



OPEN

## Molecular mechanisms detected in yak lung tissue via transcriptome-wide analysis provide insights into adaptation to high altitudes

Qianyun Ge<sup>1</sup>, Yongbo Guo<sup>2</sup>, Wangshan Zheng<sup>2</sup>, Shengguo Zhao<sup>1</sup>✉, Yuan Cai<sup>1</sup>✉ & Xuebin Qi<sup>2</sup>✉

Due to their long-term colonization of and widespread distribution in plateau environments, yaks can serve as an ideal natural animal model for the adaptive evolution of other plateau species, including humans. Some studies reported that the lung and heart are two key organs that show adaptive transcriptional changes in response to high altitudes, and most of the genes that show differential expression in lung tissue across different altitudes display nonlinear regulation. To explore the molecular mechanisms that are activated in yak lung tissue in response to hypoxia, the mRNAs, lncRNAs and miRNAs of lung tissue from 9 yaks living at three different altitudes (3400 m, 4200 m and 5000 m), with three repetitions per altitude, were sequenced. Two Zaosheng cattle from 1500 m were selected as low-altitude control. A total of 21,764 mRNAs, 14,168 lncRNAs and 1209 miRNAs (305 known and 904 novel miRNAs) were identified. In a comparison of yaks and cattle, 4975 mRNAs, 3326 lncRNAs and 75 miRNAs were differentially expressed. A total of 756 mRNAs, 346 lncRNAs and 83 miRNAs were found to be differentially expressed among yaks living at three different altitudes (fold change  $\geq 2$  and P-value  $< 0.05$ ). The differentially expressed genes between yaks and cattle were functionally enriched in long-chain fatty acid metabolic process and protein processing, while the differentially expressed genes among yaks living at three different altitudes were enriched in immune response and the cell cycle. Furthermore, competing endogenous RNA (ceRNA) networks were investigated to illustrate the roles of ceRNAs in this process, the result was also support the GO and KEGG analysis. The present research provides important genomic insights for discovering the mechanisms that are activated in response to hypoxia in yak lung tissue.

The organic mechanisms that mediate adaptation to high-altitude environments have attracted widespread attention in recent years. The genomes of some mammals that inhabit highlands, including human highlanders, have been sequenced, and many genes are associated with adaptation to high altitudes<sup>1–9</sup>. Focusing on the mechanisms underlying transcriptomic changes can provide insights into the adaptive evolution of other plateau species, including humans. Due to their long-term colonization of and widespread distribution in the plateau, Tibetans and yaks are two ideal models for studying the ability to adapt to plateau environments<sup>5,10,11</sup>. Compared with those of lowland cattle, yak lungs have developed physiological characteristics that are adapted to high-altitude hypoxia, including a larger pulmonary alveolar area per unit area, a thinner alveolar septum, a thinner blood–air barrier and smooth muscles within the arteriole wall of the microartery with a diameter of  $< 50$   $\mu\text{m}$ , whereas lowland cattle do not have such a structure. These physiological features can promote more efficient blood flow for the transport of oxygen under hypobaric hypoxia. The lung is a central functional organ in the respiratory system, and it plays a substantial role in adaptation to hypoxia in plateau environments. In addition to allowing an animal's body to adapt to external environmental stimuli through a series of physiological changes, gene expression, as an intermediate phenotype linking DNA sequences and physiological traits, plays an important

<sup>1</sup>College of Animal Science and Technology, Gansu Agricultural University, Lanzhou 730070, China. <sup>2</sup>State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. ✉email: zhaosg@gsau.edu.cn; caiyuan@gsau.edu.cn; qixuebin@mail.kiz.ac.cn

Contrasts	mRNA		miRNA		lncRNA	
	Up	Down	Up	Down	Up	Down
CON VS T1	1011	1019	19	12	740	713
CON VS T2	779	703	5	5	489	454
CON VS T3	691	772	16	18	454	476
T1 VS T2	117	158	11	11	46	84
T1 VS T3	120	160	6	11	48	95
T2 VS T3	119	82	23	21	48	25

**Table 1.** Differentially expressed mRNAs, lncRNAs and miRNAs. ZSlg01 and ZSlg02 indicate libraries derived from the lung tissue of Zaosheng cattle in two biological replicates belonging to the CON group. MQlg01, MQlg02, and MQlg03 indicate libraries derived from the lung tissue of yaks living at an altitude of 3,400 m in three biological replicates belonging to the T1 group. LZlg01, BGl02, and DXlg03 indicate libraries derived from the lung tissue of yaks living at an altitude of 4,200 m in three biological replicates belonging to the T2 group. ADlg01, ADlg02, and ADlg03 indicate libraries derived from the lung tissue of yaks living at an altitude of 5,000 m in three biological replicates belonging to the T3 group.

role in revealing molecular pathways/networks associated with genetic adaptation<sup>12,13</sup>. There is growing evidence that changes in gene expression are also essential for adaptation to high altitudes<sup>14–16</sup>.

Some studies investigated the differences in gene expression in the heart between yaks and cattle and reported that some genes associated with the oxygen supply system and the defense systems that respond to hypoxic conditions are differentially expressed<sup>17</sup>. Further study also revealed that the lung and heart are two key organs that show adaptive transcriptional changes in response to high altitudes, and most of the genes that are differentially expressed in the lung tissue across different altitudes display nonlinear regulation<sup>18</sup>. To identify organic mechanisms that mediate adaptation of the lung tissue to hypoxia in yaks, we conducted transcriptome analysis of lung tissue responses to hypoxia in yaks and cattle. To further study the changes in gene regulation that occur in response to hypoxia at increasing altitudes, we collected lung tissue samples from 9 yaks with long-term residence at three different altitudes (3400 m, 4200 m, and 5000 m), three repetitions per altitude. Since yaks are a native plateau species, they are only distributed between 3000 and 5000 m above sea level, so there is no lowland control. In phylogeny, the taurine cattle (*Bos taurus*) is the lowland bovine species that is most closely related to yaks and can be used as a lowland control<sup>19</sup>. We chose two Zaosheng cattle, China's native species, as a lowland control that lives at 1500 m and is geographically adjacent to the Tibetan Plateau. We performed RNA sequencing (RNA-Seq) and mapped the detailed transcriptome of lung tissue. The findings of this study improve our understanding of the molecular mechanisms that are activated in response to hypoxia and are useful for developing strategies to improve the adaptation of other plateau species.

## Results

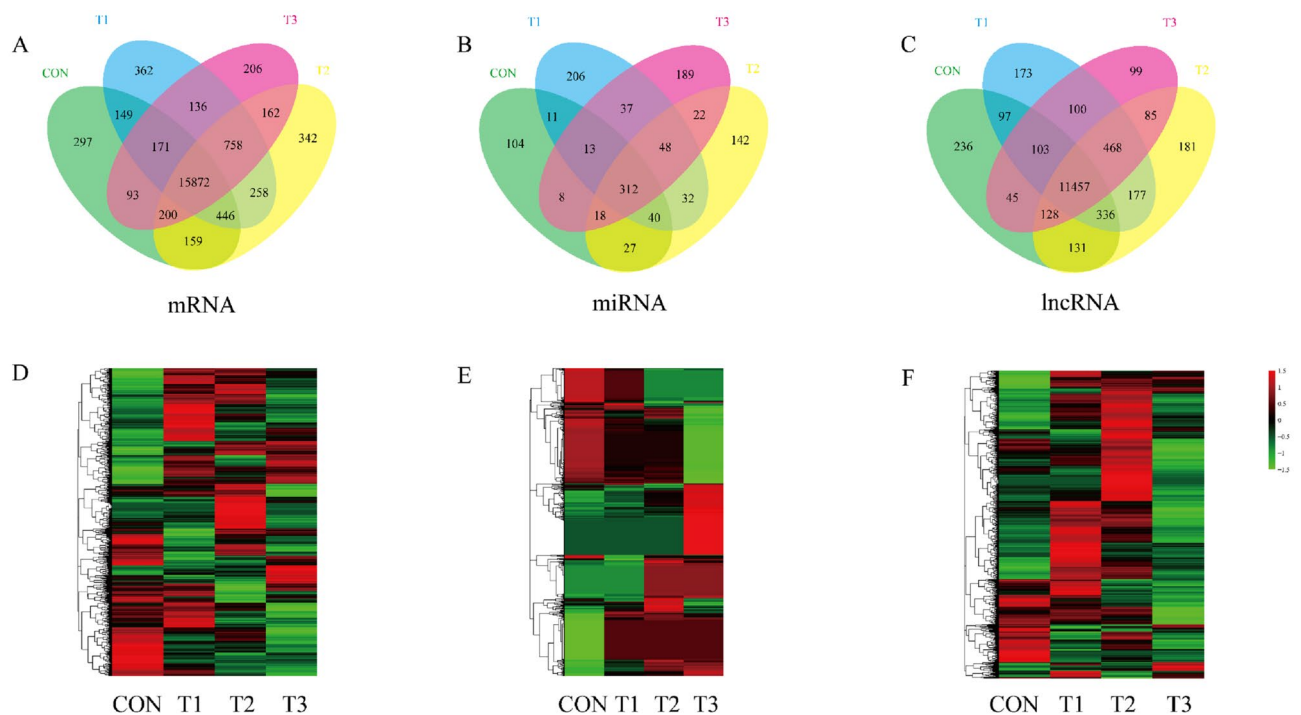
**Sequencing and mapping of mRNAs and lncRNAs.** Transcriptome data sets were generated by RNA-seq to develop a comprehensive catalogue of mRNAs and lncRNAs from the bovine lung. Illumina sequencing of bovine lung tissues yielded a total of 932.336 M raw reads. The Phred Quality Score of the samples was more than 91.7% for Q30. The clean data included 83.501 M reads on average. Subsequently, the clean reads were mapped onto the bovine reference genome using HiSAT2<sup>20</sup>. The average mapping rate of the eleven samples was 86.28%, and the unmapped rates were between 8.51% and 37.01%. The clustering of samples based on the variations in the gene expression was shown in Figure S1. A detailed summary of the sequencing results is shown in Table S1; the results indicated that the transcriptome sequencing data were of high quality and had a suitable mapping rate (Figure S2).

