lung tissues showed significant changes between Tibetan sheep (high altitude) and Hu sheep (low altitude). The haematological parameters including RBC, WBC, HGB, HCT, MCV, and PLT became significantly higher as the altitude increased (P <0.05). The reason for the difference might be due to the high cold and hypoxic factors, especially the value of HGB increases from the increase of altitude. Under

low PO_2, HGB dissociates from oxygen to provide the body of the oxygen needed for energy metabolism to better adapt to the low oxygen environment. The biochemical parameters including AST, TP, ALB, GLO, ALP, and LDH significantly increased with increasing altitude, while the ALT and PCHE decreased from increasing altitude. In particular, the values of TP increased to the altitude, which is helpful to enhance the immune function of sheep and adapt to the high-altitude ecological environment. Moreover, the blood gas indices including PCO_2, PO_2, O_2S, SBC, TCO_2, and SBE all significantly decreased with increasing altitude. Related studies have shown that Tibetan sheep can reduce tissue oxygen demand and cell metabolism levels through specific physiological changes, and adapt to the plateau hypoxia environment at the molecular level by regulating hypoxia inducible factor[13]. Additionally, the morphology of lung tissue was observed and we found that the terminal bronchioles, the number of alveolar counted per unit area, the alveolar septum thickness and the number of vessels per unit area significantly increased with increasing altitude. These changes in tissue structure are conducive to accelerating blood oxygen transport and increasing alveolar ventilation, increasing Tibetan sheep's lung blood flow and the lung gas exchange rate of Tibetan sheep in a hypoxic environment to a certain extent. Early studies have shown that these changes are the key characteristics of Tibetan sheep adapting to high altitude environments[35,36].

Analysis of lncRNAs and their target genes

Owing to the key roles of lncRNAs in many important biological processes, these are currently of particular interest[37,38]. The rapid development of high throughput sequencing methods had led to the discovery of thousands of lncRNAs in recent years. The studies have reported that the lncRNAs involved in primary wool follicle induction in carpet wool sheep[39], sheep fat-tail development[40], sheep skeletal muscle development[41], prolificacy in Hu sheep[42] and sheep testicular maturation[19,39] with high throughput sequencing technology. But expression and function of lncRNAs in Tibetan sheep adapting to high altitude hypoxia are still unclear. To provide some insights into the biological functions of lncRNAs in Tibetan sheep adaption to high altitude hypoxia, a comprehensive analysis of lncRNA and

mRNA profiling data from Tibetan sheep and Hu sheep, together with data from a public database was performed. We identified the core lncRNAs and their target genes, and validated their expression by qRT-PCR. Overall, our work uncovered an interlaced transcripts network that is involved in high altitude hypoxia environment.

By the analysis of common DE genes among Tibetan sheep and Hu sheep groups, 2 common DE lncRNAs TCONS_00139593 and TCONS_00332125 in liver and 1 common DE lncRNA TCONS_00377466 in lung were found. Moreover, the lncRNA–mRNA interaction network of liver sample showed that TCONS_00306477, TCONS_00306029, TCONS_00029720, TCONS_00145870, TCONS_00139593, TCONS_00380986, TCONS_00309307, TCONS_00225957, TCONS_00321529, and TCONS_00100469 interacted with more target genes and suggested as hub genes that related to high altitude hypoxia adaptation; the lncRNA–mRNA interaction network in lung sample showed that TCONS_00293272, TCONS_00313398, TCONS_00344932, TCONS_00078812, TCONS_00352306, TCONS_00380999, TCONS_00088235, TCONS_00467816, TCONS_00078180, and TCONS_00315164 interacted with more target genes and suggested as hub genes.

For the research of candidate genes that are associated with hypoxia responses at high altitudes, an early research reported genome-wide scans that revealed positive selection in several regions that contained genes whose products are likely to be involved in high altitude adaptation[43]. Finally, a set of 247 functional candidate genes were identified. The functional candidate genes categories included detection of oxygen (GO: 0003032), NO metabolic process (GO: 0046209), oxygen sensor activity (GO: 0019826), oxygen binding (GO: 0019825), oxygen transport (GO: 0015671), oxygen transporter activity (GO: 0005344), response to hypoxia (GO: 0001666), response to oxygen levels (GO: 0070482), Vasodilation (GO: 0042311), and hypoxia response via *HIF* activation (P00030) in panther pathway. In this study, we found that the target genes including *MB*, *PIK3R1*, *CYP1A1*, *MMP14*, and *TGFB1* belong to the list of 247 hypoxia genes. In addition, the candidate genes (*MMP14*[44], *TUBB4B*[45], *PSMD13*[46], *COL3A1*, *COL1A2*[47], *DSG3*[48,49], and *ATP6*[50]

were also identified as candidate genes associated with high altitude adaptation by previous functional studies.

Myoglobin, encoded by *MB*, is a haemoprotein present in cardiac, skeletal and smooth muscle and serves as a reserve supply of oxygen and facilitates the movement of molecular oxygen from the cell membrane to mitochondria[51]. The study has demonstrated that *PIK3R1* involved in the *HIF-1α* signaling pathway plays a critical role in mediating adipose tissue insulin sensitivity[52]. Previous study shows *CYP1A1* transcriptional activation was significantly decreased upon *PCB 126* stimulation under conditions of hypoxia. Additionally, hypoxia pre-treatment reduced *PCB 126* induced AhR binding to *CYP1* target gene promoters[53]. Moore research showed that *MMP14* is upregulated in hypoxic conditions and that this occurs by the interaction of *HIF-1α* and the *MMP14* gene promoter region[54]. Chen *et al* suggested that *TGF-b1* encoded by *TGFB1* decreases hypoxia-reoxygenation injury and attenuates alterations in NOS and PKB phosphorylation in myocytes exposed to hypoxia-reoxygenation[55].

Yang et al[56] generated whole-genome sequences from 77 native sheep and detected a novel set of candidate genes as well as pathways and GO categories that were putatively associated with hypoxia responses at high altitudes. Specifically included several positively selected genes within or regulating the *HIF-1* pathway, the *VEGF* pathway, the *VSMC* pathway, and glycolysis and lipid for energy metabolism. The network of relevant pathways indicated that hypoxia-induced factors, angiogenesis, vasodilatation and glycolysis metabolism were the most important factors that allowed sheep to manage extreme hypoxic environmental pressure. Seven sheep breeds representing both highland and lowland breeds from different areas of China were genotyped for a genome-wide collection of single-nucleotide polymorphisms (SNPs)[57]. Then detected selection events spanning genes were involved in angiogenesis, energy production and erythropoiesis played a crucial role in hypoxia adaption. Here, we found the target genes that *PIK3R1*, *IGF1R* and *PDK1* in the classical *HIF-1* pathway and *FZD4*, *IFNB2, ATF3*, *PPCK1*, *PFKFB2* in the corresponding downstream *VEGF* and glycolysis/gluconeogenesis pathways, which played a central role in regulating cellular responses to hypoxia[43,58,59].

