scientific reports

OPEN

Expression patterns of housekeeping genes and tissue-specific genes in black goats across multiple tissues

Chaobin Qin[1,6], Dong Wang[1,6], Hongbing Han[4], Yanhong Cao[2], Xiaobo Wang[5], Zeyi Xuan[2], Mingsong Wei[2✉], Zhipeng Li[1✉] & Qingyou Liu[3✉]

Black goats are a significant meat breed in southern China. To investigate the expression patterns and biological functions of genes in various tissues of black goats, we analyzed housekeeping genes (HKGs), tissue-specific genes (TSGs), and hub genes (HUBGs) across 23 tissues. Additionally, we analyzed HKGs in 13 tissues under different feeding conditions. We identified 2968 HKGs, including six important ones. Interestingly, HKGs in grazing black goats demonstrated higher and more stable expression levels. We discovered a total of 9912 TSGs, including 134 newly identified ones. The number of TSGs for mRNA and lncRNA were nearly equal, with 127 mRNA TSGs expressed solely in one tissue. Additionally, the predicted functions of tissue-specific long non-coding RNAs (lncRNAs) targeting mRNAs corresponded with the physiological functions of the tissues.Weighted gene co-expression network analysis (WGCNA) constructed 30 modules, where the dark grey module consists almost entirely of HKGs, and TSGs are located in modules most correlated with their respective tissues. Additionally, we identified 289 HUBGs, which are involved in regulating the physiological functions of their respective tissues. Overall, these identified HKGs, TSGs, and HUBGs provide a foundation for studying the molecular mechanisms affecting the genetic and biological processes of complex traits in black goats.

The goat, as one of the earliest domesticated animals with a history spanning over ten thousand years, is an integral part of human agricultural civilization[1]. Among the five major livestock species (goats, sheep, cattle, pigs, and chickens)[2,3], goats play an indispensable role in both commercial and subsistence farming systems, particularly in tropical, subtropical, and Mediterranean regions[4]. Goats provide high-quality milk[5], skin[6], wool[7], and red meat[8], earning them the title of "the poor man's cow"[9]. In China, goats have a long history of domestication, with the largest population of goats in the world. Due to the mountainous terrain in southern China, which is suitable for goat farming, they are extensively raised, contributing significantly to the economy. With the rapid development of RNA-seq technology in recent years, a large amount of transcriptome data from various animal tissues and developmental stages has been generated, providing new insights into the physiologicalfunctions and genetic traits of various animals[10]. Studying the transcriptomic profiles of different goat tissues contributes to a better understanding of goat breeding, growth, and development[11,12].

Since the introduction of RNA-seq, researchers have extensively sequenced numerous tissues from various mammals, yielding vast sequencing data, and explored them through large-scale gene expression profiling. For example, Poklukar et al. uncovered expression disparities among entire males (EM), immunocastrates (IC) and surgical castrates (SC) pigs utilizing RNA-seq and candidate gene expression methodologies[13]. La et al.

[1]School of Animal Science and Technology, Guangxi University, Nanning 530004, China. [2]Guangxi Vocational University of Agriculture, Nanning 530007, Guangxi, China. [3]Guangdong Provincial Key Laboratory of Animal Molecular Design and Precise Breeding, School of Life Science and Engineering, Foshan University, Foshan 528225, China. [4]Beijing Key Laboratory of Animal Genetic Improvement, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China. [5]Henan Academy of Crops Molecular Breeding/The Shennong Laboratory, Zhengzhou 450099, Henan, China. [6]These authors contributed equally on this work. ✉email: 175159461@qq.com; zp.li@gxu.edu.cn; qyliu-gene@gxu.edu.cn

detected significant variances in piRNA expression profiles across different developmental stages of yak testes[14]. Zhang et al. showcased that dietary probiotic supplements could enhance muscle development and the ultimate meat quality in Sunit lambs by modifying the expression profiles of genes associated with pivotal pathways[15]. Consequently, comprehending the expression patterns of genes across different tissues aids in unraveling their varied biological functions. Within goat breeding initiatives, enhancing reproductive and productive traits stands as one of the utmost critical tasks. Researchers frequently delve into the genetic mechanisms governing these distinct traits and tissues through transcriptomic profiling of various tissues, including ovaries[16], testes[17], hypothalamus[18], muscles[19], and others.