**Identification of miRNAs.** Eleven small RNA libraries were constructed and sequenced independently to identify miRNAs in yak and cattle lung tissue. A total of 109.949 M raw reads were generated, and 100.664 M clean reads were obtained, which accounted for more than 92% of the raw data. Among eleven individual libraries, the minimum rates of removed reads were at a suitable level, and the Q30 values of the raw or clean data were high enough to indicate high-quality small RNA sequencing (Table S2). The length of most sequences was distributed in the range of 21.7–22.2 nt, and the highest percentage was observed for a length of 22 nt, which is consistent with the typical size of miRNAs (Figure S2). In total, 305 known and 904 novel miRNAs were identified using miRBase (Release 21) and miRDeep2<sup>21</sup> (Table S3).

**Identification of differentially expressed mRNAs, lncRNAs and miRNAs.** The identification of differentially expressed mRNAs, lncRNAs and miRNAs was performed using edgeR with a threshold of fold change  $\geq 2$  and P-value  $< 0.05$ . A higher number of differentially expressed mRNAs, miRNAs and lncRNAs was obtained from the CON/T1, CON/T2 and CON/T3 pairwise comparisons than from the T1/T2, T1/T3 and T2/T3 pairwise comparisons (Tables 1, 2). A total of 4975 mRNAs, 3326 lncRNAs and 75 miRNAs were found to be differentially expressed between the yak and cattle groups. A total of 756 mRNAs, 346 lncRNAs and 83 miRNAs were found to be differentially expressed among the yaks living at three different altitudes. The details and distribution about differentially expressed mRNAs, miRNAs and lncRNAs in each pairwise comparison were

Sample	Category	Sample size	Sampling site	Altitude (m)	Duplicates
CON	Zaosheng cattle	2	Ningxian, Gansu, China	1500	ZSlg01 ZSlg02
T1	Yak	3	Maqu, Gannan Tibetan Autonomous Prefecture, Gansu, China	3400	MQlg01 MQlg02 MQlg03
T2	Yak	3	Dangxiong, Linzhou and Bange, Tibet Autonomous Region, China	4200	LZlg01, BGlg02 DXlg03
T3	Yak	3	Anduo, Tibetan Autonomous Region, China	5000	ADlg01 ADlg02 ADlg03

**Table 2.** Characteristics of samples.



**Figure 1.** Venn analysis of mRNAs (A), miRNAs (B) and lncRNAs (C) detected among four groups. Cluster analysis of expression of mRNAs (D) miRNAs (E) and lncRNAs (F). The data are depicted as matrices in which each row represents one mRNA, miRNA, or lncRNA and each column represents one of the samples. Relative mRNA, miRNA, or lncRNA expression is depicted according to the color scale shown at the top. Red and green represent high and low relative expression, respectively. The magnitude of deviation from the median is represented by color saturation.

shown in Table 1. Cluster analysis of differentially expressed mRNAs, miRNAs and lncRNAs was conducted with a heatmap (Fig. 1). Then, these differentially expressed noncoding RNAs (ncRNAs) were intersected with predicted target genes. As a result, 13,283 and 3673 genes targeted by miRNAs and lncRNAs, respectively, were screened out (Tables S4 and S5).

**Altitude-related transcriptomic profiles.** We collected nine lung samples in total with three from each of the three elevations. In total, 275 genes showed significant differential expression between the yak lung samples from the T1 and T2 groups, and 280 and 201 differentially expressed genes (DEGs) were identified from the comparisons of the T1 and T3 groups and the T2 and T3 groups, respectively. We performed a STEM analysis to group the DEGs of yaks at three different altitudes on the basis of expression pattern. With an FC threshold of 2, the number of DEGs was too low to identify the enriched terms, and we thus adjusted the FC threshold to 1.5. Using this criteria, 3060 genes were partitioned into four significantly ( $P < 0.05$ ) enriched altitude profiles (Fig-

ure S3). The dominance of the cross-altitude differentially expressed genes in lung display a nonlinear regulation, implying a generally nonlinear correlation between expression level and elevation in lung.

**Functional enrichment analyses of differentially expressed ncRNA target genes and differentially expressed mRNAs.** To investigate the functions of differentially expressed lncRNAs and miRNAs, enrichment analyses of their target genes were performed. In a comparison of yaks and cattle, for target genes of the differentially expressed miRNAs (Table S6), metabolic pathways were significantly enriched, including glycerophospholipid metabolism (KO00564), oxidative phosphorylation (KO00190) and glycosaminoglycan biosynthesis-heparan sulfate/heparin (KO00534). For target genes of differentially expressed lncRNAs (Table S7), metabolic pathways were also significantly enriched, including metabolism of xenobiotics by cytochrome P450 (KO00980), drug metabolism-cytochrome P450 (KO00982), drug metabolism-other enzymes (KO00983), retinol metabolism (KO00830), ascorbate and aldarate metabolism (KO00053), pentose and glucuronate interconversion (KO00040) and steroid hormone biosynthesis (KO00140). However, the immune system was significantly enriched for the target genes of the differentially expressed miRNAs among the yaks living at three different altitudes (Table S8), including leukocyte transendothelial migration (KO04670), hematopoietic cell lineage (KO04640) and T cell receptor signaling pathway (KO04660). Environmental information processing, such as the FoxO signaling pathway (KO04068), cell adhesion molecules (CAMs) (KO04514) and the sphingolipid signaling pathway (KO04071), was also highly represented. Cellular processes, such as apoptosis (KO04210) and the p53 signaling pathway (KO04115), were also highly represented. For target genes of differentially expressed lncRNAs (Table S9), organismal systems were significantly enriched, including pancreatic secretion (KO04972), renin secretion (KO04924), Toll-like receptor signaling pathway (KO04620), complement and coagulation cascades (KO04610), circadian rhythm (KO04710) and cardiac muscle contraction (KO04260). Cellular processes were also significantly enriched, including gap junctions (KO04540), adherens junctions (KO04520) and tight junctions (KO04530). These results suggested that abnormal ncRNA expression may affect metabolism, the immune system, cellular processes and environmental information processing in the bovine lung by regulating target genes.

As the expression of mRNAs is directly related to biological characteristics, the function of differentially expressed mRNAs was analyzed by KEGG and GO enrichment. When we compared yaks with cattle, differentially expressed mRNAs were classified into three major functional categories ('biological processes', 'molecular function', and 'cellular components') and were enriched in 563 GO categories. The top 10 included seven in "biological process", two in "molecular function" and one in "cellular components", of which 'long-chain fatty acid metabolic process' and 'protein targeting to peroxisome' terms were most significant (Fig. 2, Table S10). Differentially expressed mRNAs were enriched in 62 GO categories among yaks living at three different altitudes. The top 10 included 8 in "biological process", one in "molecular function" and one in "cellular components", of which the 'immune response' term was most significant (Fig. 2, Table S11). We also performed KEGG pathway enrichment analyses to evaluate the biological significance of the DEGs. A hypergeometric test with a P-adjusted value cutoff of 0.05 was used as the criterion for pathway detection (Fig. 2). To confirm the gene expression patterns, some mRNAs and lncRNAs were randomly selected to be validated by q-PCR. The results were in concordance with the RNA-seq data (Figure S4). RT-qPCR was performed with the internal control  $\beta$ -actin using the  $2^{-\Delta\Delta Ct}$  method.