546 Hypoxia regulates *IGF1* expression through *HIF-1α*, and the inhibition of *HIF-1α* or
547 *IGF1R* decreased CD133- and Oct4-positive GRPs under hypoxia[60]. Mora et al found that
548 the *PDK1* signalling network plays an important role in regulating cardiac viability and
549 preventing heart failure, the deficiency of *PDK1* in cardiac muscle results in heart failure
550 and increased sensitivity to hypoxia[61]. *ATF3* is a stress-induced transcription factor that
551 plays important roles in regulating immune and metabolic homeostasis. Overwhelming
552 evidence confirms that the *ATF3* gene is activated in many tissues by a variety of stress
553 signals, including proinflammatory cytokines, ischemia and hypoxia[62]. Parra *et al* found
554 that the mRNA levels of glycolytic markers *HK2*, *PFKFB2* and *GLUT1* increased
555 accordance with a metabolic shift towards non-mitochondrial ATP generation during
556 hypoxia[63]. The *VEGF* pathway downstream of *HIF-1* and glycolysis is an important
557 mechanism of energy metabolism for sheep under extreme hypoxic conditions. The
558 dysregulation of genes in these pathways indicated that hypoxia-induced factors,
559 angiogenesis, and glycolysis metabolism were the most important factors that allowed
560 sheep to manage extreme hypoxic environmental pressure.

561 The genes *CYP2C31*, *CYP2B4*, *CYP2B5*, and *CYCS* were functionally involved in
562 oxygen binding, oxygen transport, and heme binding. In humans, indirect evidence
563 suggests that hypoxia reduces the rate of biotransformation of drugs cleared by cytochrome
564 P450 subfamilies *CYP1A*, *2B*, and *2C*. Fradette *et al* found that hypoxia down-regulates
565 rabbit hepatic *CYP1A1*, *1A2*, *2B4*, *2C5*, and *2C16* and up-regulates *CYP3A6*. *CYP3A11* and
566 P-glycoprotein were up-regulated in the livers of hypoxic rats[64]. In addition, the genes
567 (*TUBB4B*, *PSMD13*, *COL3A1*, *COL1A2*, *DSG3,* and *ATP6*) were also identified as
568 candidate genes associated with high altitude adaptation by previous functional studies.
569 Kharrati-Koopaee *et al* found that *PSMD13* gene was associated with the hypoxia by whole
570 genome sequencing of lowland and highland chickens[46]. Qi *et al* conducted a cross-tissue,
571 cross-altitude, and cross-species study to characterize the transcriptomic landscape of
572 domestic yaks. They found that lung and heart are two key organs showing adaptive
573 transcriptional changes, five of collagen genes (*COL1A2*, *COL3A1*, *COL5A2*, *COL14A1*,
574 and *COL15A1*) highlighting the crucial role of collagen involved pathways in high altitude

adaptation[47]. The previous exome sequencing of five Chinese cashmere goat breeds revealed a candidate gene, *DSG3*, responsible for high altitude adaptation of the Tibetan goat. And the mutations significantly segregated high- and low-altitude goats in two clusters, indicating the contribution of *DSG3* to the high altitude hypoxia adaptation in the Tibetan goat[49]. Wang *et al* sequenced the *ATP8* and *ATP6* genes in 66 Tibetan yaks and 81 domestic cattle found that haplotypes H4 in *ATP8* and H5 in *ATP6* present only in Tibetan yaks were suggested to be positively associated with high altitude adaptation[50].

Overall, the expression profile of lncRNAs and mRNAs in liver and lung tissue based on the comparative transcriptome analysis between high- and low- altitude sheep indicate that lung and liver are two key organs showing adaptive transcriptional changes. Moreover, the candidate genes involved in HIF-1, VEGF, and glycolysis/gluconeogenesis pathways, as well as oxygen binding, oxygen transport, and heme binding molecular function that were putatively associated with hypoxia responses at high altitudes were screened. These findings, in combination with the results of physiological and biochemical indices analysis, are valuable to understanding the genetic mechanism of hypoxic adaptation in sheep.

Conclusions

In summary, we demonstrated the expression profiles of mRNAs and lncRNAs in Tibetan sheep and Hu sheep in order to understand their regulatory roles in Tibetan sheep adaption to high altitude hypoxia environment. A collection of aberrantly expressed lncRNA, a set of target genes and biological pathways known to be relevant for altitude adaptation were identified by comparative transcriptome analysis between Tibetan sheep and Hu sheep. Our study might contribute to find out the core DE lncRNAs between Tibetan sheep and Hu sheep. In addition, further study of these lncRNAs could provide useful insights into the genetic regulation mechanism of lncRNAs in Tibetan sheep adaption to high altitude hypoxia environment.

Availability of data and materials

The datasets generated and/or analyzed during the current study are included in this published article and its supplementary information files. Genomic resources Oar_v3.1 for *Ovis aries* are available from Ensembl (//ftp.ensembl.org/pub/release-78/fasta/ovis_aries/).

Acknowledgments

This work was financial supported by the National Natural Science Foundation for General Program of China (Grant No. 31872981), the Major Science and Technology Projects of Tibet Autonomous Region (Grant No. XZ202101ZD000N), the National Key Research and Development Program of China (Grant No. 2018YFD0502100), and the Modern Agricultural Technology Research System of Gansu (Grant no. GARS08). These agencies played no role in either the design of the study, the collection, analysis, and interpretation of the data, or the writing of the manuscript.

Declaration of Competing Interest

The authors declare that they have no competing interests.

References

1 Yang, X. J. & Zhao, Y. Z. Review on the genetic diversity of tibet sheep. *J Gansu Agric. Univ.* **37**, 395-400 (2002).

2 Yang, C. *et al.* Relationship between chemical composition of native forage and nutrient digestibility by Tibetan sheep on the Qinghai–Tibetan Plateau. *J. Anim. Sci.* **96**, 1140-1149 (2018).

3 Pagani, L. *et al.* High altitude adaptation in Daghestani populations from the Caucasus. *Hum. Genet.* **131**, 423-433 (2012).

4 Cheviron, Z. A. & Brumfield, R. T. Genomic insights into adaptation to high-altitude environments. *Heredity* **108**, 354-361 (2012).