The concept of HKGs was first introduced by Levy et al. in 1977[20]. This concept is loosely defined, generally describing genes that exhibit relatively stable expression levels across all tissues and organs, playing essential roles in maintaining basic cellular physiological functions[21,22]. Therefore, HKGs such as *GAPDH* and *ACTB* are frequently utilized as reference genes in research[23]. Numerous researchers have employed various methods to analyze large volumes of RNA-seq data in order to identify new HKGs. For instance, Pan et al. conducted a large-scale analysis of RNA-seq data from 14 pig tissues and successfully identified 2351 HKGs[24]. Similarly, Joshi et al. delved deeply into HKGs from the perspectives of stability and conservation using single-cell transcriptomics[25]. Despite the traditional notion that HKGs maintain stable expression across all tissues, mounting evidence suggests that many commonly used HKGs exhibit variable expression levels across different environmental conditions, such as various metabolic processes[26], developmental stages[27], and tissue types[28], rendering them unsuitable as reliable reference genes. In contrast to HKGs, TSGs are typically highly expressed in a single specific tissue or organ and are either not expressed or expressed at low levels in other tissues or organs[29]. Because TSGs have unique biological functions in specific tissues, an increasing number of researchers are using RNA-seq to study them in depth. Ryaboshapkina and colleagues thoroughly discussed TSGs as drug targets, noting that the potential of TSGs in drug research and development has not been fully tapped[30]. Sonawane and others found that transcription factor connections to target genes exhibit higher tissue specificity compared to the genes themselves, and transcription factors are less likely to be expressed in a tissue-specific manner compared to their genes[31]. Moreover, Zhang and others found that in mammals, TSGs evolve on average more rapidly than HKGs[32]. LncRNAs are non-coding RNAs consisting of more than 200 nucleotides, most of whose functions have not been studied, but increasing evidence suggests that tissue-specific lncRNA genes play significant roles. For instance, Zhang discovered that mutations in the promoter region of a fat tissue-specific lncRNA gene (*LOC100847835*) can cause changes in the ketotic traits of cattle[33]. Rogala found that a heart tissue-specific lncRNA Sweetheart regulates compensatory myocardial hypertrophy following myocardial injury in mice[34]. WGCNA is a method for analyzing gene expression patterns across multiple samples, used to describe patterns of correlation among genes across microarray samples[35]. As WGCNA is further developed and applied, it is increasingly used in the study of large-scale RNA-seq data to explore relationships between various tissues and genes, identify core modules and HUBGs in various tissues, such as in human various cancer tissues[36,37], human brain extracellular vesicles[38], pig intramuscular fat[39], beef cattle muscle and fat[40], and sheep ovarian tissues[41].

This study utilized RNA-seq technology to analyze the transcriptomes of 23 goat tissues (rumen, reticulum, omasum, abomasum, duodenum, jejunum, ileum, cecum, colon, liver, longissimus dorsi muscle, biceps femoris, tongue, cerebellum, heart, spleen, kidney, lung, eye, medulla oblongata, hypothalamus, fat, and skin) to explore the associations and differences in gene expression patterns among different tissues. Our results will provide new insights into the biological processes of various goat tissues.
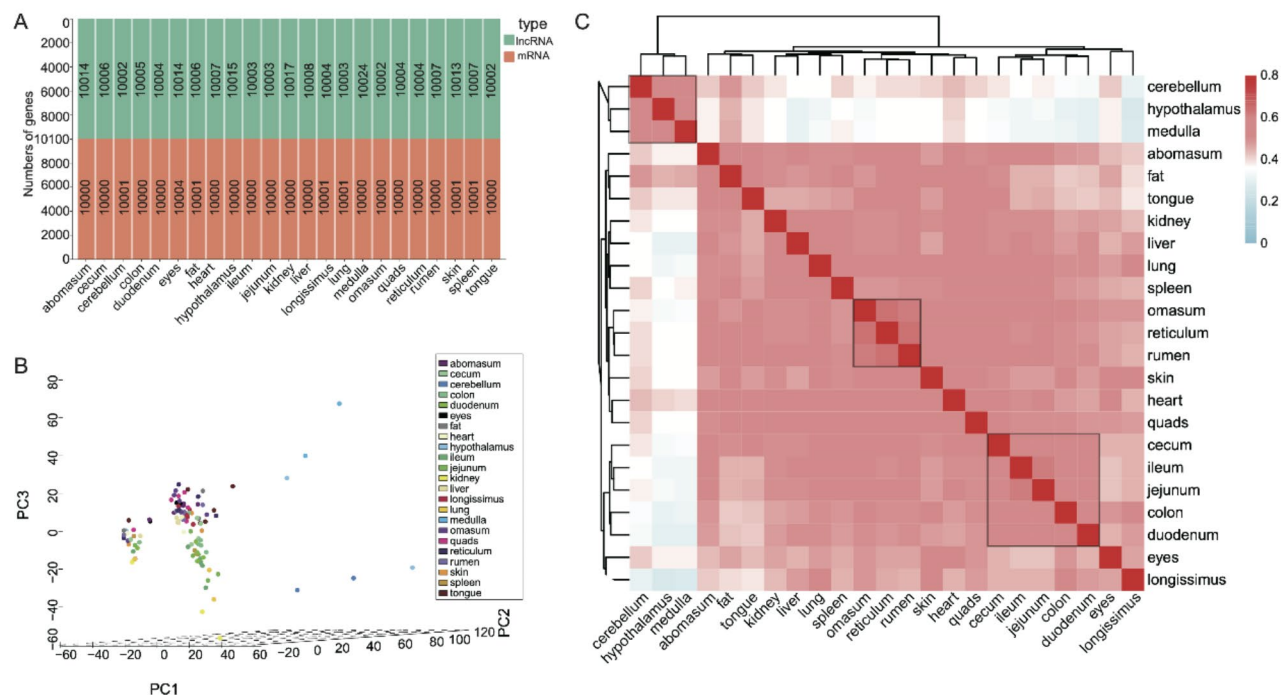
Results

