Article

The Landscape of Accessible Chromatin during Yak Adipocyte Differentiation

Zhilong Zhang [1], Yongfeng Zhang [2,3], Qi Bao [2,3], Yarong Gu [2,3], Chunnian Liang [2,3], Min Chu [2,3], Xian Guo [2,3], Pengjia Bao [2,3] and Ping Yan [1,*]

[1] College of Animal Science and Technology, Gansu Agricultural University, Lanzhou 730070, China
[2] Key Laboratory of Yak Breeding Engineering Gansu Province, Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences, Lanzhou 730050, China
[3] Key Laboratory of Animal Genetics and Breeding on Tibetan Plateau, Ministry of Agriculture and Rural Affairs, Lanzhou 730050, China
[*] Correspondence: yanping@caas.cn; Tel.: +86-931-216-4180

Abstract: Although significant advancement has been made in the study of adipogenesis, knowledge about how chromatin accessibility regulates yak adipogenesis is lacking. We here described genome-wide dynamic chromatin accessibility in preadipocytes and adipocytes by using the assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq), and thus revealed the unique characteristics of open chromatin during yak adipocyte differentiation. The chromatin accessibility of preadipocytes and adipocytes exhibited a similar genomic distribution, displaying a preferential location within the intergenic region, intron, and promoter. The pathway enrichment analysis identified that genes with differential chromatin accessibility were involved in adipogenic metabolism regulation pathways, such as the peroxisome proliferator activated receptor-γ (PPAR) signaling pathway, wingless-type MMTV integration site (Wnt) signaling pathway, and extracellular matrix-receptor (ECM–receptor) interaction. Integration of ATAC-seq and mRNA-seq revealed that genes with a high expression were associated with high levels of chromatin accessibility, especially within 1 kb upstream and downstream of the transcription start site. In addition, we identified a series of transcription factors (TFs) related to adipogenesis and created the TF regulatory network, providing the possible interactions between TFs during yak adipogenesis. This study is crucial for advancing the understanding of transcriptional regulatory mechanisms of adipogenesis and provides valuable information for understanding the adaptation of plateau species to high-altitude environments by maintaining whole body homeostasis through fat metabolism.

Keywords: ATAC-seq; chromatin accessibility; adipogenesis; yak

1. Introduction

Adipose tissue is a vital endocrine organ for energy storage and metabolism in animals, and the growth of this tissue refers to the increase in adipocyte number and size [1]. Mature adipocytes differentiate from preadipocytes under the action of a series of transcription factors (TFs), fatty acid-binding proteins, and lipid-metabolizing enzymes after growth arrest. The TFs cascade occurs prior to adipocyte gene expression during adipocyte differentiation [2]. PPARγ (peroxisome proliferator-activated receptor-γ) and C/EBP (CCAAT/enhancer binding protein) have been confirmed as pioneer TFs for inducing initial adipocyte differentiation [3]. In addition, many signaling pathways, such as insulin, glucocorticoid, BMP (bone morphogenetic protein), Wnt (wingless-type MMTV integration site) and Hedgehog signaling, have been reported to promote or inhibit differentiation of preadipocytes into adipocytes [4–8]. Numerous studies have found the regulatory role of TFs in bovine adipocyte differentiation. For example, by studying the mechanism of TFs regulating SIRT4 (silent information regulator Protein 4) gene during bovine adipogenesis, it was found that E2F1 (E2F transcription factor-1), C/EBPβ, and

HOXA5 (homeobox A5) positively regulate the expression of SIRT4, while IRF4 (interferon regulatory factor 4), PAX4 (paired box 4), and CREB1 (cAMP responsive element-binding protein 1) suppressed its expression [9]. A previous study investigated the core promoter region of the PLIN1 (perilipin1) in bovine adipogenesis and found that E2F1, PLAG1 (pleiomorphic adenoma gene 1), C/EBPβ, and SMAD3 (SMAD family member 3) can bind to this region and activate or inhibit its transcription [10]. A similar study found that KLF15 (Krüppel-like factors 15) plays a regulatory role by binding to the core promoter region of KLF3 in bovine adipogenesis [11]. However, knowledge about genome-wide TFs regulating bovine adipocyte differentiation and the interaction network between these TFs is lacking. The assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) is a method that can be used to find the open chromatin region, predict TFs, and map the nucleosome position [12,13]. Chromatin in mammalian cells is usually in active euchromatin and inactive heterochromatin states. When chromatin is in the active euchromatin state, open chromatin regions with DNA regulatory elements can be bound by TFs, which will then contribute to gene regulation. When chromatin is in the inactive heterochromatin state, TFs cannot bind to transcriptionally active regions of genes, which is not conducive to gene expression [14,15]. Therefore, ATAC-seq data analysis can identify genome-wide transcription regulatory elements and transcriptionally active genomic regions in specific biological processes.

The yak is a unique and rare ruminant living in the Qinghai–Tibet Plateau and adjacent regions. As an important livestock species for animal husbandry in this area, yaks provide local residents with agricultural products such as meat, milk, and hair while adapting to environmental conditions including cold, low oxygen, and high altitude and ultraviolet rays [16,17]. Because yaks graze year-round, their growth and development is restricted by the yield and quality of forage [18]. During the warm season, the plentiful and high-quality forage feed induces weight gain and fat deposition in yaks. However, during the long cold season, yaks have insufficient nutrient intake and need to consume the fat stored in the warm season to resist severe cold and maintain the normal metabolism [19,20]. This mechanism of maintaining the normal biological process and development of the body through fat metabolism on a plateau makes the yak an ideal model for studying the adaptation of plateau species. We here used ATAC-seq to identify open chromatin regions in yak preadipocytes and adipocytes. We also applied RNA-seq to detect gene expression during yak adipocyte differentiation. Our study objectives were to (i) compare chromatin accessibility and gene expression to identify differential chromatin accessibility and differentially expressed genes (DEGs) during yak adipocyte differentiation; (ii) explore the correlation between chromatin accessibility and gene expression; and (iii) predict TF regulatory networks during yak adipogenesis. Our results provide new insights into yak adipogenesis and further reveal the mechanisms through which open chromatin regulates gene expression during biological processes.