**Differential RNA expression identify core regulatory networks & RNAs.** To identify the key RNAs associated with hypoxic adaptation in yak lung, we analyzed the differentially expressed RNAs in the form of Venn diagrams (Fig. 3). The intersection of the differentially expressed RNAs considered the most interesting candidates, as they represented the main differences in RNA expression. We construct the networks to further interrogate these differentially expressed RNAs which in the intersection of Venn diagrams.

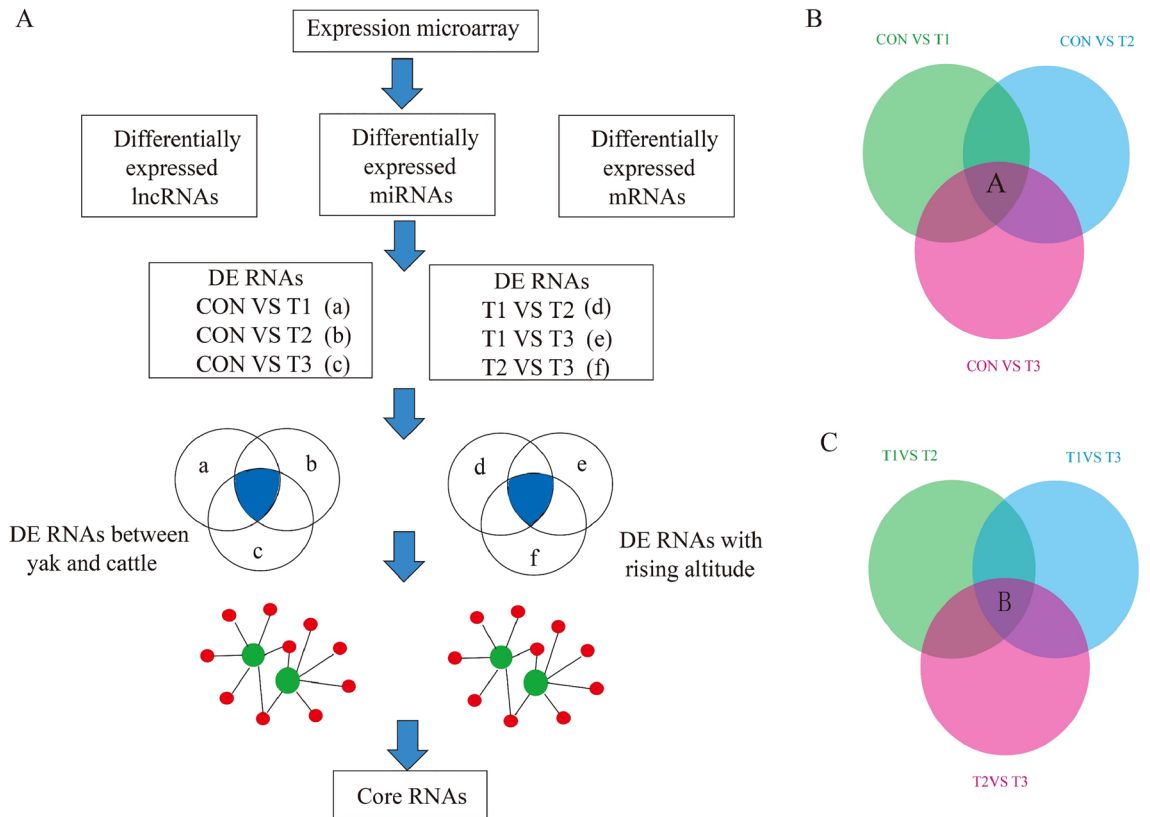
The intersection of the differentially expressed RNAs among the CON vs T1, CON vs T2 and CON vs T3 comparisons (Fig. 3B part A) were represented the main differences in RNA expression between yak and cattle. In the networks, all mRNAs and circRNAs were downregulated, and the miRNAs were upregulated. These RNAs, including ptc-miR6471\_R18-4L21, gma-miR6300\_R2-18L18, sly-miR482e-3p\_R1-21L22, ppt-miR902f-3p\_R15-1L20, stu-miR408b-5p\_R4-21L21, ath-miR5658\_R1-19L21\_8T-A, ata-miR166d-3p\_R1-20L21 and ata-miR5168-3p, were selected as the most affected RNAs between yak and cattle, and their associated RNAs, including circRNAs and mRNAs (SKIV2L2, PRKCSH, NewGene.10854.1, POR and LOC102286089), identified using miRanda were selected as the core RNAs (Fig. 4A).

The intersection of differentially expressed RNAs identified from the T1 vs T2, T1 vs T3 and T2 vs T3 comparisons represented the differences in RNA expression in yak lung with increasing altitude (Fig. 3C part B). We also constructed networks of the differentially expressed RNAs in the intersection, which included nine downregulated miRNAs (NewGene.14316.2, pab-miR396a, ppt-miR902f-3p\_R15-1L20, stu-miR408b-5p\_R4-21L21, bra-miR319-3p\_R1-20L21, ath-miR5658\_R1-19L21\_8T-A, NewGene.9265.1, NewGene.217.1 and gma-miR6300\_R2-18L18) and four upregulated miRNAs (NewGene.16064.3, NewGene.18257.2, NewGene.17679.7 and NewGene.18455.7), and their related three upregulated mRNAs (PARP4, SLTM and CLCNT) and three downregulated mRNAs (UVSSA, MFN2 and CASK). There was no related lncRNAs (Fig. 4B).

## Discussion and conclusion

Since the development of next-generation sequencing, a large number of whole-genome sequencing studies of Tibetan mammals have been performed to explore the molecular mechanism of high-altitude adaptation<sup>5,22</sup>. The yak is an ancient species that is unique to the Tibetan Plateau in China. During long-term adaptation to harsh natural environments, such as the hypoxia and low temperatures of the Tibetan Plateau, many genes specifically increase related transcriptional activities and regulate a series of metabolic activities in the body. The lung is a