Detection of genes in multiple tissues of black goats

We have gathered 108 RNA-seq samples, encompassing 23 different tissues from Black Goats, including the rumen, reticulum, omasum, abomasum, duodenum, jejunum, ileum, cecum, colon, liver, longissimus dorsi muscle, biceps femoris, tongue, cerebellum, heart, spleen, kidney, lung, eye, medulla oblongata, hypothalamus, fat, and skin (Table S1). In an effort to delve into the diversity and biological clustering of these various tissues, we conducted an analysis of the transcriptome data. Our findings revealed a total of 17,120 genes, with 16,885 of these identified in at least one sample with a TPM ≥ 1 (Table S2). Interestingly, the quantity of mRNAs and lncRNAs detected was nearly identical across all 23 tissues (Fig. 1A). Furthermore, our unbiased hierarchical clustering heat maps of the goat tissues demonstrated a tendency for tissues with similar physiological functions to cluster together. For instance, tissues from the nervous system such as the medulla oblongata, cerebellum, and hypothalamus were grouped together, as were the digestive system tissues like the rumen, reticulum, and omasum. Similarly, the duodenum, jejunum, ileum, colon, and cecum also formed a distinct cluster (Fig. 1C).

Analysis of HKGs expression in black goats

To investigate HKGs within black goats, we initially identified 11,872 HKGs through screening (Table S3). Typically, the expression patterns of HKGs are relatively constant across different tissues. However, we observed significant variation in the expression levels of these initially screened 11,872 HKGs across tissues (Fig. 2A). Based on the stability and expression levels of genes across various tissues, using coefficient of variation (CV) values, we further categorized the HKGs into three groups: low, medium, and high expression variability. The threshold values used for categorization were 1.005738828 (first quartile) and 1.818340663 (third quartile). We identified 2,968 genes with CV values below the first quartile as low expression variability HKGs (Fig. 2B, Table S4). Based on the heatmap generated from the top 100 housekeeping genes with the lowest CV (Figure S1), six important HKGs (*EEF1A1*, *RPS6*, *RPL7*, *RPL23*, *RPS15*, and *RPL19*) were ultimately identified. Compared to commonly used reference genes in black goats (such as *ACTB*, *UBC*, *BCBP3*, and *SDHA*), these six genes demonstrated higher and more stable expression levels (Fig. 2C).