2. Results

2.1. Landscape of Chromatin Accessibility during Yak Adipocyte Differentiation

To determine whether preadipocytes fully differentiated into adipocytes, Oil Red O staining and mRNA detection of adipogenic marker genes were performed. The results of Oil Red O staining revealed that preadipocytes were fully differentiated into mature lipid-filled adipocytes after induction with adipogenic agents for 12 days (Figures 1A,B and S1). At the same time, the expression of C/EBPα, PPARγ, FABP4 (fatty acid binding protein 4), and SREBF1 (sterol regulatory element-binding factor-1) was significantly higher in adipocytes than in preadipocytes, which also supports this finding (Figure 1C).
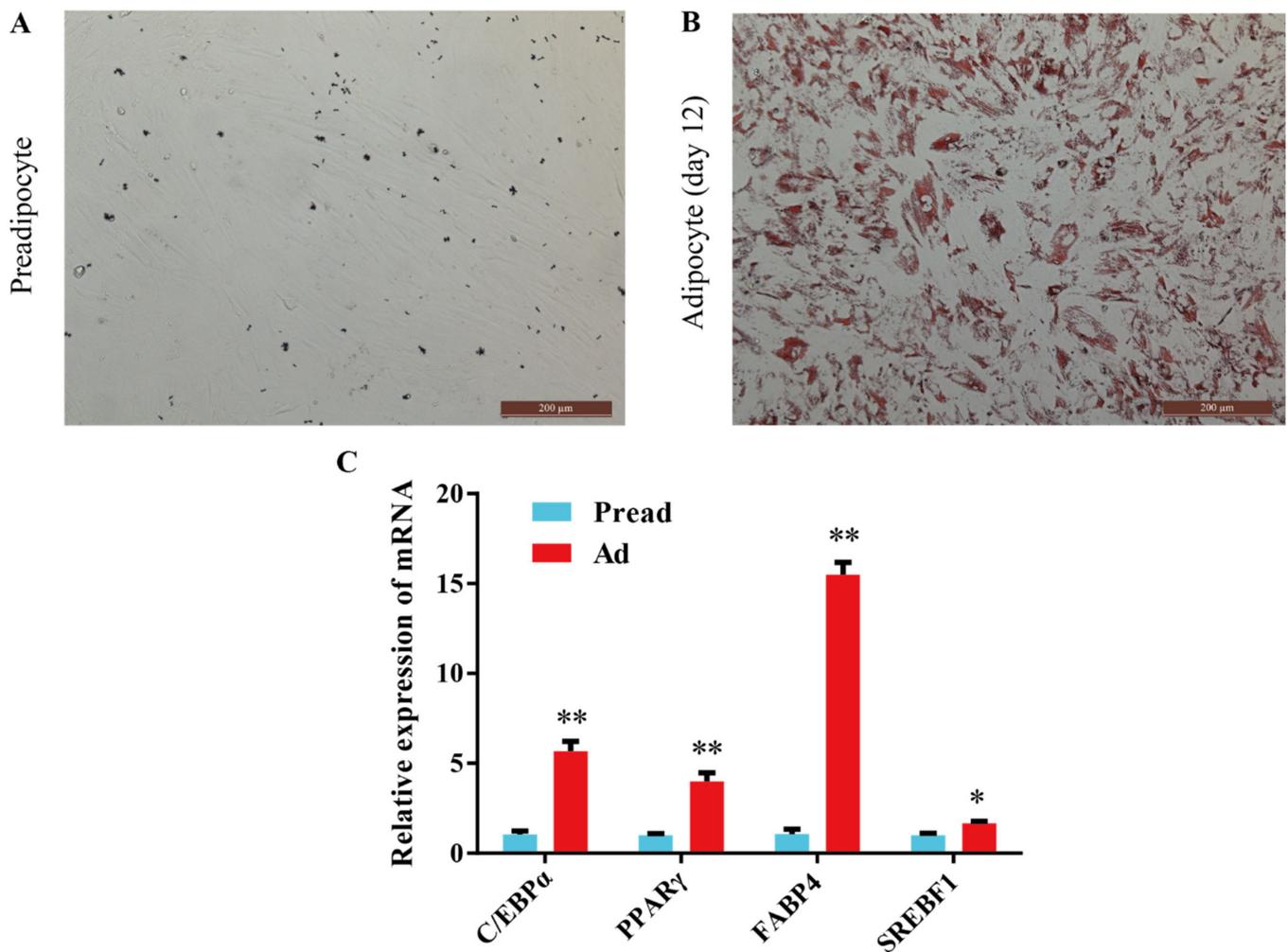
Figure 1. Induction of adipocyte differentiation. (**A,B**) Oil red O staining of preadipocytes and adipocytes. (**C**) The mRNA expression level of C/EBPα, PPARγ, FABP4 and SREBF1 between the preadipocytes and adipocytes (mean ± SEM, n = 3, * $p < 0.05$, ** $p < 0.01$).

To explore epigenetic regulation during yak adipocyte differentiation, we used ATAC-seq to identify open chromatin regions in preadipocytes and adipocytes (three biological replicates for each cell group). An average of 167,249,845 and 159,616,847 high-quality reads were obtained from Pread (preadipocytes) and Ad (adipocytes) groups, respectively. The reads information of each ATAC-seq library before and after filtering is shown in Table S1. After the reads were aligned to LU_Bosgru_v3.0 using Bowtie2, we selected the reads aligned to unique positions for downstream analyses. Because Tn5 transposase can preferentially insert into open chromatin regions, most reads are short fragments with no or only one nucleosome. Some long fragments containing multiple nucleosomes are also present. Analysis of fragment size distribution revealed that the length of most fragments was <100 bp, indicating that the reads of each library were primarily located in the open chromatin regions (Figure 2A). Statistical analysis of all reads in the 2 kb upstream and downstream of transcription start site (TSS) by using deep Tools software showed that the highest density was observed at the TSS, suggesting that the reads were distributed according to the typical characteristics of higher eukaryotes (Figure 2B). Genome-wide peak scanning was performed using MACS software to identify regions of significant enrichment in the open chromatin. In the Pread and Ad groups, 15,045 and 7683 peaks were found, respectively. Approximately 35% of peaks detected in the Pread group were also present in the Ad group. Similarly, approximately 69% of peaks detected in the Ad

group were also present in the Pread group (Figure 2C). The genomic distribution of Pread peaks showed that 43.83% of the peaks were located within the distal intergenic region, 30.48% located within the intron, 16.97% located within the promoter (2 kb upstream of the TSS), 4.13% located within 5′UTR, 2.99% located within the exon, 1.22% located within 500 bp downstream of the gene, and 0.39% located within 3′UTR. The peaks detected in the Ad group exhibited a similar genomic distribution, showing a preferential location within the distal intergenic region (48.44%), followed by the intron (26.04%), promoter (17.19%), 5′UTR (4.10%), exon (2.59%), downstream of the gene (1.17%), and 3′UTR (0.46%) (Figure 2D and Table S2).
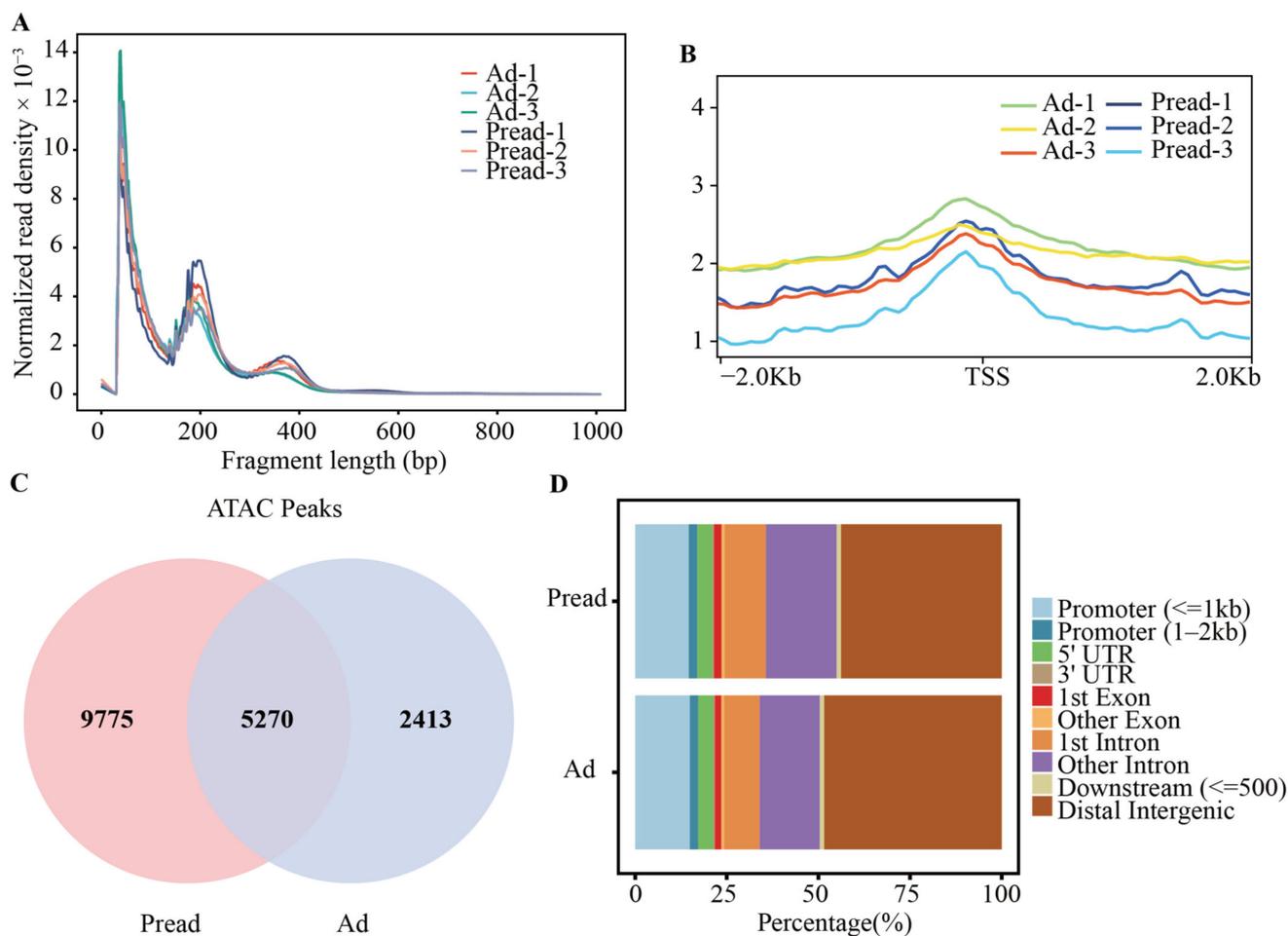

Figure 2. The characteristics of accessible chromatin in preadipocyte and adipocyte. (**A**) The fragment size distribution of each sample. A large proportion fragment size was less than 100 bp, which represents the regions of chromatin accessibility. The peaks at the 200 bp and 400 bp positions represent mononucleosomes and dinucleosomes. (**B**) Distribution of all reads relative to the position of TSS ± 2 Kb. The ordinate represents the ATAC-seq signal intensity. (**C**) Venn diagram showing the peak overlap between Pread and Ad groups. (**D**) The distribution of peaks among different genomic regions (promoter, 5′UTR, 3′UTR, exon, intron, downstream and distal intergenic) in Pread and Ad groups.