5 Sasaki, R., Masuda, S. & Nagao, M. Erythropoietin: Multiple Physiological Functions and Regulation of Biosynthesis. *Biosci. Biotechnol. Biochem.* **64**, 1775-1793, doi:10.1271/bbb.64.1775 (2014).

6 Salhan, A. K., Sharma, S. K., Chauhan, S. & Oza, M. Improved Habitability Under Extreme Environments at High Attitude. *Tanslational Research in Environmental and Occupational Stress*, pp.193-204 (2014).

7 Wang, X. X. *et al.* The Research Progress of Clinical Application of High Oxygen. *Prog. Mod. Biomed.* **14**, 2586-2589. (2014).

8 Shi, H. & Liu, K. Cerebral tissue oxygenation and oxidative brain injury during ischemia and reperfusion. *Front Biosci* **12**, 1318-1328 (2007).

9 Hochachka, P. W., Rupert, J. L. & Monge, C. Adaptation and conservation of physiological systems in the evolution of human hypoxia tolerance. *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.* **124**, 1-17 (1999).

10 Semenza, G. L. HIF-1 and mechanisms of hypoxia sensing. *Curr. Opin. Cell Biol.* **13**, 167-171 (2001).

11 Wang, G. L., Jiang, B. H., Rue, E. A. & Semenza, G. L. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *P Natl Acad Sci. USA* **92**, P.5510-5514 (1995).

12 Zhang, P. *et al.* Rac1 activates HIF-1 in retinal pigment epithelium cells under hypoxia. *Graefe's Arch. Clin. Exp. Ophthalmol.* **247**, 633-639 (2009).

13 Zhang, F., Wu, W., Deng, Z., Zheng, X. & Wang, X. High altitude increases the expression of hypoxia-inducible factor 1α and inducible nitric oxide synthase with intest-inal mucosal barrier failure in rats. *In. J Clin Exp. Pathol.* **8**, 5189-5195 (2015).

14 Saldana, M. J., Salem, L. E. & Travezan, R. High altitude hypoxia and chemodectomas. *Hum Pathol* **4**, 251-263, doi:10.1016/s0046-8177(73)80012-7 (1973).

15 Kornienko, A. E., Guenzl, P. M., Barlow, D. P. & Pauler, F. M. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol* **11**, 59-59, doi:10.1186/1741-7007-11-59 (2013).

16 Li, W., Li, J., Wang, R., Xie, H. & Jia, Z. MDR1 will play a key role in pharmacokinetic changes under hypoxia at high altitude and its potential regulatory networks. *Drug Metab. Rev.* **47**, 191-198 (2015).

17 Shang, P. *et al.* Identification of lncRNAs and Genes Responsible for Fatness and Fatty Acid Composition Traits between the Tibetan and Yorkshire Pigs. *Int J Genomics* **2019**, 5070975, doi:10.1155/2019/5070975 (2019).

18 He, X. *et al.* Long Non-coding RNA LINC-PINT and LINC00599 Polymorphisms are Associated With High-altitude Pulmonary Edema in Chinese. *Arch Bronconeumol* **56**, 360-364, doi:10.1016/j.arbres.2019.09.021 (2020).

19 Yang, H. *et al.* Comprehensive analysis of long noncoding RNA and mRNA expression patterns in sheep testicular maturation. *Biol Reprod* **99**, 650-661, doi:10.1093/biolre/ioy088 (2018).

20 Zuo, J. *et al.* Analysis of the Coding and Non-Coding RNA Transcriptomes in Response to Bell Pepper Chilling. *Int. J. Mol. Sci.* **19**, 2001- (2018).

21 Jiang, F., Chen, X. P., Hu, W. S. & Zheng, S. Q. Identification of differentially expressed genes implicated in peel color (red and green) of Dimocarpus confinis. *Springerplus* **5**, 1088, doi:10.1186/s40064-016-2743-y (2016).

22 Ai, Y. *et al.* Transcriptomic Analysis of Differentially Expressed Genes during Flower Organ Development in Genetic Male Sterile and Male Fertile Tagetes erecta by Digital Gene-Expression Profiling. *Plos One* **11**, e0150892 (2016).

23 Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. biotechnol.* **28**, 511-515, doi:10.1038/nbt.1621 (2010).

24 Li, J. *et al.* LncTar: a tool for predicting the RNA targets of long noncoding RNAs. *Brief Bioinform* **16**, 806-812, doi:10.1093/bib/bbu048 (2015).

25 Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44-57 (2008).

26 Mao, X., Tao, C., Olyarchuk, J. G. & Wei, L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* **21**, 3787-3793 (2005).

27 Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic acids res.* **45**, D362-D368, doi:10.1093/nar/gkw937 (2017).

28 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).

29 Ishizaki, T., Sakai, A., Koizumi, T., Ruan, Z. & Zhan, G. W. Blunted effect of the Kv channel inhibitor on pulmonary circulation in Tibetan sheep: A model for studying hypoxia and pulmonary artery pressure regulation. *Respirology* **9**, 125-129 (2004).

30 Rong, C. *et al.* Cardiac adaptive mechanisms of Tibetan antelope (Pantholops hodgsonii) at high altitudes. *Am. J. Vet. Res.* **73**, 809-813 (2012).

31 Liu, F. Y. & Chen, Q. H. Effect of acute hypoxia on the parameters of blood gas in Tibetan plateau sheep. *Chin. J. Zool.* **41**, 48-52 (2006).

32 Ruan, Z., Koizumi, T., Sakai, A., Ishizaki, T. & Wang, Z. Endogenous nitric oxide and pulmonary circulation response to hypoxia in high-altitude adapted Tibetan sheep. *Eur. J. Appl. Physiol.* **93**, 190-195 (2004).

33 Wang, X. R. & Wu, J. P. Study on structural features of the cerebral arterial system associated with plateau adaptability in tibetan sheet. *Acta Ecologiae Animalis Domastici* **34** (2013).