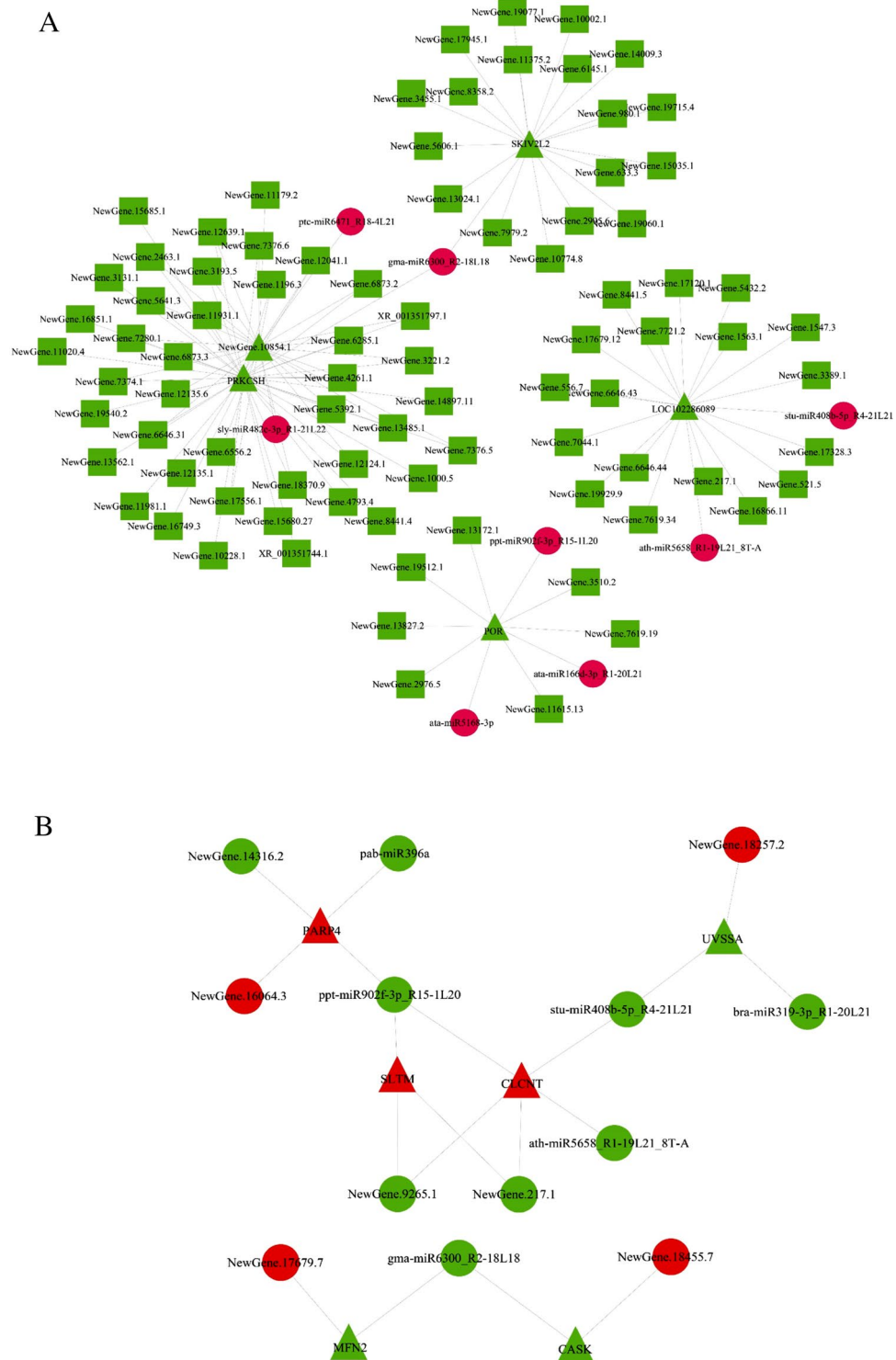




**Figure 3.** Experimental design and a computational scheme to detect key regulatory molecules and networks in lung tissues from yaks and cattles (A). Venn diagram indicated the cattles and yaks intersection (part A) (B) and the yaks living at three different altitudes intersection (part B) (C).

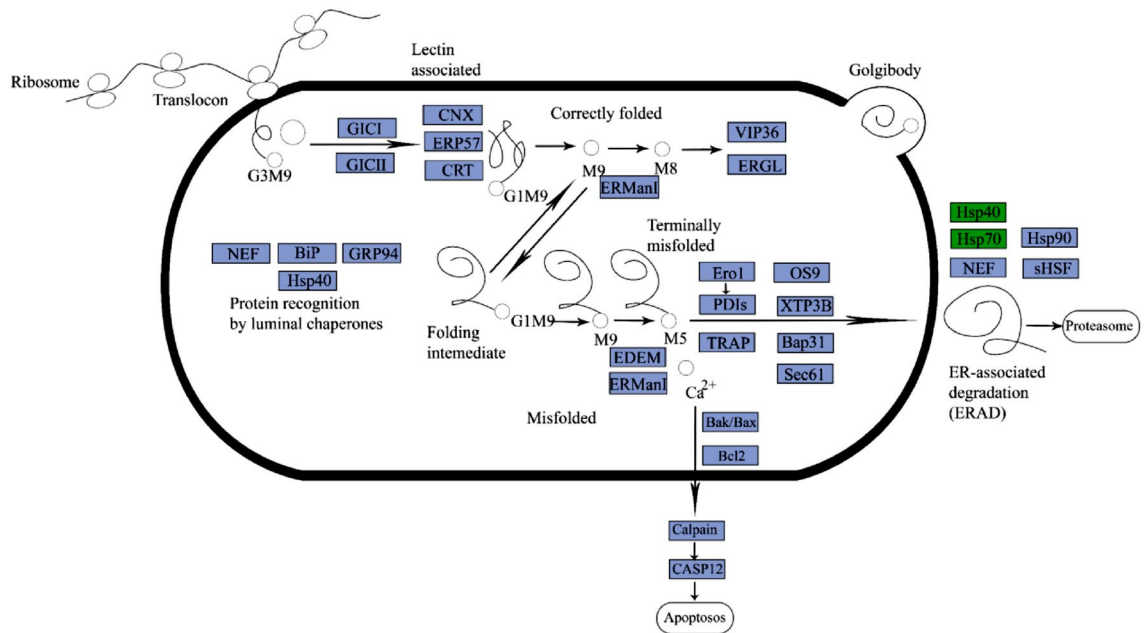
central functional organ in the respiratory system. We compared lung tissue of yaks with cattle to explore the genes associated with hypoxic adaptation. To explore the changes of regulatory mechanisms of genes in the lung tissue of yaks with the increasing of the altitude, we conducted transcriptome analysis of yaks living at three different altitudes in this study.

The transcriptomes of lung tissue in yaks and cattle were examined to identify the major biological processes and mechanisms involved in adaptation to hypoxia. GO analysis revealed that many biological processes were significantly altered. Exposure to hypoxia was mainly associated with the GO terms related to 'long-chain fatty acid metabolic process' and 'protein targeting to peroxisome'. The GO term 'protein targeting to peroxisome' is related to peroxisomes, and recent studies have shown that peroxisome proliferator activated receptor (PPAR) expression is detected in the human airway epithelium, bronchiolar submucosa, and airway smooth muscle. This finding will provide new ideas for the treatment of a number of lung-related diseases, such as pulmonary hypertension, acute lung injury, asthma and other common diseases. Therefore, PPAR has important research value for understanding the normal physiological function of the respiratory system and potential application value related to disease prevention and control. Simultaneously, PPARs can regulate the expression of multiple genes at the same time and play important roles in physiological functions such as adipocyte differentiation and regulation of lipid metabolism. Related studies showed that PPARs play an important role in the regulation of energy metabolism by long-chain fatty acids<sup>23</sup>. PPAR-related genes ACOT1 (acyl-coenzyme A thioesterase 1-like) and ACOT2 (acyl-coenzyme A thioesterase 1) were upregulated in this study, which involved in the biosynthesis and metabolism of Fatty Acyl-CoA. ACOT is a key enzyme in the process of fatty acid desaturation and elongation, which regulates the biosynthesis of long-chain fatty acids. Notably, these two genes were also upregulated in another enriched GO term in the comparison between yaks and cattle, 'long-chain fatty acid metabolic process'. Abca1 (ATP-binding cassette transporter A1) was upregulated in the 'long-chain fatty acid metabolic process'. Abca1 is a peroxisomal membrane protein that exports cellular cholesterol and phospholipids to apolipoproteins, generating nascent high-density lipoprotein (HDL) particles<sup>24</sup>. Abca1 is highly expressed in adipose tissue and is upregulated during adipogenesis<sup>25,26</sup> and its expression is critical for regulating adipose tissue cholesterol content<sup>25,27,28</sup>. PPAR is highly expressed in adipocytes and is a major regulator of adipocyte differentiation that controls adipogenesis and lipid metabolism during differentiation<sup>29,30</sup>. The role of PPAR involves transcribing and activating the expression of phosphoenolpyruvate carboxykinase (PEPC) and glycerokinase (GK) and participating in the production of triacylglycerol through different pathways. Under the action of lipoprotein lipase (LPL), PPAR adjusts the glucose content and then affects the final synthesis of triglycerides by affecting 3-phosphoglycerate (Figure S5). Therefore, PPAR plays an important role in the final synthesis of triglycerides. It is well known that there are two main types of adipose tissue, namely, white adipose tissue (WAT) and brown



**Figure 4.** Network of core RNAs in the comparison between yaks and cattles (A) and the comparison among yaks living at three different altitudes (B). Triangular nodes represent mRNAs; circular nodes represent miRNAs; square nodes represent lncRNAs. Red nodes represent the up-regulated, and the green ones represent the down-regulated transcripts. The network were analyzed using the miRanda software and were visualized by the Cytoscape V3.2 software (<http://cytoscape.org/>).

adipose tissue (BAT)<sup>30</sup>. WAT can store adipose tissue and is a very active tissue<sup>31</sup>, and BAT can burn (white) adipose tissue in the body via a heat-generating function that uncouples protein resulting in oxidation of lipid