2.2. Differential Chromatin Accessibility during Yak Adipocyte Differentiation

To demonstrate the role of open chromatin regions during adipocyte differentiation, we performed differential chromatin accessibility analysis between the Pread and Ad groups by using DiffBind ($|\log2FC| > 1$, $p < 0.05$). The results revealed 1293 differential peaks, of which 1265 were downregulated and 28 were upregulated (Figures 3A and S2A and Table S3). GO term analysis exhibited that these differential peaks were mainly implicated with cellular transcriptional regulatory and metabolic activities, such as binding,

transcription regulator activity, and molecular transducer activity (ontology: molecular function), metabolic process, regulation of the biological process, developmental process and signaling (ontology: biological process), and cell part and protein-containing complex (ontology: cellular component) (Figure 3C and Table S4). Meanwhile, KEGG enrichment analysis revealed that differential chromatin accessibility was closely correlated to adipogenic metabolism regulation pathways including the PPAR signaling pathway, Wnt signaling pathway, regulation of actin cytoskeleton, and ECM–receptor interaction (Figure 3B and Table S5). The genetic information about these signaling pathways is presented in Figure S2B and Table S6. These results suggested that changes in gene transcription induced by altered chromatin accessibility can regulate preadipocyte differentiation into adipocytes.
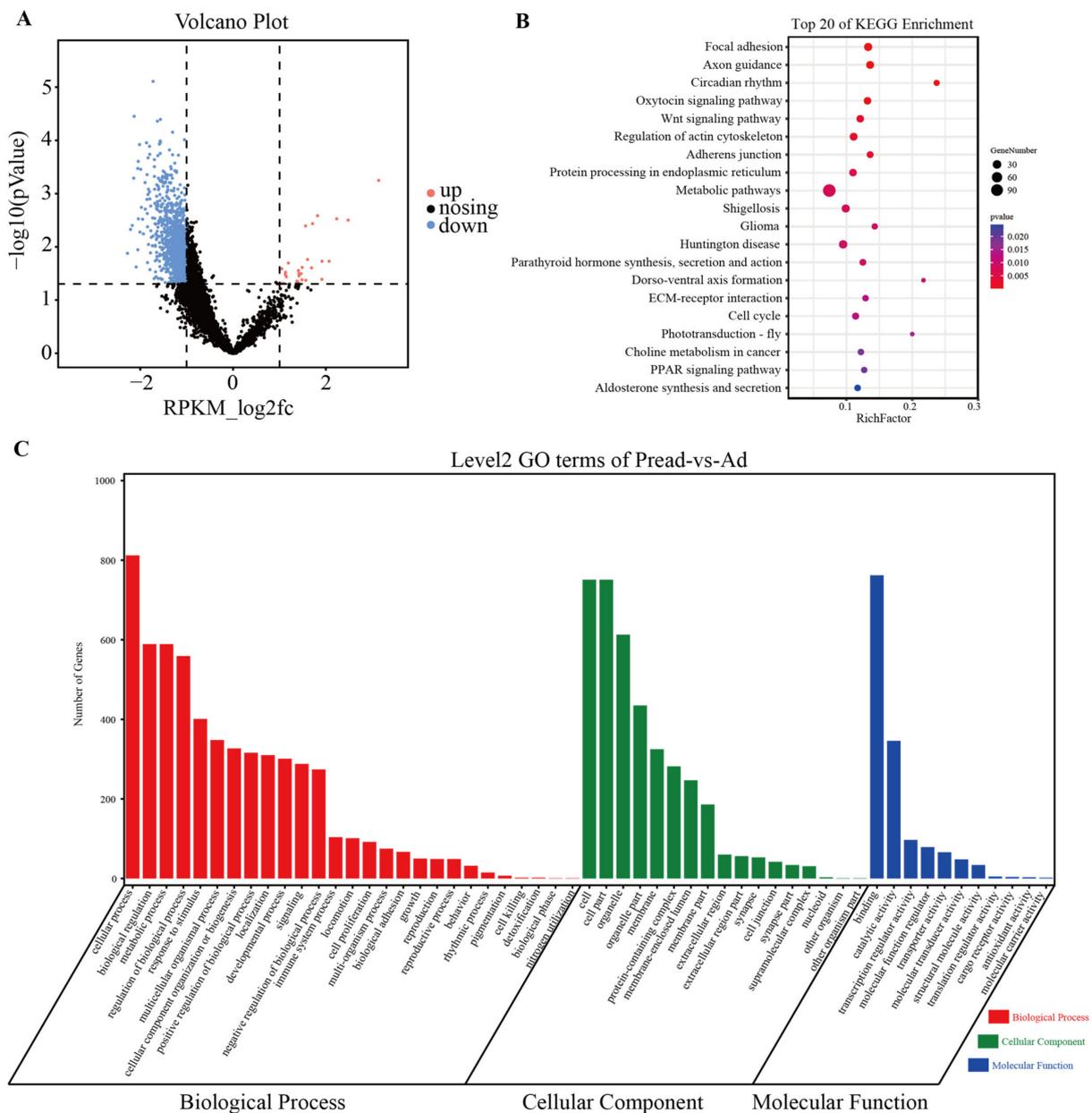

Figure 3. GO and KEGG pathway enrichment analysis of genes associated with differential chromatin accessibility. (**A**) Volcano plots showing differential ATAC peaks ($|\log2FC| > 1$, *p*-value < 0.05) between Pread and Ad groups. (**B**) The top 20 KEGG pathways for the genes of differential ATAC peaks. (**C**) GO functional annotation of differential chromatin accessibility between Pread and Ad groups.

2.3. Integration of ATAC-Seq and mRNA-Seq with Both Groups

mRNA-seq was used to detect gene expression changes during adipocyte differentiation. In total, 667 DEGs (|log2FC| > 1, FDR < 0.05) were identified, of which 381 genes were downregulated and 286 upregulated (Figure S3). The mRNA-seq results were validated by qRT-PCR (Figure S4). To explore the association between chromatin accessibility and gene expression in different cell types, we integrated ATAC-seq and mRNA-seq in preadipocytes and adipocytes, respectively. All genes were divided into three groups (high, medium and low) based on the gene activity value (RPKM), and the gene expression levels (FPKM) of each group were counted. The results showed that the gene expression level of the high group in Pread and Ad was higher than that of the medium and low groups (Figure 4A). This indicated that the more the open chromatin was present in preadipocytes and adipocytes, the higher the level of gene expression. Analogously, all genes were separated into three groups (high, medium and low) on the basis of the expression level, and the level of chromatin accessibility at different gene positions was counted. The results identified that the overall chromatin accessibility of the high group in Pread and Ad was higher than that of the medium and low groups. Interestingly, chromatin accessibility levels in the high and medium groups displayed obvious peaks near the TSS (Figure 4A). This suggested that the high level of chromatin accessibility located near the TSS contributes to the high gene expression level during adipocyte differentiation. To further investigate the importance of the region located near the TSS, we integrated chromatin accessibility within 1 kb upstream or downstream of the TSS and gene expression levels. We used the same method to classify chromatin accessibility and gene expression levels into three groups (high, medium and low) and analyzed their association in Pread and Ad. The results also revealed that within the 1 kb upstream and downstream of the TSS, the chromatin accessibility and gene expression levels of the high group were higher than those of the medium and low groups (Figure 4B). The high and medium groups had the highest levels of chromatin accessibility in promoter regions close to the TSS, demonstrating the vital role of these regions in gene transcriptional regulation (Figure 4B). To more intuitively show the regulation of chromatin accessibility on gene expression, we used IGV to simultaneously display the ATAC-seq and mRNA-seq signals of the fat metabolism-related genes CDK5 (cyclin-dependent kinase 5), HMGB2 (high mobility group box 2), and PIN1 (peptidylprolyl cis/trans isomerase, NIMA-Interacting 1). The ATAC-seq signal within promoter regions close to the TSS was obviously higher than that in the other gene regions (Figure 4C). The ATAC-seq signal levels of DEGs showed that the overall signal intensity of the Pread group was higher than that of the Ad group within the 2 kb upstream and downstream of the TSS (Figure S5A). A further analysis of DEGs found that 33 genes had differences in chromatin accessibility (Figure S5B). This indicated that differential gene expression caused by differential chromatin accessibility is a mode of gene expression regulation during adipocyte differentiation. The interaction of distal regulatory elements and proximal promoter regions is a crucial mechanism for gene regulation. Therefore, we detected the ATAC signals of the distal (>2 kb from the TSS) and proximal (TSS ± 2 kb) peaks and the gene expression level, and found that the overall signals of the distal and proximal peaks in the Pread group were higher than those in the Ad group. The gene expression also differed between the two groups (Figure S5C).