685 34 Yuan, L. G., Sun, Y. & Huang, B. M. Micromorphological characteristics of the adult tibetan plateau sheep
686 spermatic cord. *Acta Veterinaria et Zootechnica Sinica* **42**, 1450-1456 (2011).
687 35 Ge, R. L. *et al.* Higher exercise performance and lower VO2max in Tibetan than Han residents at 4,700 m altitude.
688 *J. Appl. Physiol.* **77**, 684-691 (1994).
689 36 Han, S. X, Xiao, G. & Yang, S. L. Physiologic and Molecular Mchanism of the Animal Hypoxia Adaptation.
690 *Chinese animal husbandry and veterinary medicine* **37**, 29-34 (2010).
691 37 Kopp, F. & Mendell, J. T. Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell*
692 **172**, 393-407, doi:10.1016/j.cell.2018.01.011 (2018).
693 38 Chen, X. & Yan, G. Y. Novel human lncRNA-disease association inference based on lncRNA expression profiles.
694 *Bioinformatics* **29**, 2617-2624, doi:10.1093/bioinformatics/btt426 (2013).
695 39 Cao, M., Zhao, J. & Hu, G. Genome-wide methods for investigating long noncoding RNAs. *Biomed.*
696 *Pharmacother.* **111**, 395-401 (2019).
697 40 Bakhtiarizadeh, M. R. & Salami, S. A. Identification and Expression Analysis of Long Noncoding RNAs in
698 Fat-Tail of Sheep Breeds. *G3 Genesgenetics* **9**, 1263-1276 (2019).
699 41 Ren, C. *et al.* Genome-Wide Analysis Reveals Extensive Changes in LncRNAs during Skeletal Muscle
700 Development in Hu Sheep. *Genes (Basel)* **8**, doi:10.3390/genes8080191 (2017).
701 42 Feng, X. *et al.* Genome-wide differential expression profiling of mRNAs and lncRNAs associated with
702 prolificacy in Hu sheep. *Biosci. Rep.* **38**, BSR20171350 (2018).
703 43 Simonson, T. S. *et al.* Genetic evidence for high-altitude adaptation in Tibet. *Science* **329**, 72-75,
704 doi:10.1126/science.1189406 (2010).
705 44 Sakamoto, T. *et al.* Hypoxia-Inducible Factor 1 Regulation through Cross Talk between mTOR and MT1-MMP.
706 *Mol. Cell Biol.* **34**, 30-42 (2014).
707 45 Capitanio, D. *et al.* TCA cycle rewiring fosters metabolic adaptation to oxygen restriction in skeletal muscle from
708 rodents and humans. *Sci. Rep.* **7**, 9723-9723 (2017).
709 46 Kharrati-Koopaee, H., Ebrahimie, E., Dadpasand, M., Niazi, A. & Esmailizadeh, A. Genomic analysis reveals
710 variant association with high altitude adaptation in native chickens. *Sci. Rep.* **9**, 9224 (2019).
711 47 Qi, X., Zhang, Q., He, Y., Yang, L. & Zhang, X. The Transcriptomic Landscape of Yaks Reveals Molecular
712 Pathways for High Altitude Adaptation. *Genome Bio. Evol.* **11**, 72-85 (2019).
713 48 Guo, J. *et al.* Comparative genome analyses reveal the unique genetic composition and selection signals
714 underlying the phenotypic characteristics of three Chinese domestic goat breeds. *Genet. Sel. Evol.* **51**, 1-18
715 (2019).
716 49 Kumar, C. *et al.* Sequence Characterization of DSG3 Gene to Know Its Role in High-Altitude Hypoxia
717 Adaptation in the Chinese Cashmere Goat. *Frontiers in Genetics* **9**, 553 (2018).@@@
718 50 Wang, J. *et al.* Genetic diversity of ATP8 and ATP6 genes is associated with high-altitude adaptation in yak.
719 *Mitochondrial Dna Part A* **29**, 385-393 (2018).
720 51 Rayner, B. S., Hua, S., Sabaretnam, T. & Witting, P. K. Nitric oxide stimulates myoglobin gene and protein
721 expression in vascular smooth muscle. *Biochemical Journal* **423**, 169 (2009).
722 52 Mccurdy, C. E. *et al.* Attenuated Pik3r1 Expression Prevents Insulin Resistance and Adipose Tissue Macrophage
723 Accumulation in Diet-Induced Obese Mice. *Diabetes* **61**, 2495-2505 (2012).
724 53 Vorrink, S. U., Severson, P. L., Kulak, M. V., Futscher, B. W. & Domann, F. E. Hypoxia perturbs aryl
725 hydrocarbon receptor signaling and CYP1A1 expression induced by PCB 126 in human skin and liver-derived
726 cell lines. *Toxicology & Applied Pharmacology* **274**, 408-416 (2014).
727 54 Moore, A. D. A Study of Effects on MMP14 Transcriptional Regulation and Angiogenesis by Hypoxia and Statins.
728 (2014).

55 Chen, H. J., Li, D. Y., Saldeen, T. & Mehta, J. L. TGF-b1 modulates NOS expression and phosphorylation of Akt/PKB in rat myocytes exposed to hypoxia-reoxygenation. *Ajp Heart & Circulatory Physiology* **281**, H1035-1039 (2001).

56 Yang, J. *et al.* Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Molecular Biology & Evolution* **33**, 2576–2592 (2016).

57 Wei, C. *et al.* Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. *Scientific Reports* **6**, 26770 (2016).

58 Frede, S. & Fandrey, J. Cellular and Molecular Defenses Against Hypoxia. In: Swenson E., Bärtsch P. (eds) High Altitude. Springer, New York, NY. (2014)

59 Feng, C. C. *et al.* Hypoxia suppresses myocardial survival pathway through HIF-1α-IGFBP-3-dependent signaling and enhances cardiomyocyte autophagic and apoptotic effects mainly via FoxO3a-induced BNIP3 expression. *Growth Factors Journal* **34**, 73-86 (2016).

60 Murakami, A. *et al.* Hypoxia Increases Gefitinib-Resistant Lung Cancer Stem Cells through the Activation of Insulin-Like Growth Factor 1 Receptor. *Plos One* **9**, e86459 (2014).

61 Mora & A. Deficiency of PDK1 in cardiac muscle results in heart failure and increased sensitivity to hypoxia. *Embo Journal* **22**, 4666-4676 (2014).

62 Jiang, X. *et al.* Role of activating transcription factor 3 (ATF3) in sublytic C5b-9-induced glomerular mesangial cell apoptosis. *Cell Mol Immunol* **7**, 143-151 (2010).

63 Parra, V., Bravo-Sagua, R., Norambuena-Soto, I., Hernández-Fuentes, C. P. & Castro, P. F. Inhibition of mitochondrial fission prevents hypoxia-induced metabolic shift and cellular proliferation of pulmonary arterial smooth muscle cells. *Biochim Biophys Acta* **1863**, 2891-2903 (2017).

64 Fradette, C., Batonga, J., Teng, S., Piquettemiller, M. & Souich, P. D. Animal models of acute moderate hypoxia are associated with a down-regulation of CYP1A1, 1A2, 2B4, 2C5, and 2C16 and up-regulation of CYP3A6 and P-glycoprotein in liver. *Drug Metabolism & Disposition the Biological Fate of Chemicals* **35**, 765 (2007).