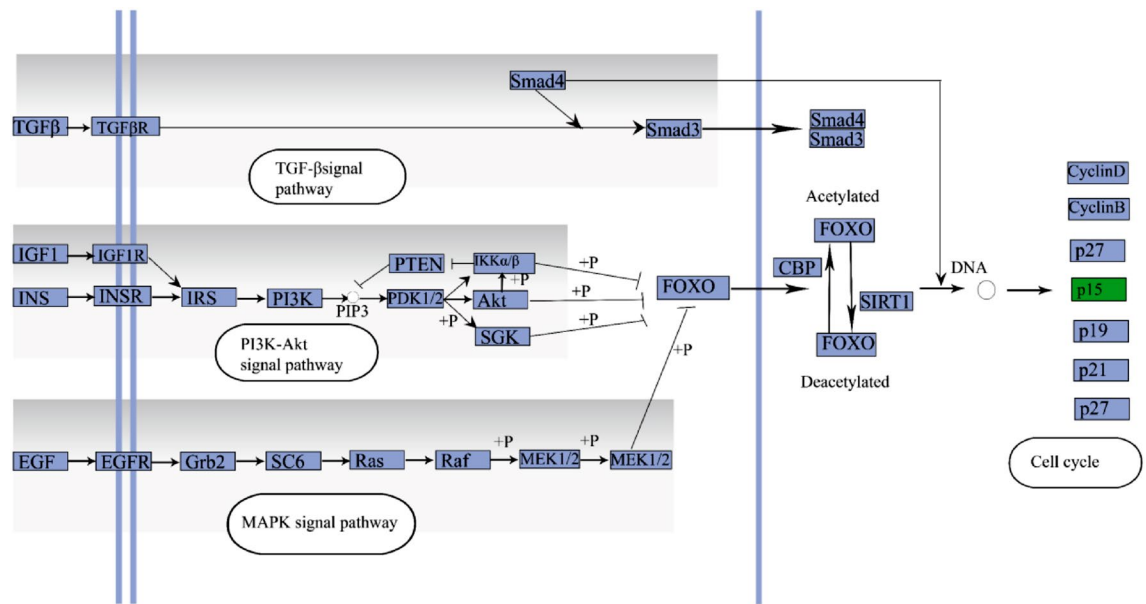
**Figure 5.** Protein processing in the endoplasmic reticulum at high altitudes (The purple boxes represent enzymes or genes, green box represent down-regulated). The pathway was visualized by the Adobe Illustrator software (<https://www.adobe.com/cn/products/illustrator.html>).

in mitochondria. BAT is essential for animals under cold plateau conditions, which is a hot area of research on the plateau animals adapt to the cold environment in the plateau area. PPAR is a key regulator of BAT cell differentiation and heat production<sup>32</sup>. The PPAR signaling pathway can promote the differentiation of mesenchymal progenitor cells into BAT cells, helping to produce heat and increase energy consumption. And there were also studies have shown that BAT expresses heat-producing markers that include PPAR<sup>33</sup>. PPAR helps animals resist the cold environment of the plateau by regulating BAT differentiation and heat production and increasing nonshivering thermogenesis to maintain a constant body temperature<sup>34</sup>. We infer that yaks maybe increase BAT cell differentiation and heat production by upregulated PPAR to resist the cold environment of the plateau.

The most significantly enriched pathway in the lungs of yaks compared to those of cattle in response to hypoxia was ‘protein processing in the endoplasmic reticulum (ER)’. Disruption at any stage of the protein process may affect the function of the protein and even impact the normal function of the whole cell due to animals suffer from the effects of cold and hypoxia in plateau environments. It is hypothesize that the protein-folding mechanisms would increase accordingly, as high altitudes may impact protein processing. In the ER, proteins are subjected to strict quality control (QC) to ensure correct folding or to facilitate degradation of misfolded polypeptides through a series of tightly regulated processes known as ER-associated degradation (ERAD). The QC system consists of calreticulin (CRT) and lectin chaperones calnexin (CNX), as well as the cochaperone ERp57, a glycoprotein-specific thiol-disulfide oxidoreductase; ‘calreticulin/calnexin cycle’ is composed of these factors (Fig. 5)<sup>35</sup>. In this study, the expression of protein disulfide isomerase A4 (PDIA4) was upregulated to promote correct protein folding. PDI may interact to recover part of the function of CNX, thus promoting the recovery and maintenance of normal lung function in yaks under high-altitude conditions<sup>36</sup>. The misfolded proteins in the ER lumen are recognized by luminal chaperones such as HSP40, NEF, GRP94 and BIP, which transfer misfolded proteins to the ER membrane for ubiquitin-dependent degradation (Fig. 5)<sup>37</sup>. Since the accumulation of misfolded proteins hinders ER function and can ultimately lead to cell death, rapid removal of misfolded polypeptides through the ERAD pathway is a critical process. ERAD is involved in chaperones of the HSP family that act synergically with ubiquitin ligases<sup>38</sup>. The HSP family is a well-known gene family that responds to heat stress. In response to low temperature, the expression levels of HSP70 and HSP40, which participate in the ERAD process, were downregulated (Fig. 5), which is likely to promote the degradation of misfolded proteins and to maintain cell homeostasis. The hypoxia-induced gene HYOU1 was also upregulated in the ‘protein processing in the ER’ pathway, indicating that hypoxia tolerance and temperature tolerance are associated with each other<sup>39</sup> and suggesting that yaks adapt to cold and hypoxia simultaneously. BAX and CAPN1, which are involved in apoptosis-related physiological processes, are upregulated by hypoxia in the ‘protein processing in the ER’ pathway. Apoptosis and programmed cell death are critical for the removal of dead cells and are linked to various biological processes. Defects in apoptosis may lead to cancer and autoimmune diseases<sup>40</sup>. The upregulation of BAX and CAPN1 expression in yaks is involved in adaptation to hypoxia. In the horrible conditions of Tibetan plateau, yaks increase protein-folding rates to maintain the function of whole cells.

In a comparison of yaks living at three different altitudes to identify which processes were significantly enriched with increasing altitude, the most overrepresented biological function in the lung in response to environmental stress was ‘immune response’, including ‘regulation of lymphocyte migration’ ‘immunoglobulin production’ ‘complement activation’ ‘regulation of complement activation’ ‘Fc-gamma receptor signaling pathway





**Figure 6.** Cell cycle at high altitudes (The purple boxes represent enzymes or genes, green box represent down-regulated). The pathway was visualized by the Adobe Illustrator software (<https://www.adobe.com/cn/products/illustrator.html>).

involved in phagocytosis' and 'leukocyte migration'. The overall atmospheric oxygen content of the Tibetan Plateau is less than 60% of that of the plain. In the state of hypoxia, the functional damage to the body caused by insufficient oxygen supply or impaired oxygen use affects normal life activities, the permeability of the yak lung tissue is increased, and the expression of various inflammatory mediators and their receptors is significantly increased. Previous studies found that many genes were related to DNA repair, blood circulation, and immunity in Tibetan pigs compared with local Chinese pig breeds<sup>41</sup>. Yaks can live in an extremely hypoxic environment for a long time without any hypoxic injury to the lungs. The significant accumulation of the complement cascade in the lungs of yaks is a good illustration of how yaks living in a hypoxic environment have developed molecular mechanisms of resistance to avoid damage caused by hypoxia to respiratory organs such as the lungs.

The FOXO signaling pathway was also enriched in a comparison of yaks living at three different altitudes. The FOXO signaling pathway can regulate the expression of the cell cycle and apoptosis-related genes to inhibit the cell cycle and promote apoptosis, thereby inhibiting cell proliferation<sup>42</sup>. The phosphoinositide-3-kinase (PI3K)/protein kinase B (Akt) signaling pathway, mitogen-activated protein kinase (MAPK) signaling pathway and transforming growth factor-beta (TGF-beta) pathway are the three main upstream pathways of FOXO (Fig. 6). The PI3K/AKT signaling pathway plays an extremely important role in FOXO regulation. Akt is the first kinase that has been shown to inhibit FOXO function. The MAPK signaling pathway is another important proliferative and antiapoptotic pathway upstream of FOXO. Its activation can affect the activity of effector molecules such as downstream cell cycle regulators and apoptosis-related proteins, leading to uncontrolled cell cycle and apoptosis abnormalities that play an important role in the occurrence, development, invasion and metastasis of tumors<sup>43,44</sup>. The MAPK pathway and PI3K pathway are two important pathways for transmitting cell proliferation signals, and these pathways are known to be related to adaptation to hypoxia<sup>18,45</sup>. There is a certain degree of interaction between these two pathways. The FOXO protein is at the intersection of these two signaling pathways and activates the activity of FOXO transcription factors involved in cell proliferation and apoptosis regulation<sup>46</sup>. After stimulation by cell proliferation signals including growth factors and other factors<sup>47</sup>, these two signaling pathways are activated, conserved sites in FOXO protein molecules are phosphorylated, this leads to the loss of transcriptional activity and the loss of regulatory effects on the cell cycle and apoptosis, thereby promoting cell proliferation<sup>48,49</sup> and influencing the cell cycle and apoptotic events via the regulation of downstream cell cycle and apoptosis-related signaling molecules<sup>50–53</sup>. Transforming growth factor-beta (TGF-beta) is a multifunctional regulatory peptide family that can control a variety of cell biological functions, including cell proliferation, differentiation, metastasis, apoptosis, adhesion and angiogenesis. TGF regulates the expression of many cell cycle regulators, such as the cytostatic factor p15, p19 and p21, by activating Smad transcription factors to make it bind to the DNA-binding domains of FOXO1, FOXO3a, and FOXO4<sup>54,55</sup>. Under hypoxic conditions, cells are exposed to environmental stresses, and apoptosis is induced by hyperpermeability of the inner mitochondrial membrane, the generation of reactive oxygen species (ROS) or the activation of stress-activated protein kinase (SAPK)<sup>56</sup>. Yaks regulate FOXO through the PI3K-AKT, MAPK and TGF-beta pathways and downregulate P15 to promote cell survival and proliferation, preventing cell apoptosis under low-oxygen stress at high altitudes (Fig. 6). Our observations related to the FOXO pathway highlight the potential molecular mechanism used to overcome hypobaric hypoxia. It has been widely reported that activating PI3K/AKT can protect cells from apoptosis in rats<sup>57</sup> and multiple cancers<sup>58</sup>. Yaks have used the same mechanism to respond to hypoxic stress.