2.4. Prediction of TF Regulatory Networks during Yak Adipocyte Differentiation

TFs are a class of protein molecules that bind to promoters and cooperate with RNA polymerase II to initiate gene transcription. We used AME from the MEME suite (E-Value < 0.01) to identify the enriched TFs based on conservative motifs of differential peaks. A total of 26 TFs were enriched in these differential peaks, among which FOS (FBJ osteosarcoma oncogene), JUNB (jun B proto-cologene) and KLF5 (Kruppel-like factor 5) were confirmed to be related to the adipogenic metabolism regulation (Figure 5A and Table S7). The mechanism of action of TFs is usually combined with other TFs to work together. To identify the interaction between TFs, we constructed a global landscape of

the TF regulatory network during yak adipocyte differentiation. The most connected TFs were FOS, JUNB, and JUND (jun D proto-cologene), and they could cooperate with many other TFs in the regulatory network. The expression levels of FOSL2, TEAD1 (TEA domain transcription factor 1) and RUNX2 were significantly different (FDR < 0.01 and |log2FC| > 0.8) between Pread and Ad groups (Figure 5B and Table S8).
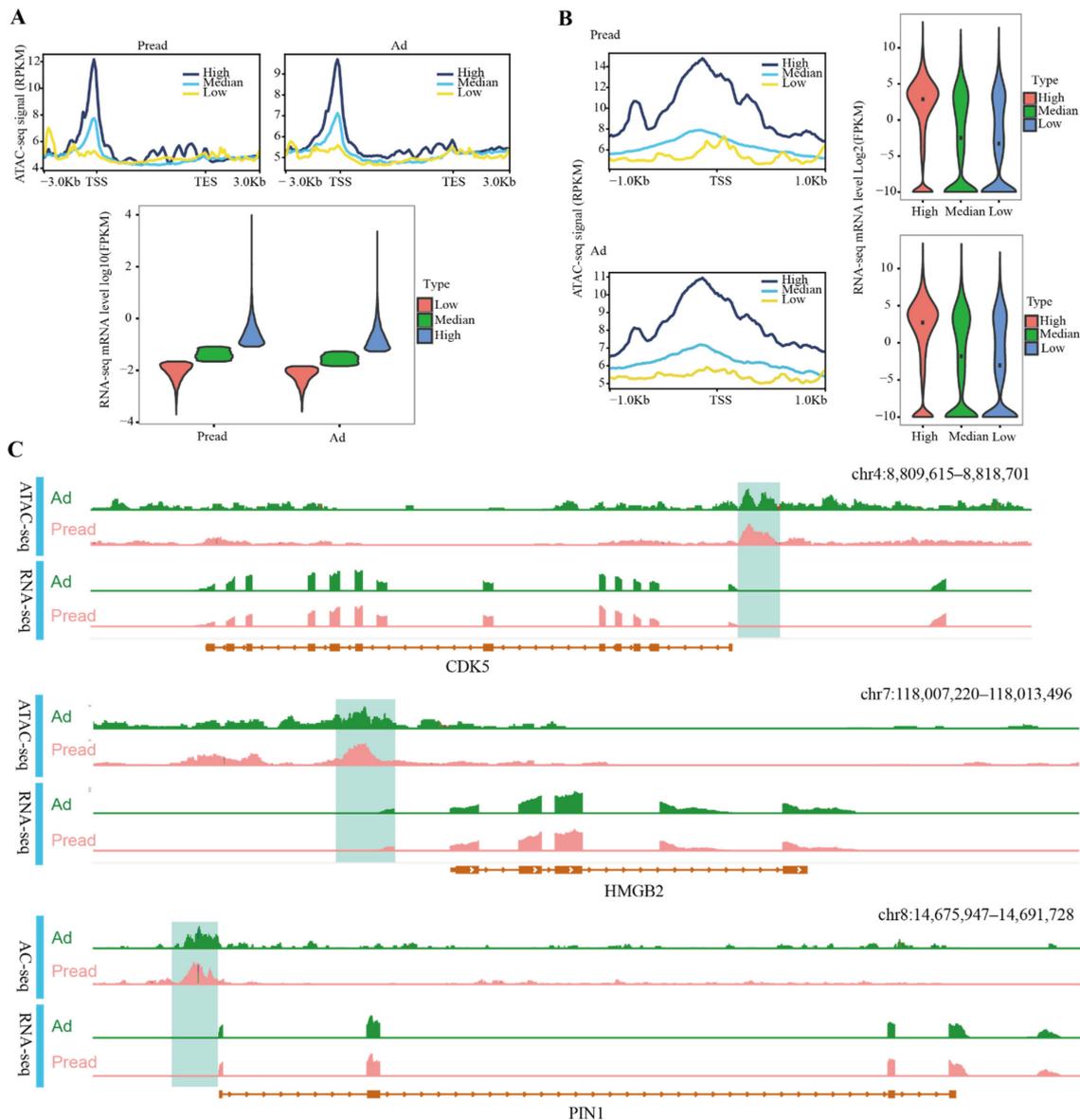