Figure and Table Legends

Fig. 1: The physiological and biochemical indices of five sheep populations from different altitude. (A) Partial pressure of carbon dioxide (PO_2). (B) Haemoglobin contents (HGB). (C) Total protein (TP). (D) The number of alveolar counted per unit area (MAN). The columns with different letters are significantly different at $P < 0.05$.

Fig. 2: Identification of long non-coding RNAs (lncRNAs) and mRNAs involved in high altitude hypoxia adaptation in lung tissue. Coding potentiality filter using Coding Potential Calculator, Pfam, Phylogenetic codon substitution frequency, and Coding-Non-Coding Index (A). The boxplot shows the expression levels (log10 (FPKM +1)) of lncRNAs and mRNAs (B). Transcript lengths distribution of mRNAs (C) and lncRNAs (D). Exon number distribution of mRNAs (E) and lncRNAs (F). Open reading frame (ORF) length distribution of mRNAs (G) and lncRNAs (H).

Fig. 3: The up and down regulation of DE genes of five sheep populations. (A) Number of DE lncRNAs in liver samples. (B) Number of DE lncRNAs in lung samples. (C) Number of

DE mRNAs in liver samples. (D) Number of DE mRNAs in lung samples. Red: up-regulated; Green: down-regulated.

Fig. 4: Venn diagram of common differential expression genes. (A) The common DE lncRNAs in pairwise comparison between lung and liver tissue. (B) The common DE mRNAs in pairwise comparison between lung and liver tissue. (C) The common DE lncRNAs in four comparison groups of liver samples (AW vs HS, HB vs HS, GJ vs HS, WT vs HS). (D) The common DE mRNAs in four comparison groups of liver samples (AW vs HS, HB vs HS, GJ vs HS, WT vs HS). (E) The common DE lncRNAs in four comparison groups of lung samples (AW vs HS, HB vs HS, GJ vs HS, WT vs HS). (F) The common DE mRNAs in four comparison groups of lung samples (AW vs HS, HB vs HS, GJ vs HS, WT vs HS).

Fig. 5: Confirmation of expression patterns of the eight selected differential expression genes using qRT-PCR. The trends of qRT-PCR results are consistent with RNA-seq data. (A) The common DE lncRNA TCONS_00139593 in five sheep populations of liver samples (AW, HB, GJ, WT, HS). (B) The common DE lncRNA TCONS_00332125 in five sheep populations of liver samples (AW, HB, GJ, WT, HS). (C) The common DE lncRNA TCONS_00377466 in five sheep populations of lung samples (AW, HB, GJ, WT, HS). (D) The common DE lncRNA TCONS_101107056 in five sheep populations of liver samples (AW, HB, GJ, WT, HS). (E) The common DE mRNAs TFF3 in five sheep populations of liver samples (AW, HB, GJ, WT, HS). (F) The common DE mRNAs TFF3 in five sheep populations of lung samples (AW, HB, GJ, WT, HS). (G) The common DE mRNAs ATF3 in five sheep populations of lung samples (AW, HB, GJ, WT, HS). (H) The common DE mRNAs PTGS2 in five sheep populations of lung samples (AW, HB, GJ, WT, HS).

Table 1: Sampling information of five sheep populations (*Ovis aries*) for transcriptome analysis.

Table 2: The candidate genes putatively associated with the plateau adaptation of Tibetan sheep.