Unlike those of malignant tumors, the growth and division of yak cells are still under control, which may provide insights for biomedical research on treatments for cancer.

After detecting abundant differentially expressed RNAs between yak and cattle, we constructed a computational scheme (Fig. 3) to identify the key regulatory molecules and networks involved in hypoxic adaptation. Here, we found many novel RNAs to be candidates for hypoxic adaptation compared with previous studies that performed microarray analyses. First, the overlap RNAs in Venn diagrams showed the intersection (part A and part B in Fig. 3) according to the experimental design. Differentially expressed RNAs in part A were majorly associated with difference between yak and cattle whereas RNAs in part B were associated with difference among yaks living at three different altitudes. Second, although the mRNA-miRNA-lncRNA interaction networks included predictable relationship between mRNA, miRNA and lncRNA according to the current database and bioinformatics analysis, the regulatory networks provided important information regarding core hypoxic adaptation-associated RNAs. Third, the core mRNAs in part A (differences of yak and cattle) predominantly encoded protein processing in endoplasmic reticulum and heme binding, such as PRKCSH and POR. Conversely, the core mRNA-encoded proteins in part B (differences of yaks living at three different altitudes) were mainly enriched in apoptosis, p53 signaling pathway, DNA repair and immunity, such as PARP4, UVSSA and MFN2. The function of core mRNAs in two intersectional parts was also confirmed with GO and KEGG analysis mentioned above, respectively.

Our study has many advantages over previous microarray about hypoxic adaptation in lung tissue of yak. Instead of focusing on single genes or a group of candidate genes previously demonstrated to be associated with hypoxic adaptation, the high-throughput sequencing method provided complete coverage and greater dynamic range of the transcriptome, and allowed an unbiased evaluation of the biological processes underlying hypoxic adaptation in lung tissue of yak. Finally, generating network interactions of coding and ncRNA enabled us to elucidate the functional complexity of the yak transcriptome and identify key regulatory molecules and networks.

## Methods

**Ethics statement.** The procedures for animal care were approved by the Gansu Agricultural University Animal Care and Use Committee of the Gansu Agricultural University, and all experiments were conducted in accordance with approved relevant guidelines and regulations. The study was carried out in compliance with the ARRIVE guidelines (<http://www.nc3rs.org.uk/page.asp?id=1357>).

**Sample collection and transcriptome sequencing.** We collected lung tissue samples from 9 indigenous adult male yaks with a gross body weight of 200 kg and an age of 4–6 years old from three different altitudes (3400 m, 4200 m and 5000 m), with three repetitions per altitude, and two indigenous adult male cattle with a gross body weight of 400 kg and an age of 4–6 years old from 1500 m as a low-altitude control. These samples included three yaks from Maqu County in the Gannan Tibetan Autonomous Prefecture of Gansu Province representing an altitude of 3400 m, three yaks from Dangxiong County, Linzhou County and Bange County representing an altitude of 4200 m, three yaks from Anduo County in the Tibetan Autonomous Region representing an altitude of 5000 m, and two indigenous Zaosheng cattle from Ningxian County in Gansu Province, as the T1, T2, T3 and CON groups, respectively (Table 2, Figure S6). Animals of breeding age are reported to have relatively stable gene expression, so we chose animals of this age. Before collecting samples, we performed a physical examination of the animals to ensure the health of the animals. All tissues were immediately frozen in liquid nitrogen until RNA extraction. Total RNA was extracted using TRIzol reagent, and the purity of the isolated RNA was determined by agarose gel electrophoresis. Libraries were prepared from the resulting total RNA and sequenced on the Illumina HiSeq 2500 platform with paired-end 150 bp reads. For each samples, 10G of data were generated.

**Transcriptome assembly.** RNA-seq reads were preprocessed, and low-quality areas that affected the data quality were removed for subsequent analysis. The raw data were filtered under a series of steps as follow. For RNA data, The BWA algorithm was used to remove the low quality reads with a mass threshold of 30. Then joint sequence, unknown base calls (N) and sequences of length less than 60 bp were trimmed out from RNA raw data. For the raw data of miRNA, unknown base calls (N), reads without 3' adapters, low quality reads (> 15% of bases whose Phred scores were  $\leq 20$ ), reads containing poly-A/T or the length was not within the required range (15–34nt) were filtered to generate clean data. The clean reads were mapped to the reference genome [ftp://ftp.ncbi.nlm.nih.gov/genomes/refseq/vertebrate\\_mammalian/Bos\\_mutus/annotation\\_releases/current/GCF\\_000298355.1\\_BosGru\\_v2.0/GCF\\_000298355.1\\_BosGru\\_v2.0\\_genomic.fna.gz](ftp://ftp.ncbi.nlm.nih.gov/genomes/refseq/vertebrate_mammalian/Bos_mutus/annotation_releases/current/GCF_000298355.1_BosGru_v2.0/GCF_000298355.1_BosGru_v2.0_genomic.fna.gz) using HISAT2 (v2.0.5) (<http://ccb.jhu.edu/software/hisat2/index.shtml>). The resulted bam files of each sample were feed to StringTie (v1.3.2d) to obtain transcript information of each sample, then merged all transcript to get results of assembly.

**Identification of lncRNA and miRNA.** Potential lncRNAs were identified from the assembled transcripts through the following highly stringent criterion: (1) transcript length is not less than 200 nt; (2) transcript expression is more than 3 reads; (3) transcripts with reads coverage of less than 5 were eliminated; (4) compared with the annotation file of the species to screen the known mRNA and other non-coding RNA (rRNA, tRNA, snoRNA and snRNA) according to the gffcompare software; (5) according to the information of class code (“u”, “i”, “x”), the potential lincRNA, intronic lncRNA and anti-sense lncRNA were screened; (6) the calculation of the protein-coding potency of transcripts was performed using four software programs: CNCI (Coding-Non-Coding-Index) (score < 0), CPC (Coding Potential Calculator) (score < 0), Pfam-scan (E-value < 0.001), and PhyloCSF. Predicted transcripts with coding potential in any or all of the four tools were filtered out.

For miRNA, unique reads of Quality Control first aligned to the Rfam RNA<sup>59</sup> families database to filter other small RNA species. BLAST parameters: (1) E value no more than 1; (2) Number of mismatched bases no more than 1; (3) Score  $\geq 15$  (At least 15 bases). Compare the mature miRNA sequences of this species in the miRBase database<sup>60</sup>. Those that could not be compared were then compared with the mature miRNA sequences of other species in the miRBase database to identify known miRNA from sequencing data. For sequences not identified as known miRNA, bowtie2<sup>61</sup> was first used to locate them on the reference genome. Then MIREAP was used to predict the novel miRNA based on the results of localization.

**Qualification and differential analysis of RNAs expression level.** Previously, read counts of each sample were obtained, for mRNA and lncRNA, FPKM (Fragments Per Kilobase of transcript per Million fragments mapped) of each samples were obtained, for miRNA, RPM (reads per million) values were calculated according to total counts of each sample. Then completed the analysis of differentially expressed RNAs (mRNA, lncRNA and miRNA) between CON vs T1, CON vs T2, CON vs T3, T1 vs T2, T1 vs T3, T2 vs T3 by DESeq2 (<http://www.bioconductor.org/packages/release/bioc/html/DESeq.html>), considering calculated P value accomplished multiple hypothesis testing (Benjamini Y, Hochberg Y 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57: 12. ), therefore, an adjusted P value  $< 0.05$  and  $|\log_2(\text{foldchange})| > 1$  was applied to detect differential expression RNAs<sup>62</sup>.