Figure 4. Association between ATAC signal and gene expression in Pread and Ad groups. (**A**) Association between ATAC signal within gene bodies and its 3kb upstream and downstream and gene expression. All genes were separated into three groups (high, medium, low) by expression level. Profile plots (above) display the difference in ATAC signal of each group. The horizontal axis represents different gene positions, and the vertical axis represents ATAC-seq signal intensity. Violin plot (below) show mRNA levels of high, medium and low groups divided by the gene activity value (RPKM). Plot width represents the density of repeated values in the range. (**B**) Association between ATAC signal within 1kb upstream and downstream of the TSS and gene expression. The same method is used to divide the high, medium and low groups according to RPKM and gene expression levels (FPKM). (**C**) The enrichment of ATAC-seq and RNA-seq signal near CDK5, HMGB2 and PIN1. Direction of the transcription is represented by an arrow.
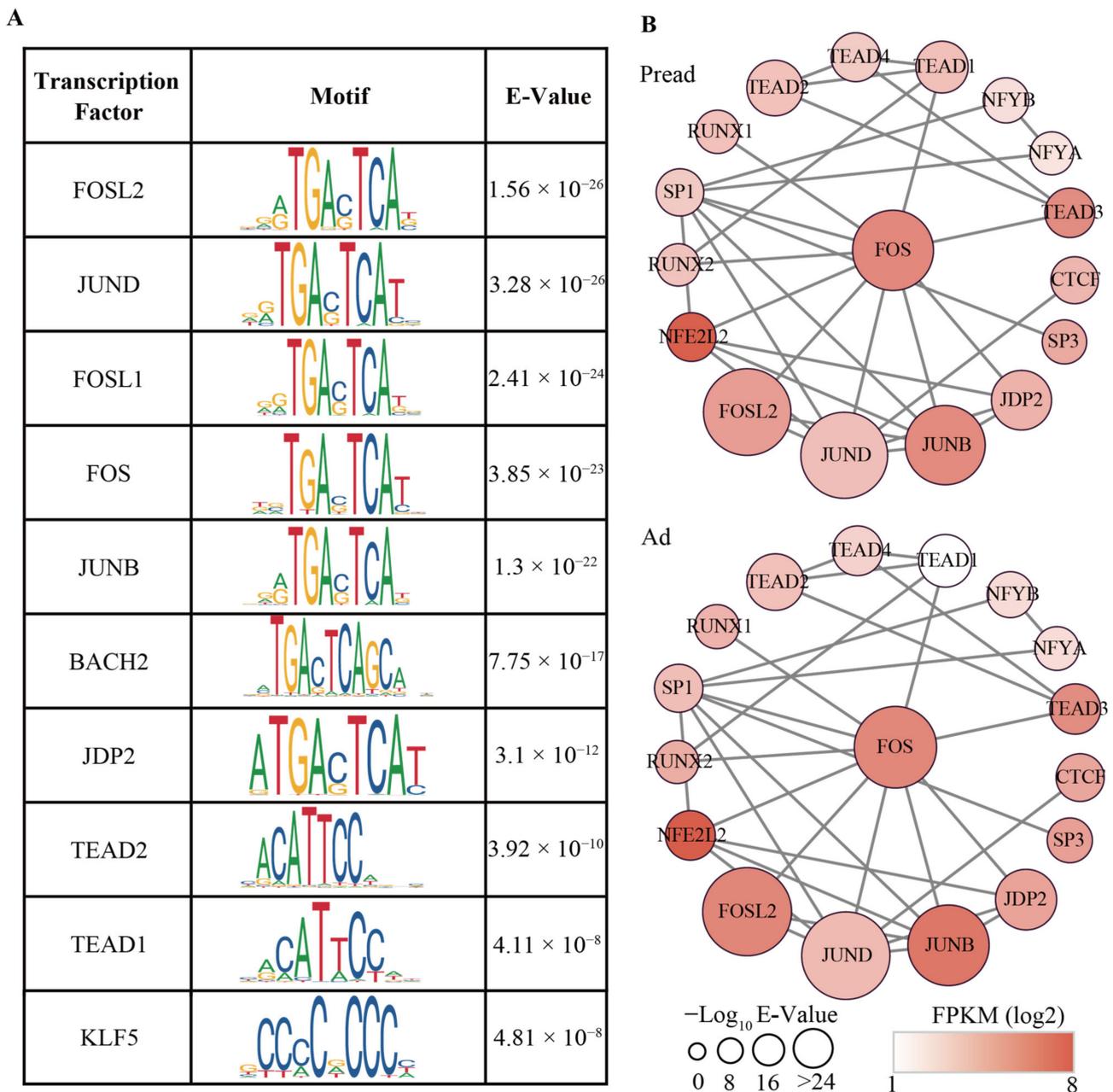
Figure 5. Transcriptional regulatory network during adipocyte differentiation. (**A**) Top ten transcription factors enriched in motifs of differential accessible peaks. The E-value is the adjusted *p*-value multiplied by the number of motifs in the motif file. (**B**) The interaction network between transcription factors. The size of the nodes represents the E-value of the significance. The color of the nodes represents expression levels (FPKM) of TFs.

3. Discussion

To our knowledge, this study is the first to use ATAC-seq technology to detect the landscape of chromatin accessibility during yak adipocyte differentiation. This will provide a new perspective on the molecular mechanism of adipogenesis in plateau species. Transcriptional regulatory elements of genes include enhancers, upstream activators, and proximal promoters. When the chromatin region containing transcriptional regulatory elements is in an open state, TFs can bind to the region and recruit RNA polymerase to the core promoter to initiate transcription. By contrast, condensed chromatin prevents TFs from binding to regulatory elements to silence gene expression [21]. ATAC-seq is a

whole genome epigenetic detection technology that utilizes Tn5 transposase to add the adapter into open chromatin regions located primarily in promoters. The most prominent advantages of ATAC-seq are low cell input and short time of sample preparation and experiment [22]. ATAC-seq is used to study chromatin accessibility regulation during embryonic development, cell differentiation, and disease occurrence [23]. Genome-wide open chromatin detected through ATAC-seq in mouse preimplantation embryos revealed that accessible chromatin regions were widely distributed within cis-regulatory sequences whose activities diminished prior to major zygotic genome activation. Further integration of cis-regulatory sequences with single-cell transcriptomes led to the identification of essential lineage-specific regulators [24]. Use of ATAC-seq to study bovine early embryo development revealed that chromatin accessibility significantly increased during major embryonic genome activation, and its signals were strong at the TSS and transcription end sites [25]. A study on bovine myogenic differentiation found that chromatin accessibility identified through ATAC-seq exhibited dynamic changes at different time points (0, 2, and 4 days) of differentiation, indicating that open chromatin was among the epigenetic regulatory effects regulating cell differentiation [26]. Yak adipogenesis is the differentiation of preadipocytes into adipocytes; therefore, ATAC-seq can be used to identify chromatin accessibility changes during this process. Our results also suggested that open chromatin regulates yak adipogenesis, and this finding can be helpful for fat metabolism research.

We noted that the number of Pread peaks (15,045) was much higher than that of Ad peaks (7683), suggesting that ATAC-seq worked better in preadipocytes, or that differentiation would lead to reduced chromatin accessibility in yak adipocytes. The peaks of Pread and Ad groups exhibited a similar genomic distribution, demonstrating the preferential location within the distal intergenic region, intron and promoter. This distribution is consistent with the chromatin accessibility features reported in bovine myogenic differentiation, suggesting the vital role of non-promoter cis-regulatory elements in the regulation of cell differentiation [26]. Although TF-binding sites (TFBS) are mainly located in the promoter region, an increasing number of studies have shown that intron regions are also rich in TFBS. A study of regulatory elements in the human coagulation factor VIII (hFVIII) gene found that 31% of TFBS were located in intron regions and had binding preferences located far from intronic splice sites [27]. Genome-wide detection of cis-regulatory elements in early human adipogenesis revealed chromatin accessibility changes in PPARγ intron regions rich in TFBS [28]. Consequently, we inferred that the gene regulation effect of TFs binding in the intron region was also a regulatory mechanism during yak adipogenesis. Enhancer elements with sequence conservation are known to be present in intergenic regions [29]. These enhancer DNA elements are commonly recognized and bound by specific TFs to ensure transcriptional activation of target genes [30]. Meanwhile, increasing evidence suggests that enhancer–promoter interactions are also crucial for gene expression regulation [31]. Studies have shown that the majority of chromatin accessibility found during induction of T cell activation by omni ATAC-seq were located in intergenic enhancer regions [32]. A large fraction of chromatin accessibility in our study was located in intergenic regions, further suggesting that enhancers may interact with TFs and promoters to regulate gene expression during yak adipocyte differentiation.

GO terms of differential chromatin accessibility were mainly implicated with binding, transcription regulator activity, molecular transducer activity, metabolic process, regulation of biological processes, developmental processes, signaling, cell part and protein-containing complex being enriched, further illustrating that adipocyte differentiation was a complex transcriptional regulatory activity mediated by a cascade of multiple TFs. The KEGG pathway enrichment analysis found that the signaling pathways of differential chromatin accessibility was significantly associated with adipogenesis, such as the PPAR signaling pathway, Wnt signaling pathway, regulation of actin cytoskeleton, and ECM–receptor interaction. Among them, the PPAR signaling pathway, Wnt signaling pathway, and ECM–receptor interaction are the classical pathways of fat metabolism. The actin cytoskeleton is a highly dynamic structure involved in the maintenance of cell morphology and structural

stability [33]. Insulin metabolism is related to actin cytoskeleton participation in the insulin signaling pathway [34]. Furthermore, investigation of the role of actin cytoskeleton in adipocyte development found that the interaction of the remodeled actin cytoskeleton and insulin signaling would affect adipocyte size [35].