Figures

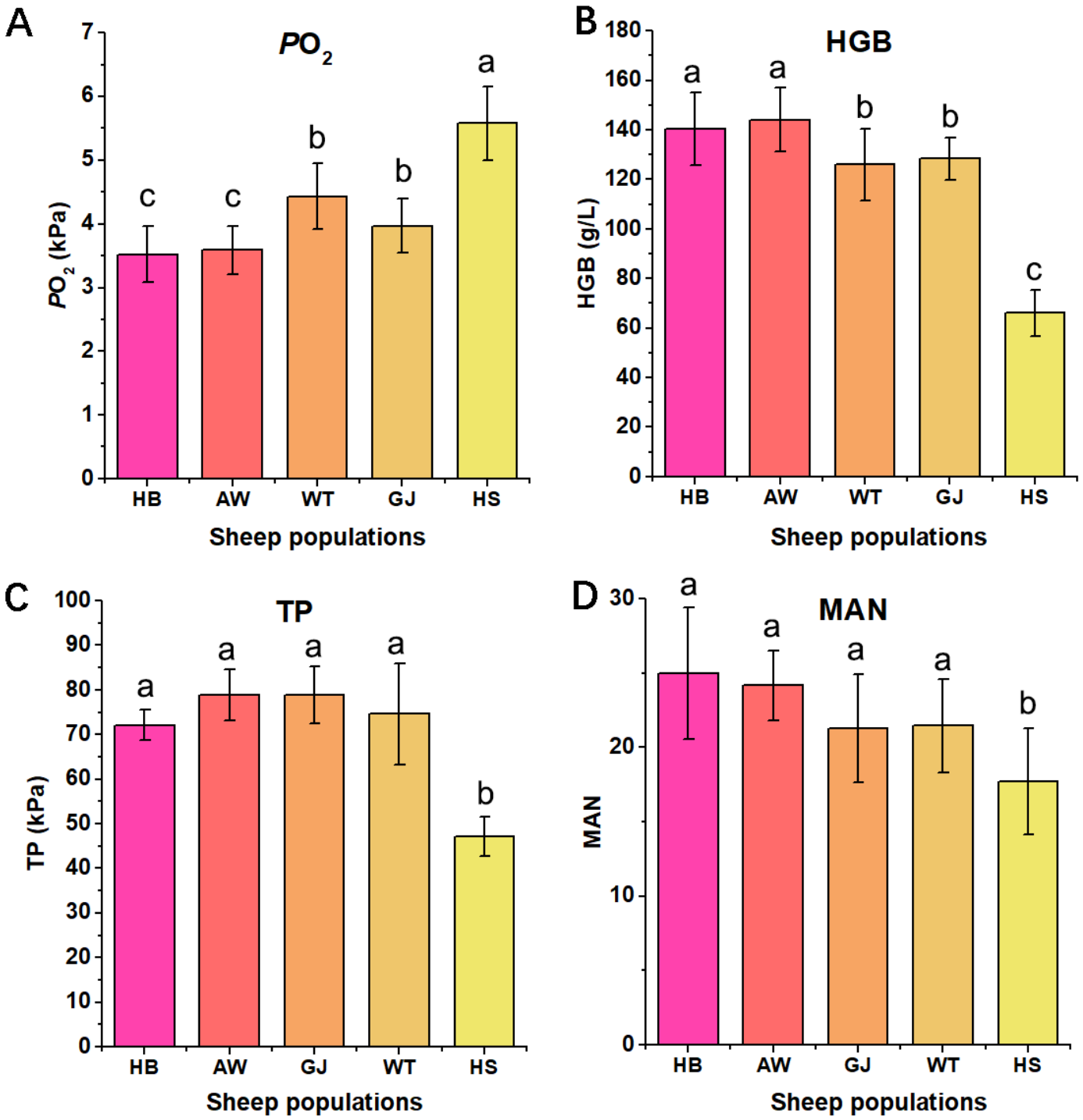

Figure 1

The physiological and biochemical indices of five sheep populations from different altitude. (A) Partial pressure of carbon dioxide (PO2). (B) Haemoglobin contents (HGB). (C) Total protein (TP). (D) The number of alveolar counted per unit area (MAN). The columns with different letters are significantly different at $P < 0.05$.

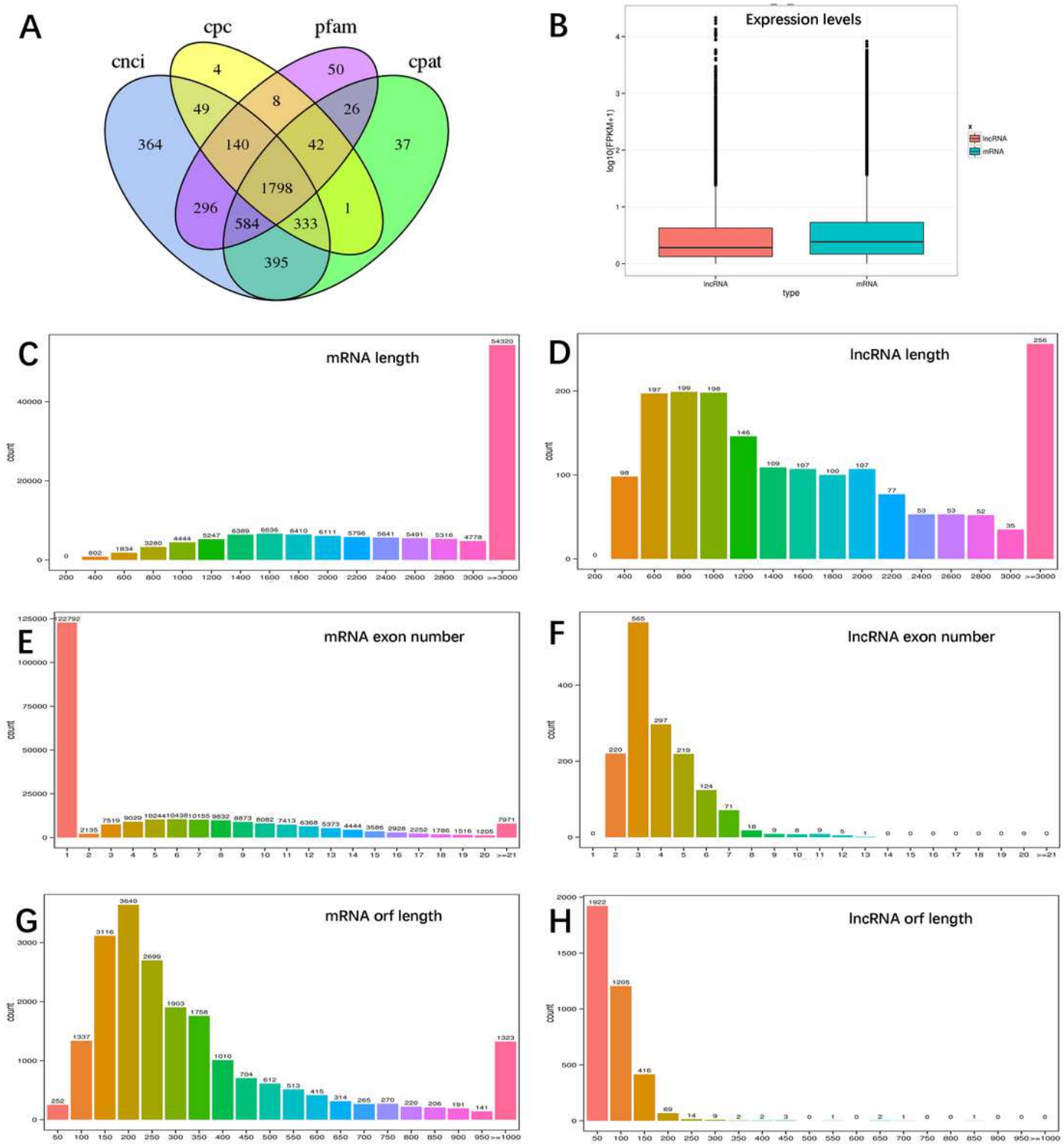

Figure 2

Identification of long non-coding RNAs (lncRNAs) and mRNAs involved in high altitude hypoxia adaptation in lung tissue. Coding potentiality filter using Coding Potential Calculator, Pfam, Phylogenetic codon substitution frequency, and Coding-Non-Coding Index (A). The boxplot shows the expression levels (log10 (FPKM +1)) of lncRNAs and mRNAs (B). Transcript lengths distribution of mRNAs (C) and lncRNAs

(D). Exon number distribution of mRNAs (E) and lncRNAs (F). Open reading frame (ORF) length distribution of mRNAs (G) and lncRNAs (H).