**Target gene prediction of lncRNAs and miRNAs.** For lncRNAs, mRNAs with a high Spearman correlation coefficient ( $P \geq 0.9$ ) were selected as the trans-targets. mRNAs with distances less than 50 kb were selected as cis-targets. For miRNAs, miRanda<sup>63</sup> and RNAhybrid<sup>64</sup> was used to predict targets of known or novel miRNAs. The principle of miRanda prediction was based on seed region sequence alignment. The results were filtered by the parameters -sc 160 -en -20.

**Enrichment analysis.** For the analysis of differentially expressed RNAs (mRNA, miRNA, lncRNA), we applied clusterProfiler (Guangchuang Yu [guangchuangyu@gmail.com](mailto:guangchuangyu@gmail.com)) Packag in R to identify gene ontology (GO) and KEGG pathway categories, a cutoff value of 0.05 from the chi-square test to filter significantly enriched GO and KEGG pathway categories<sup>65,66</sup>. All tests completed multiple hypothesis testing according to Benjamini-Hochberg approach (Benjamini Y, Hochberg Y 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57: 12).

**Quantitative real-time (RT)-qPCR analysis.** RT-qPCR was performed on 7 mRNAs and 5 lncRNAs elected from the RNAs sequencing data according to potential functional importance. Primers were designed according to sequencing data from the lung transcriptome of yaks using Primer Premier 5.0. PCR amplification and specificity were examined via melting curve analysis. The relative expression of target gene transcripts was calculated using the comparative Ct method, and SPSS software was used for statistical analysis.

### Data availability

Sequence data have been deposited with the GenBank Data Libraries under Accession GSE153956 and GSE153962.

Received: 29 October 2020; Accepted: 19 March 2021

Published online: 08 April 2021

### References

1. Simonson, T. S. *et al.* Genetic evidence for high-altitude adaptation in Tibet. *Science* **329**, 72–75 (2010).
2. Yi, X. *et al.* Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75–78 (2010).
3. Peng, Y. *et al.* Genetic variations in Tibetan populations and high-altitude adaptation at the Himalayas. *Mol. Biol. Evol.* **28**, 1075–1081. <https://doi.org/10.1093/molbev/msq290> (2011).
4. Xu, S. *et al.* A genome-wide search for signals of high-altitude adaptation in Tibetans. *Mol. Biol. Evol.* **28**, 1003–1011. <https://doi.org/10.1093/molbev/msq277> (2011).
5. Qiu, Q. *et al.* The yak genome and adaptation to life at high altitude. *Nat. Genet.* **44**, 946–949. <https://doi.org/10.1038/ng.2343> (2012).
6. Cho, Y. S. *et al.* The tiger genome and comparative analysis with lion and snow leopard genomes. *Nat. Commun.* **4**, 2433. <https://doi.org/10.1038/ncomms3433> (2013).
7. Ge, R. L. *et al.* Draft genome sequence of the Tibetan antelope. *Nat. Commun.* **4**, 1858. <https://doi.org/10.1038/ncomms2860> (2013).
8. Li, M. *et al.* Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. *Nat. Genet.* **45**, 1431–1438. <https://doi.org/10.1038/ng.2811> (2013).
9. Li, Y. *et al.* Population variation revealed high-altitude adaptation of tibetan mastiffs. *Mol. Biol. Evol.* **31**, 1200–1205. <https://doi.org/10.1093/molbev/msu070> (2014).
10. Zhang RC. China Yak. Lanzhou, P.R. China: Gansu Scientific & Technology Press, ISBN:7-5424-0237-4 (1989).
11. Wiener G, Jianlin H, Ruijun L. The Yak. Bangkok, Thailand: The Regional Office for Asia and the Pacific, Food and Agriculture Organization of the United Nations, <http://www.fao.org/3/ad347e/ad347e00.htm> (2003).
12. Graf, A. *et al.* Fine mapping of genome activation in bovine embryos by RNA sequencing. *Proc. Natl. Acad. Sci. USA* **111**, 4139–4144. <https://doi.org/10.1073/pnas.1321569111> (2014).
13. Reed, R. D. *et al.* Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* **333**, 1137–1141. <https://doi.org/10.1126/science.1208227> (2011).
14. Chen, F. *et al.* Transcriptome and network changes in climbers at extreme altitudes. *PLoS ONE* **7**, e31645. <https://doi.org/10.1371/journal.pone.0031645> (2012).
15. Malik, A. *et al.* Transcriptome analysis of the spalax hypoxia survival response includes suppression of apoptosis and tight control of angiogenesis. *BMC Genom.* **13**, 615. <https://doi.org/10.1186/1471-2164-13-615> (2012).