Open chromatin regions can provide binding sites for TFs, thereby making initiation of target gene transcription possible. Therefore, integration of genome-wide chromatin accessibility and gene expression in specific biological processes can reveal the role that open chromatin plays in gene expression regulation. Chromatin accessibility has a significant correlation with gene expression during adipogenesis of human adipose-derived stem cells [28], bovine early embryo development [25], and bovine myogenic differentiation [26]. Our study results also revealed that genes expressed at high levels were associated with high levels of chromatin accessibility, especially within the 1 kb upstream and downstream of the TSS. Moreover, highly expressed genes exhibited the highest levels of chromatin accessibility in promoter regions close to the TSS. These results suggested that open chromatin can regulate gene expression during yak adipogenesis. We noted that only 33 DEGs identified by RNA-seq exhibited differences in chromatin accessibility, which could be explained as follows: (i) activation of target genes requires more enhancer elements in an open chromatin state; (ii) additional cofactors are required to co-ordinate TFs that bind to open chromatin regions to initiate gene transcription; and (iii) open chromatin regions may be affected by DNA methylation or other epigenetic modifications.

In the present study, differential chromatin accessibility was enriched in some crucial TFs, including FOSL2 (FOS-like antigen 2), JUND, FOS, and JUNB. FOSL2 and FOS are members of the FOS gene family. A study on the adipogenesis of deep intra-abdominal preadipocytes found that the induction of FOS protein was essential for preadipocyte differentiation into adipocytes [36]. Studies on osteoblasts from FOSL2 knockout mice in vitro found that the expression levels of adipogenic genes (C/EBPα, C/EBPβ, and PPARγ) were increased and adipocyte formation was accelerated [37]. Compared with the control mice, JUNB knockout mice exhibited reduced fat mass, higher insulin sensitivity, and increased adipose triglyceride lipase and hormone-sensitive lipase levels, demonstrating the critical regulatory role of JUNB in fat metabolism [38]. JUND can bind to the promoter of PPARγ, which contributes to the expression of FAS (fatty acid synthetase), CD36 (cluster of differentiation 36), LPL (lipoprotein lipase), and Plin5 (Perilipin 5) genes related to triglyceride synthesis, uptake, hydrolysis, and storage [39]. Based on the aforementioned findings, we infer that FOSL2, JUND, FOS, and JUNB may be involved in regulating yak adipogenesis. However, complex biological processes require multiple TFs to cooperate with each other. The members of FOS and JUN proteins can combine to form transcriptionally active complexes and exert regulatory functions during cell proliferation, differentiation, and embryonic development [40,41]. Studies have shown that the Fos–Jun complex can bind to the promoter region of the lipid-binding protein adipocyte P2 (aP2) to regulate 3T3-F442A adipocyte differentiation [42]. In our study, the core TF FOS interacted with JUNB and JUND in the TF regulatory network, further demonstrating that this interaction occurs during yak adipocyte differentiation. Interactions between other TFs were also observed in the regulatory network, and further investigation is required to clarify the mode of action of these TFs in adipogenesis.

4. Materials and Methods

4.1. Ethics Statement

All experimental procedures involved in this study were reviewed and confirmed by the Animal Administration and Ethics Committee of Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences (SYXK-2014–0002).

4.2. Preadipocyte Isolation

Three healthy 3-day-old infant Datong yaks from Datong Yak Breeding Center (Datong County, Qinghai, China) were selected as experimental animals for this study. The

collected subcutaneous adipose tissue samples were first rinsed with 0.9% NaCl and then with phosphate buffered saline (PBS) containing penicillin (200 U/mL) and streptomycin (200 U/mL). The adipose tissue was then cut into approximately 1 mm^3 pieces in a sterile environment. Cells were isolated through type I collagenase (1 mg/mL) digestion with constant stirring for 60 min in a 37 °C in water bath. Subsequently, the digested tissue was sequentially filtered through 100 μm and 70 μm nylon mesh films, and the filtrate was centrifuged at 1400× *g* for 5 min. Cell pellets obtained through centrifugation were incubated with red blood cell lysis buffer (Beyotime, Shanghai, China) for 10 min. After washing with PBS, the mixture was centrifuged at 1400× *g* for 5 min to obtain the final cell pellets. Finally, the pellets were dissolved with DME-F12 (Hyclone, Logan, UT, USA) containing 10% fetal bovine serum (Gibco, Waltham, MA, USA), inoculated into flasks, and cultured under conditions at 37 °C with 5% CO_2.

4.3. Staining of Oil Red O and Quantitative Real-Time PCR

When the density of preadipocytes reached 70–80%, the cells were induced using adipogenic agents containing 3-isobutyl-methylxanthine (0.5 mM), dexamethasone (1 μM), and insulin (10 μg/mL) (Sigma, St. Louis, MO, USA) for 2 days. The cells were then cultured with adipogenic agents including only insulin (10 μg/mL) until day 12. Finally, the adipocytes were washed three times with PBS, fixed with 4% formaldehyde crosslink for 1 h, and reacted with Oil Red O for 30 min. The stained samples were observed through light microscopy.

RNA extraction and reverse transcription of the cells were performed according to the manual of TriZol reagent (Transgen Biotch, Beijing, China) and PrimeScript™ 1st Strand cDNA Synthesis Kit (TaKaRa Bio Inc., Dalian, China), respectively. Quantitative RT-PCR was conducted on the LightCycler® 96 Instrument (Roche, Basel, Switzerland) with the SYBR Green dye. Primer sequences used to determine the credibility of RNA-seq are listed in Table S9.

4.4. ATAC-Seq Library Preparation and Sequencing

We performed ATAC-seq using preadipocytes (n = 3) and adipocytes (n = 3). Nuclear suspensions of samples were incubated in the transposition reaction mix containing Tn5 transposase at 37 °C for 30 min. When Tn5 entered the nuclei, open chromatin regions were preferentially fragmented. Meanwhile, an adapter was added to the ends of the fragments. Immediately following transposition, the sample was purified using the QIAGEN MinElute PCR purification kit (Tiangen Biotech, Beijing, China) [43]. The libraries were sequenced using Illumina HiSeqTM 4000 by Gene Denovo Biotechnology Co. (Guangzhou, China).

4.5. RNA-Seq Library Preparation and Sequencing

Total RNA of preadipocytes and adipocytes was isolated using the TriZol reagent kit in accordance with the manufacturer's manual. RNA quality was verified through RNase-free agarose gel electrophoresis and evaluated on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). Then, mRNA was enriched using oligo(dT) beads. Subsequently, the enriched mRNA was cut into short fragments using fragmentation buffer and reverse transcribed into cDNA by utilizing NEBNext Ultra RNA Library Prep Kit for Illumina (NEB #7530, New England Biolabs, Ipswich, MA, USA). The purified double-stranded cDNA fragments were end fixed, and a base was added. Then, the fragments were connected to Illumina sequencing adapters. The connection reaction was purified using AMPure XP beads (1.0×). Ligated fragments were subjected to size selection through polymerase chain reaction (PCR) amplification and agarose gel electrophoresis. The resulting cDNA library was sequenced using Illumina Novaseq6000 by Gene Denovo Biotechnology Co. (Guangzhou, China).

4.6. ATAC-Seq Data Processing and Analyses

To ensure the high quality of ATAC-seq reads, we conducted three stringent filtering standard procedures including removing reads with adapters, more than unknown nucleotides (N), and more than 50% of low-quality (Q-value $\leq$ 20) bases. Filtered reads of each sample were mapped to LU_Bosgru_v3.0 (Ensembl_release104) by using Bowtie2 [44] (version 2.2.8; parameters: -X 2000). Reads mapped to the + and − strands were offset by +4 bp and −5 bp, respectively. Peaks were called using MACS [45] (version 2.1.2, Yong Zhang, Boston, MA, USA) with parameters "--nomodel --shift -100 --extsize 200 -B -q 0.05". We used ChIPseeker [46] (version v1.16.1, Guangchuang Yu, Guangzhou, China) to confirm peak-related genes and distribution of peaks in different genome regions. Significant differential peaks were identified using DiffBind [47] (version 2.8.0, Rory Stark, Cambridge, UK) with $p < 0.05$ and $|\log2FC| > 1$. The MEME-ChIP and MEME-AME suite was used to identify the motifs. The ChIP and AME were selected to examine motifs with high credibility through peak regions and confirm the existences of any specific known motifs, respectively. Gene Ontology (GO) enrichment analysis was performed to recognize the main biological functions of peak-related genes. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed to identify metabolic pathways or signal transduction pathways associated with peak-related genes. GO terms and pathways were considered to be significantly enriched if they met the threshold of $p < 0.05$.