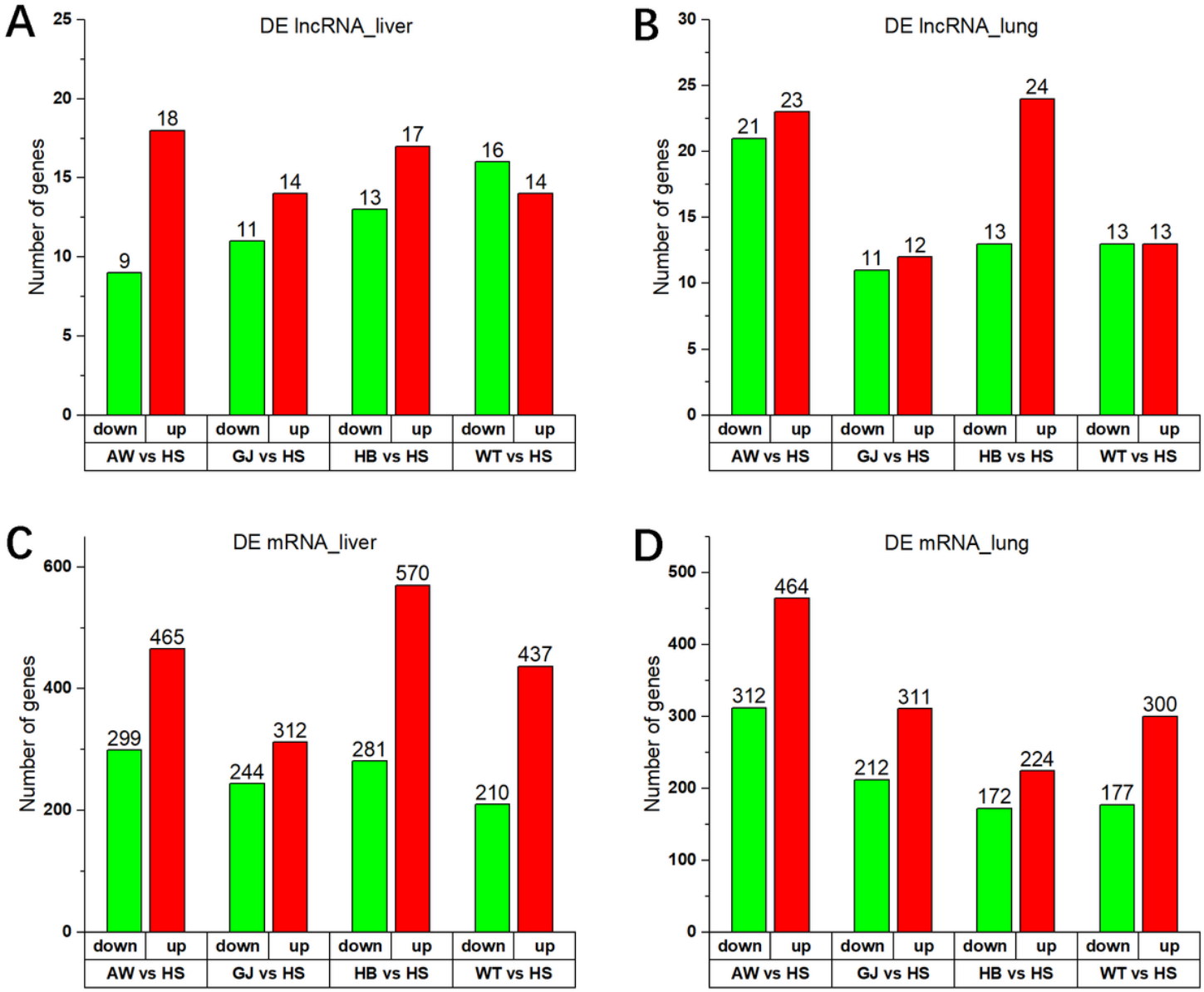

Figure 3

The up and down regulation of DE genes of five sheep populations. (A) Number of DE lncRNAs in liver samples. (B) Number of DE lncRNAs in lung samples. (C) Number of DE mRNAs in liver samples. (D) Number of DE mRNAs in lung samples. Red: up-regulated; Green: down-regulated.

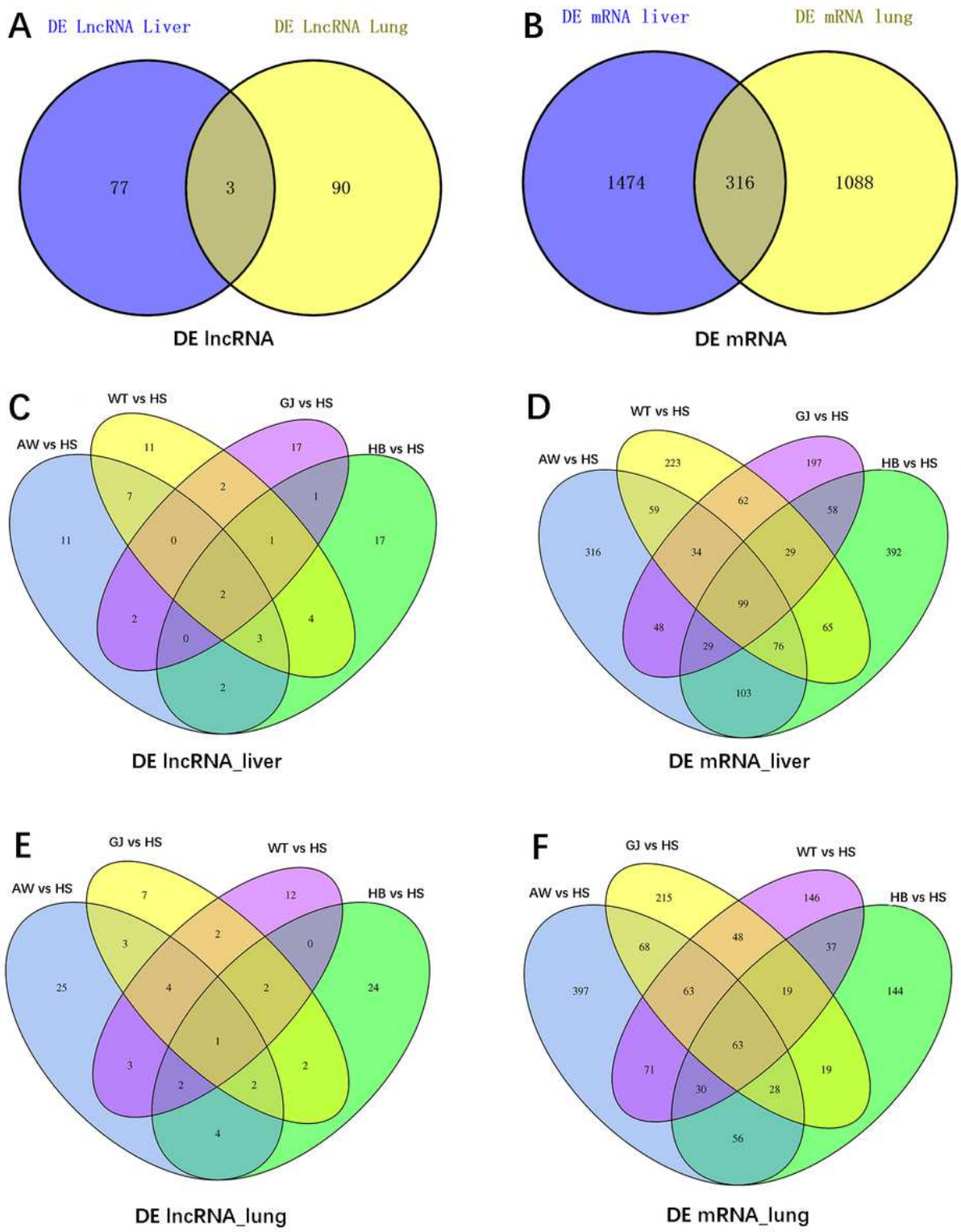

Figure 4

Venn diagram of common differential expression genes. (A) The common DE lncRNAs in pairwise comparison between lung and liver tissue. (B) The common DE mRNAs in pairwise comparison between lung and liver tissue. (C) The common DE lncRNAs in four comparison groups of liver samples (AW vs HS, HB vs HS, GJ vs HS, WT vs HS). (D) The common DE mRNAs in four comparison groups of liver samples (AW vs HS, HB vs HS, GJ vs HS, WT vs HS). (E) The common DE lncRNAs in four comparison