16. Choudhry, H. *et al.* Extensive regulation of the non-coding transcriptome by hypoxia: role of HIF in releasing paused RNAPol2. *EMBO Rep.* **15**, 70–76. <https://doi.org/10.1002/embr.201337642> (2014).
17. Wang, K. *et al.* Different gene expressions between cattle and yak provide insights into high-altitude adaptation. *Sticht. Int. Found. Anim. Gen.* **47**, 28–35. <https://doi.org/10.1111/age.12377> (2015).
18. Xuebin, Q. *et al.* The transcriptomic landscape of Yaks reveals molecular pathways for high altitude adaptation. *Genome Biol. Evol.* **11**, 72–85. <https://doi.org/10.1093/gbe/evy264> (2018).
19. Wu, D.-D. *et al.* Pervasive introgression facilitated domestication and adaptation in the Bos species complex. *Nat. Ecol. Evol.* **2**, 1139–1145. <https://doi.org/10.1038/s41559-018-0562-y> (2018).
20. Perte, M., Kim, D., Perte, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667. <https://doi.org/10.1038/nprot.2016.095> (2016).
21. Friedlander, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **40**, 37–52. <https://doi.org/10.1093/nar/gkr688> (2012).
22. Ai, H. Population history and genomic signatures for high-altitude adaptation in Tibetan pigs. *BMC Genom.* **15**, 834. <https://doi.org/10.1186/1471-2164-15-834> (2014).
23. Akbar, H. *et al.* Dietary lipid during late-pregnancy and early-lactation to manipulate metabolic and inflammatory gene network expression in dairy cattle liver with a focus on PPARs. *Gene Regul. Syst. Bio.* **7**, 103–123. <https://doi.org/10.4137/GRSB.S12005> (2013).
24. Zhu, X. & Parks, J. S. New roles of HDL in inflammation and hematopoiesis. *Annu. Rev. Nutr.* **32**, 161–182. <https://doi.org/10.1146/annurev-nutr-071811-150709> (2012).
25. Chung, S., Sawyer, J. K., Gebre, A. K., Maeda, N. & Parks, J. S. Adipose tissue ATP binding cassette transporter A1 contributes to high-density lipoprotein biogenesis in vivo. *Circulation* **124**, 1663–1672. <https://doi.org/10.1161/CIRCULATIONAHA.111.025445> (2011).
26. Wellington, C. L. *et al.* ABCA1 mRNA and protein distribution patterns predict multiple different roles and levels of regulation. *Lab. Invest.* **82**, 273–283. <https://doi.org/10.1038/labinvest.3780421> (2002).
27. Zhang, Y. *et al.* Adipocyte modulation of high-density lipoprotein cholesterol. *Circulation* **121**, 1347–1355. <https://doi.org/10.1161/CIRCULATIONAHA.109.897330> (2010).
28. Cuffe, H. *et al.* Targeted deletion of adipocyte abca1 (ATP-binding cassette transporter A1) impairs diet-induced obesity. *Arterioscler. Thromb. Vasc. Biol.* **38**, 733–743. <https://doi.org/10.1161/ATVBAHA.117.309880> (2018).
29. Zhang, H., You, L. & Zhao, M. Rosiglitazone attenuates paraquat-induced lung fibrosis in rats in a PPAR gamma-dependent manner. *Eur. J. Pharmacol.* **851**, 133–143. <https://doi.org/10.1016/j.ejphar.2019.02.037> (2019).
30. Araki, M. *et al.* The peroxisome proliferator-activated receptor alpha (PPAR alpha) agonist pemafibrate protects against diet-induced obesity in mice. *Int. J. Mol. Sci.* **19**, 2148. <https://doi.org/10.3390/ijms19072148> (2018).
31. Festuccia, W. T. & Deshaies, Y. Depot specificities of PPARy ligand actions on lipid and glucose metabolism and their implication in PPARy-mediated body fat redistribution. *Clin. Lipidol.* **4**, 633–642. <https://doi.org/10.2217/clp.09.45> (2009).
32. Chang, J. S. & Ha, K. A truncated PPAR gamma 2 localizes to mitochondria and regulates mitochondrial respiration in brown adipocytes. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0195007> (2018).
33. Contreras, C. *et al.* Hypothalamus and thermogenesis: Heating the BAT, browning the WAT. *Mol. Cell Endocrinol.* **438**, 107–115. <https://doi.org/10.1016/j.mce.2016.08.002> (2016).
34. Storz, J. *et al.* Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. *Proc. Natl. Acad. Sci. U S A* **106**, 14450–14455. <https://doi.org/10.1073/pnas.0905224106> (2009).
35. Gupta, G. S. Lectins in quality control: calnexin and calreticulin. *Springer Vienna* **5**, 29–56. [https://doi.org/10.1007/978-3-7091-1065-2\\_2](https://doi.org/10.1007/978-3-7091-1065-2_2) (2012).
36. Zhang, H., He, J., Ji, Y., Kato, A. & Song, Y. The effect of calnexin deletion on the expression level of PDI in *Saccharomyces cerevisiae* under heat stress conditions. *Cell. Mol. Biol. Lett.* **13**, 38–48. <https://doi.org/10.2478/s11658-007-0033-y> (2008).
37. Plemper, R. K., Böhmeler, S., Bordallo, J., Sommer, T. & Wolf, D. H. Mutant analysis links the translocon and BiP to retrograde protein transport for ER degradation. *Nature* **388**, 891–895. <https://doi.org/10.1038/42276> (1997).
38. Goldberg, A. L. Protein degradation and protection against misfolded or damaged proteins. *Nature* **426**, 895–899. <https://doi.org/10.1038/nature02263> (2003).
39. Pörtner, H. O. & Peck, M. A. Climate change effects on fishes and fisheries: towards a cause-and-effect understanding. *Fish. Biol.* **77**, 1745–1779. <https://doi.org/10.1111/j.1095-8649.2010.02783.x> (2010).
40. Czabotar, P. E., Lessene, G., Strasser, A. & Adams, J. M. Control of apoptosis by the BCL-2 protein family: implications for physiology and therapy. *Nat. Rev. Mol. Cell Biol.* **15**, 49–63. <https://doi.org/10.1038/nrm3722> (2014).
41. Li, M. *et al.* Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. *Nat. Genet.* **45**, 1431–1438. <https://doi.org/10.1038/ng.2811> (2013).
42. Sisci, D. *et al.* The estrogen receptor is the key regulator of the bifunctional role of FoxO3a transcription factor in breast cancer motility and invasiveness. *Cell Cycle* **12**, 3405–3420. <https://doi.org/10.4161/cc.26421> (2013).
43. Chen, P. N. *et al.* Silibinin inhibits invasion of oral cancer cells by suppressing the MAPK pathway. *J. Dent. Res.* **85**, 220–225. <https://doi.org/10.1177/154405910608500303> (2006).
44. Hill, K. *et al.* (2000) Specific requirement for the p85-p110alpha phosphatidylinositol 3-kinase during epidermal growth factor-stimulated actin nucleation in breast cancer cells. *J. Biol. Chem.* **275**, 3741–3744. <https://doi.org/10.1074/jbc.275.6.3741> (2000).
45. Sang, N. *et al.* MAPK signaling up-regulates the activity of hypoxia-inducible its effects on p300. *J. Biol. Chem.* **278**, 14013–14019. <https://doi.org/10.1074/jbc.M209702200> (2003).
46. Roy, S. K., Srivastava, R. K. & Shankar, S. Inhibition of PI3K/AKT and MAPK/ERK pathways causes activation of FOXO transcription factor, leading to cell cycle arrest and apoptosis in pancreatic cancer. *J. Mol. Signal* **5**, 10. <https://doi.org/10.1186/1750-2187-5-10> (2010).
47. Boscolo, E. *et al.* AKT hyper-phosphorylation associated with PI3K mutations in lymphatic endothelial cells from a patient with lymphatic malformation. *Angiogenesis* **18**, 151–162. <https://doi.org/10.1007/s10456-014-9453-2> (2015).
48. Obsilova, V. *et al.* 14-3-3 Protein interacts with nuclear localization sequence of forkhead transcription factor FoxO4. *Biochemistry* **44**, 11608–11617. <https://doi.org/10.1021/bi050618r> (2005).
49. Kim, Y. S. *et al.* Decreased interaction between FoxO3a and Akt correlates with seizure-induced neuronal death. *Epilepsy Res.* **108**, 367–378. <https://doi.org/10.1016/j.eplepsyres.2014.01.003> (2014).
50. Ramjaun, A. R. *et al.* Upregulation of two BH3-only proteins, Bmf and Bim, during TGF beta-induced apoptosis. *Oncogene* **26**, 970–981. <https://doi.org/10.1038/sj.onc.1209852> (2007).
51. Kuiperij, H. B., Van der Horst, A. & Raijmakers, J. Activation of Foxo transcription factors contributes to the antiproliferative effect of cAMP. *Oncogene* **24**, 2087–2095. <https://doi.org/10.1038/sj.onc.1208450> (2005).
52. Weidinger, C. *et al.* Forkhead box-O transcription factor: critical conductors of cancer's fate. *Endocr. Relat. Cancer* **15**, 917–929. <https://doi.org/10.1677/ERC-08-0153> (2008).
53. Gilley, J. *et al.* FOXO transcription factors directly activate bim gene expression and promote apoptosis in sympathetic neurons. *Cell Biol.* **162**, 613–622. <https://doi.org/10.1083/jcb.200303026> (2003).
54. Palazuelos, J., Klingener, M. & Agurru, A. Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation. *Cell* **117**, 211–223. [https://doi.org/10.1016/s0092-8674\(04\)00298-3](https://doi.org/10.1016/s0092-8674(04)00298-3) (2004).

55. Nakayama, M. *et al.* Attenuation of the PI3kinase/Akt signaling pathway by *Porphyromonas gingivalis* gingipains Rgp A, Rgp B, and Kgp. *J. Biol. Chem.* **290**, 5190–5202. <https://doi.org/10.1074/jbc.M114.591610> (2015).
56. Greijer, A. E. & van der Wall, E. The role of hypoxia inducible factor 1 (HIF-1) in hypoxia induced apoptosis. *J Clin Pathol* **57**, 1009–1014. <https://doi.org/10.1136/jcp.2003.015032> (2004).
57. Alvarez-Tejado, M. *et al.* Hypoxia induces the activation of the phosphatidylinositol 3-kinase/Akt cell survival pathway in PC12 cells: protective role in apoptosis. *J. Biol. Chem.* **276**, 22368–22374. <https://doi.org/10.1074/jbc.M011688200> (2001).
58. Kilic-Eren, M., Boylu, T. & Tabor, V. Targeting PI3K/Akt represses Hypoxia inducible factor-1alpha activation and sensitizes Rhabdomyosarcoma and Ewing's sarcoma cells for apoptosis. *Cancer Cell Int.* **13**, 36. <https://doi.org/10.1186/1475-2867-13-36> (2013).
59. Burge, S. W. *et al.* Rfam 110: 10 years of RNA families. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gks1005> (2012).
60. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *NAR* **42**, D68–D73. <https://doi.org/10.1093/nar/gkt1181> (2014).
61. Langmead, B. & Salzberg, S. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359. <https://doi.org/10.1038/nmeth.1923> (2012).
62. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, 106. <https://doi.org/10.1186/gb-2010-11-10-r106> (2010).
63. Enright, A. J. *et al.* MicroRNA targets in *Drosophila*. *Genome Biol.* **5**, R1. <https://doi.org/10.1186/gb-2003-5-1-r1> (2003).
64. Rehmsmeier, M., Stefffen, P., Hoeschmann, M. & Giegerich, R. Fast and effective prediction of microRNA/target duplexes. *RNA* **10**, 1507–1517. <https://doi.org/10.1261/rna.5248604> (2004).
65. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29. <https://doi.org/10.1038/75556> (2000).
66. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. <https://doi.org/10.1093/nar/28.1.27> (2000).

## Acknowledgements

This work was supported by the Discipline construction fund project of Gansu Agricultural University (GAU-XKJS-2018-049).

## Author contributions

S.Z. conceived and designed the experiments; Q.G. performed the experiments; Y.C. and X.Q. analyzed the data; Y.G. and W.Z. contributed reagents/materials/analysis tools; Q.G. wrote the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-87420-7>.

**Correspondence** and requests for materials should be addressed to S.Z., Y.C. or X.Q.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021