4.7. RNA-Seq Data Processing and Analyses

To obtain high-quality reads, raw reads containing adapters or low-quality bases were filtered using fastp [48] (version 0.18.0, Shifu Chen, Shenzhen, China). The filtered reads were first aligned to the ribosome RNA (rRNA) database by Bowtie2 [44] (version 2.2.8, Ben Langmead, Maryland, USA). After removing rRNA reads, all clean reads were aligned to LU_Bosgru_v3.0 (Ensembl_release104) using HISAT (version 2.2.4, Daehwan Kim, Maryland, USA) [49] with "-rna-strandness RF" and other parameters set as a default. Reads aligning to each sample were assembled using StringTie (version 1.3.1, Mihaela Pertea, Maryland, USA) [50]. Meanwhile, the FPKM (fragment per kilobase of transcript per million mapped reads) value was calculated using RSEM [51] software to quantify gene abundance. DEGs between Pread and Ad groups were identified using DESeq2 (version 1.14.1, Michael I Love, Heidelberg, Germany) [52]. DEGs with false discovery rate (FDR) of <0.05 and fold change of >2 were regarded as significant DEGs.

4.8. Integration of ATAC-Seq and RNA-Seq

The reads belonging to each of the ATAC-seq peaks were converted to RPKM (reads per kilobase per million mapped reads). We split all genes into three groups (high, medium and low) by using a threshold value determined by dividing the gene activity value (RPKM) in three quantile groups based on their means (Hmisc::cut2 R function [53]). Then, the gene expression levels (FPKM) in each group were calculated. We used the same method to divide mRNA values (FPKM) of all genes into high, medium, and low groups and calculate RPKM in different gene positions. Integrative Genomics Viewer (IGV) (version 2.12.2, James T. Robinson, MA, USA) was applied to visualize ATAC-seq and mRNA-seq signals at the same location. The TF regulatory networks were generated using the STRING database and visualized in Cytoscape (version 3.7.1, Paul Shannon, MA, USA).

4.9. Statistical Analysis

Student's *t*-test in SPSS (version 22, IBM, Chicago, IL, USA) was used to evaluate statistics. The results were displayed as mean ± SEM, and $p < 0.05$ was defined as statistically significant.

5. Conclusions

In summary, our study described genome-wide dynamic chromatin accessibility and gene expression during adipocyte differentiation in yaks, which are an ideal model for

studying the adipogenesis of plateau species. Integration of ATAC-seq and mRNA-seq revealed that genes expressed at high levels were associated with high levels of chromatin accessibility, especially within 1 kb upstream and downstream of the TSS. Additionally, we identified a series of TFs and created the TF regulatory network, clarifying the possible interactions between TFs during yak adipogenesis. Taken together, our study sheds light on the mechanism by which open chromatin regulates gene expression during adipogenesis and provides the theoretical and material bases for research on its epigenetic roles in fat metabolism.

Abbreviations

ATAC	seq assay for transposase-accessible chromatin with high-throughput sequencing
DEGs	differentially expressed genes
FDR	false discovery rate
FPKM	fragment per kilobase of transcript per million mapped reads
RPKM	reads per kilobase per million mapped reads
qRT	PCR real-time quantitative PCR
PCR	polymerase chain reaction

References

1. Tang, Q.Q.; Lane, M.D. Adipogenesis: From stem cell to adipocyte. *Annu. Rev. Biochem.* **2012**, *81*, 715–736. [CrossRef]
2. Gregoire, F.M.; Smas, C.M.; Sul, H.S. Understanding Adipocyte Differentiation. *Physiol. Rev.* **1998**, *78*, 783–809. [CrossRef] [PubMed]
3. Ghaben, A.L.; Scherer, P.E. Adipogenesis and metabolic health. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 242–258. [CrossRef] [PubMed]
4. Haeusler, R.A.; McGraw, T.E.; Accili, D. Biochemical and cellular properties of insulin receptor signalling. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 31–44. [CrossRef] [PubMed]
5. Tomlinson, J.J.; Boudreau, A.; Wu, D.; Abdou Salem, H.; Carrigan, A.; Gagnon, A.; Mears, A.J.; Sorisky, A.; Atlas, E.; Hache, R.J. Insulin sensitization of human preadipocytes through glucocorticoid hormone induction of forkhead transcription factors. *Mol. Endocrinol.* **2010**, *24*, 104–113. [CrossRef]

6. Son, J.W.; Kim, M.K.; Park, Y.M.; Baek, K.H.; Yoo, S.J.; Song, K.H.; Son, H.S.; Yoon, K.H.; Lee, W.C.; Cha, B.Y.; et al. Association of serum bone morphogenetic protein 4 levels with obesity and metabolic syndrome in non-diabetic individuals. *Endocr. J.* **2011**, *58*, 39–46. [CrossRef]

7. Ross, S.E.; Hemati, N.; Longo, K.A.; Bennett, C.N.; Lucas, P.C.; Erickson, R.L.; MacDougald, O.A. Inhibition of adipogenesis by Wnt signaling. *Science* **2000**, *289*, 950–953. [CrossRef]

8. Suh, J.M.; Gao, X.; McKay, J.; McKay, R.; Salo, Z.; Graff, J.M. Hedgehog signaling plays a conserved role in inhibiting fat formation. *Cell Metab.* **2006**, *3*, 25–34. [CrossRef]

9. Hong, J.; Li, S.; Wang, X.; Mei, C.; Zan, L. Study of expression analysis of SIRT4 and the coordinate regulation of bovine adipocyte differentiation by SIRT4 and its transcription factors. *Biosci. Rep.* **2018**, *38*, BSR20181705. [CrossRef] [PubMed]

10. Shijun, L.; Khan, R.; Raza, S.H.A.; Jieyun, H.; Chugang, M.; Kaster, N.; Gong, C.; Chunping, Z.; Schreurs, N.M.; Linsen, Z. Function and characterization of the promoter region of perilipin 1 (PLIN1): Roles of E2F1, PLAG1, C/EBPbeta, and SMAD3 in bovine adipocytes. *Genomics* **2020**, *112*, 2400–2409. [CrossRef]

11. Guo, H.; Khan, R.; Raza, S.H.A.; Ning, Y.; Wei, D.; Wu, S.; Hosseini, S.M.; Ullah, I.; Garcia, M.D.; Zan, L. KLF15 promotes transcription of KLF3 gene in bovine adipocytes. *Gene* **2018**, *659*, 77–83. [CrossRef]

12. Qu, K.; Zaba, L.C.; Giresi, P.G.; Li, R.; Longmire, M.; Kim, Y.H.; Greenleaf, W.J.; Chang, H.Y. Individuality and variation of personal regulomes in primary human T cells. *Cell Syst.* **2015**, *1*, 51–61. [CrossRef] [PubMed]

13. Li, Z.; Schulz, M.H.; Look, T.; Begemann, M.; Zenke, M.; Costa, I.G. Identification of transcription factor binding sites using ATAC-seq. *Genome Biol.* **2019**, *20*, 45. [CrossRef] [PubMed]

14. Buenrostro, J.D.; Wu, B.; Chang, H.Y.; Greenleaf, W.J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr. Protoc. Mol. Biol.* **2015**, *109*, 21–29. [CrossRef]

15. Yan, F.; Powell, D.R.; Curtis, D.J.; Wong, N.C. From reads to insight: A hitchhiker's guide to ATAC-seq data analysis. *Genome Biol.* **2020**, *21*, 22. [CrossRef]

16. Liu, Y.; Luo, J.; Dou, J.; Yan, B.; Ren, Q.; Tang, B.; Wang, K.; Qiu, Q. The sequence and de novo assembly of the wild yak genome. *Sci. Data* **2020**, *7*, 66. [CrossRef]

17. Wang, H.; Zhong, J.; Zhang, C.; Chai, Z.; Cao, H.; Wang, J.; Zhu, J.; Wang, J.; Ji, Q. The whole-transcriptome landscape of muscle and adipose tissues reveals the ceRNA regulation network related to intramuscular fat deposition in yak. *BMC Genom.* **2020**, *21*, 347. [CrossRef]

18. Shikui, D.; Ruijun, L.; Muyi, K.; Xiaogeng, P.; Yanjun, G. Effect of urea multinutritional molasses block supplementation on liveweight change of yak calves and productive and reproductive performances of yak cows. *Can. J. Anim. Sci.* **2003**, *83*, 141–145. [CrossRef]