groups of lung samples (AW vs HS, HB vs HS, GJ vs HS, WT vs HS). (F) The common DE mRNAs in four comparison groups of lung samples (AW vs HS, HB vs HS, GJ vs HS, WT vs HS).

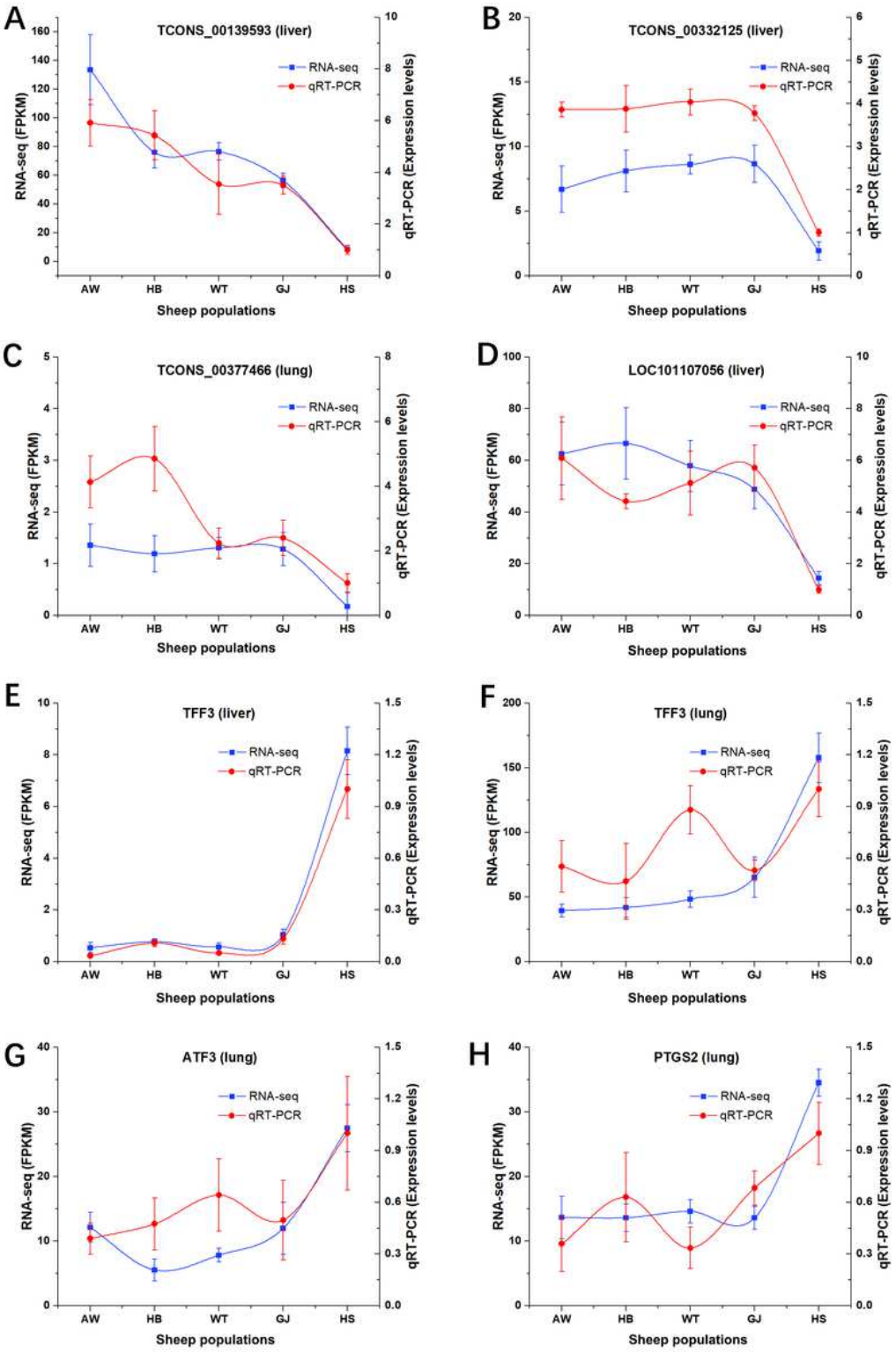

Figure 5

Confirmation of expression patterns of the eight selected differential expression genes using qRT-PCR. The trends of qRT-PCR results are consistent with RNA-seq data. (A) The common DE lncRNA TCONS_00139593 in five sheep populations of liver samples (AW, HB, GJ, WT, HS). (B) The common DE

lncRNA TCONS_00332125 in five sheep populations of liver samples (AW, HB, GJ, WT, HS). (C) The common DE lncRNA TCONS_00377466 in five sheep populations of lung samples (AW, HB, GJ, WT, HS). (D) The common DE lncRNA TCONS_101107056 in five sheep populations of liver samples (AW, HB, GJ, WT, HS). (E) The common DE mRNAs TFF3 in five sheep populations of liver samples (AW, HB, GJ, WT, HS). (F) The common DE mRNAs TFF3 in five sheep populations of lung samples (AW, HB, GJ, WT, HS). (G) The common DE mRNAs ATF3 in five sheep populations of lung samples (AW, HB, GJ, WT, HS). (H) The common DE mRNAs PTGS2 in five sheep populations of lung samples (AW, HB, GJ, WT, HS).

Supplementary Files

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

- Table1.pdf
- Table2.pdf
- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- TableS5.xlsx
- TableS6.xlsx
- TableS7.xlsx
- TableS8.xlsx
- TableS9.xlsx
- TableS10.xlsx
- FigureS1.tif
- FigureS2.tif
- FigureS3.tif
- FigureS4.tif