19. Dong, Q.M.; Zhao, X.Q.; Ma, Y.S.; Xu, S.X.; Li, Q.Y. Live-weight gain, apparent digestibility, and economic benefits of yaks fed different diets during winter on the Tibetan plateau. *Livest. Sci.* **2006**, *101*, 199–207. [CrossRef]

20. Ding, X.Z.; Guo, X.; Yan, P.; Liang, C.N.; Bao, P.J.; Chu, M. Seasonal and nutrients intake regulation of lipoprotein lipase (LPL) activity in grazing yak (*Bos grunniens*) in the Alpine Regions around Qinghai Lake. *Livest. Sci.* **2012**, *143*, 29–34. [CrossRef]

21. Gottesfeld, J.M.; Carey, M.F. Introduction to the Thematic Minireview Series: Chromatin and transcription. *J. Biol. Chem.* **2018**, *293*, 13775–13777. [CrossRef] [PubMed]

22. Sun, Y.; Miao, N.; Sun, T. Detect accessible chromatin using ATAC-sequencing, from principle to applications. *Hereditas* **2019**, *156*, 29. [CrossRef] [PubMed]

23. Corces, M.R.; Granja, J.M.; Shams, S.; Louie, B.H.; Seoane, J.A.; Zhou, W.; Silva, T.C.; Groeneveld, C.; Wong, C.K.; Cho, S.W.; et al. The chromatin accessibility landscape of primary human cancers. *Science* **2018**, *362*, eaav1898. [CrossRef] [PubMed]

24. Wu, J.; Huang, B.; Chen, H.; Yin, Q.; Liu, Y.; Xiang, Y.; Zhang, B.; Liu, B.; Wang, Q.; Xia, W.; et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **2016**, *534*, 652–657. [CrossRef]

25. Ming, H.; Sun, J.; Pasquariello, R.; Gatenby, L.; Herrick, J.R.; Yuan, Y.; Pinto, C.R.; Bondioli, K.R.; Krisher, R.L.; Jiang, Z. The landscape of accessible chromatin in bovine oocytes and early embryos. *Epigenetics* **2021**, *16*, 300–312. [CrossRef]

26. Li, Q.; Wang, Y.; Hu, X.; Zhang, Y.; Li, H.; Zhang, Q.; Cai, W.; Wang, Z.; Zhu, B.; Xu, L.; et al. Transcriptional states and chromatin accessibility during bovine myoblasts proliferation and myogenic differentiation. *Cell Prolif.* **2022**, *55*, e13219. [CrossRef]

27. Haddad, A.; Hemmat, J.; Aslamkhan, M. Intronic regions of the human coagulation factor VIII gene harboring transcription factor binding sites with a strong bias towards the short-interspersed elements. *Heliyon* **2020**, *6*, e04727. [CrossRef]

28. Li, S.; Zong, X.; Zhang, L.; Li, L.; Wu, J. A chromatin accessibility landscape during early adipogenesis of human adipose-derived stem cells. *Adipocyte* **2022**, *11*, 239–249. [CrossRef]

29. Fazel, S.; Esau, C.; Lesage, C.; Monis, S.; Poitras, L.; Yu, M.; Perin, S.; Hatch, G.; Ekker, M. Increased Sociability in Mice Lacking Intergenic Dlx Enhancers. *Front. Neurosci.* **2021**, *15*, 718948. [CrossRef]

30. Herz, H.M.; Hu, D.; Shilatifard, A. Enhancer malfunction in cancer. *Mol. Cell.* **2014**, *53*, 859–866. [CrossRef]

31. Mumbach, M.R.; Satpathy, A.T.; Boyle, E.A.; Dai, C.; Gowen, B.G.; Cho, S.W.; Nguyen, M.L.; Rubin, A.J.; Granja, J.M.; Kazane, K.R.; et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat. Genet.* **2017**, *49*, 1602–1612. [CrossRef]

32. Iqbal, M.M.; Serralha, M.; Kaur, P.; Martino, D. Mapping the landscape of chromatin dynamics during naive CD4+ T-cell activation. *Sci. Rep.* **2021**, *11*, 14101. [CrossRef]

33. Rottner, K.; Faix, J.; Bogdan, S.; Linder, S.; Kerkhoff, E. Actin assembly mechanisms at a glance. *J. Cell Sci.* **2017**, *130*, 3427–3435. [CrossRef] [PubMed]

34. Tsakiridis, T.; Tong, P.; Matthews, B.; Tsiani, E.; Bilan, P.J.; Klip, A.; Downey, G.P. Role of the actin cytoskeleton in insulin action. *Microsc. Res. Techniq.* **1999**, *47*, 79–92. [CrossRef]

35. Hansson, B.; Moren, B.; Fryklund, C.; Vliex, L.; Wasserstrom, S.; Albinsson, S.; Berger, K.; Stenkula, K.G. Adipose cell size changes are associated with a drastic actin remodeling. *Sci. Rep.* **2019**, *9*, 12941. [CrossRef]

36. Lacasa, D.; Garcia, E.; Henriot, D.; Agli, B.; Giudicelli, Y. Site-Related Specificities of the Control by Androgenic Status of Adipogenesis and Mitogen-Activated Protein Kinase Cascade/c-fos Signaling Pathways in Rat Preadipocytes. *Endocrinology* **1997**, *138*, 3181–3186. [CrossRef]

37. Bozec, A.; Bakiri, L.; Jimenez, M.; Schinke, T.; Amling, M.; Wagner, E.F. Fra-2/AP-1 controls bone formation by regulating osteoblast differentiation and collagen production. *J. Cell Biol.* **2010**, *190*, 1093–1106. [CrossRef]

38. Pinent, M.; Prokesch, A.; Hackl, H.; Voshol, P.J.; Klatzer, A.; Walenta, E.; Panzenboeck, U.; Kenner, L.; Trajanoski, Z.; Hoefler, G.; et al. Adipose triglyceride lipase and hormone-sensitive lipase are involved in fat loss in JunB-deficient mice. *Endocrinology* **2011**, *152*, 2678–2689. [CrossRef]

39. Costantino, S.; Akhmedov, A.; Melina, G.; Mohammed, S.A.; Othman, A.; Ambrosini, S.; Wijnen, W.J.; Sada, L.; Ciavarella, G.M.; Liberale, L.; et al. Obesity-induced activation of JunD promotes myocardial lipid accumulation and metabolic cardiomyopathy. *Eur. Heart J.* **2019**, *40*, 997–1008. [CrossRef]

40. Wagner, E.F. Functions of AP1 (Fos/Jun) in bone development. *Ann. Rheum. Dis.* **2002**, *61*, ii40–ii42. [CrossRef]

41. Chinenov, Y.; Kerppola, T.K. Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. *Oncogene* **2001**, *20*, 2438–2452. [CrossRef] [PubMed]

42. Herrera, R.; Ro, H.S.; Robinson, G.S.; Xanthopoulos, K.G.; Spiegelman, B.M. A direct role for C/EBP and the AP-I-binding site in gene expression linked to adipocyte differentiation. *Mol. Cell Biol.* **1989**, *9*, 5331–5339. [CrossRef] [PubMed]

43. Buenrostro, J.D.; Giresi, P.G.; Zaba, L.C.; Chang, H.Y.; Greenleaf, W.J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **2013**, *10*, 1213–1218. [CrossRef] [PubMed]

44. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef]

45. Zhang, Y.; Liu, T.; Meyer, C.A.; Eeckhoute, J.; Johnson, D.S.; Bernstein, B.E.; Nusbaum, C.; Myers, R.M.; Brown, M.; Li, W.; et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **2008**, *9*, R137. [CrossRef]

46. Yu, G.; Wang, L.G.; He, Q.Y. ChIPseeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **2015**, *31*, 2382–2383. [CrossRef]

47. Stark, R.; Brown, G. DiffBind Differential Binding Analysis of ChIP-Seq Peak Data. R Package Version. 2011, Volume 100. Available online: https://bioc.ism.ac.jp/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf (accessed on 18 May 2021).

48. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef]

49. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods.* **2015**, *12*, 357–360. [CrossRef]

50. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [CrossRef]

51. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [CrossRef]

52. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]

53. Ruiz, J.L.; Ranford, L.C.; Gomez, E. The regulatory genome of the malaria vector Anopheles gambiae: Integrating chromatin accessibility and gene expression. *NAR Genom. Bioinform.* **2021**, *3*, lqaa113. [CrossRef]