Fig. 1. Gene expression profile across 23 different tissues in Black Goats. (**A**) The quantities of mRNA and lncRNA in various tissues. (**B**) principal component analysis (PCA) performed on all tissues based on the expression of genes detected in at least one sample with TPM ≥ 1. (**C**) Heatmap of unbiased hierarchical clustering based on Pearson Correlation Coefficients. Red indicates high correlation, while blue represents low correlation.
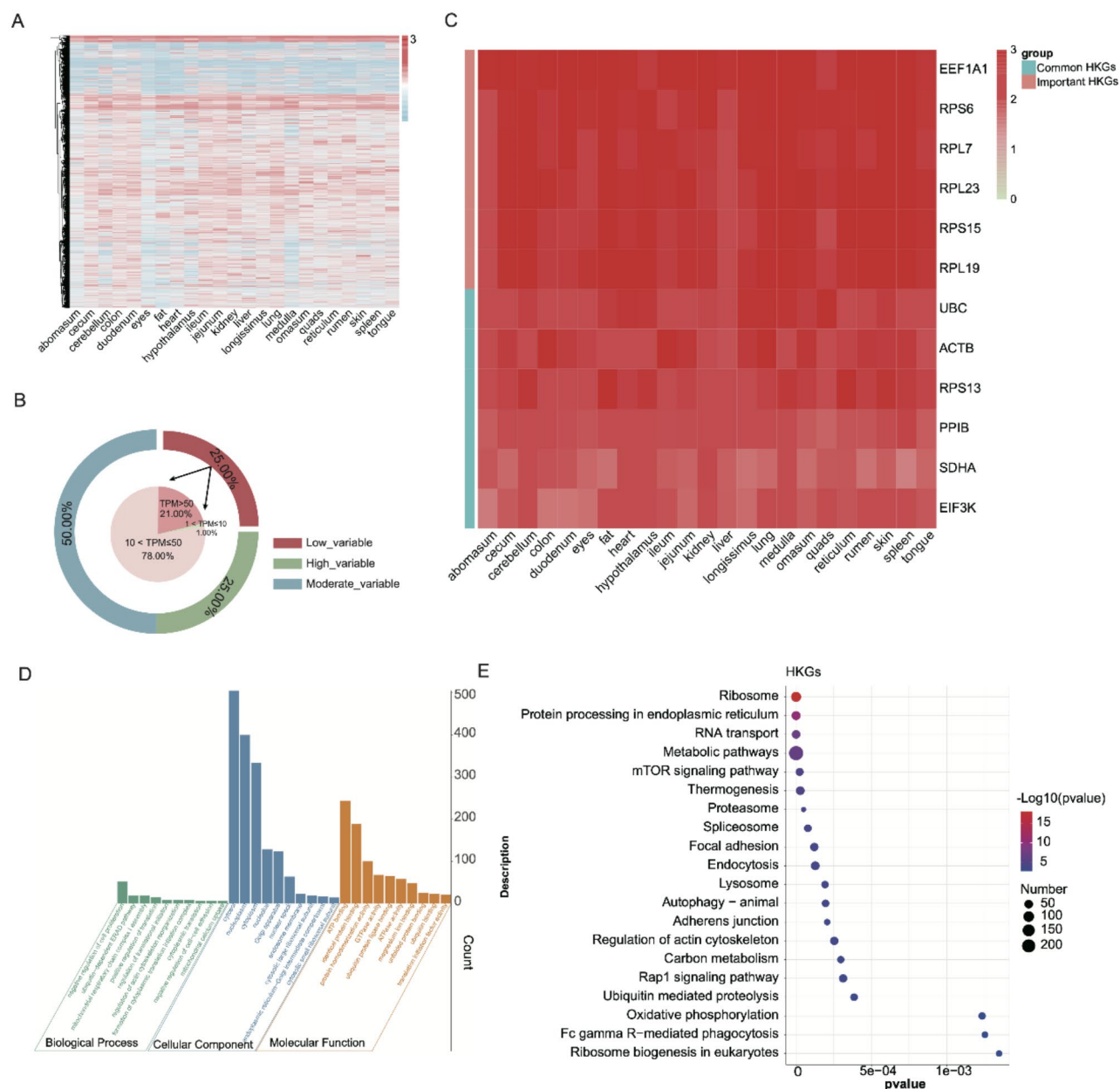
To gain deeper insights into the biological functions of these HKGs, we found that ubiquitin-dependent ERAD pathway, assembly of mitochondrial respiratory chain complex I, regulation of translation initiation, cytoplasmic translation, cytosolic ribosome, Golgi apparatus, ubiquitin-protein ligase binding, identical protein binding, and ATP binding were enriched in the GO analysis (Fig. 2D). Additionally, the KEGG analysis showed that ribosome, protein processing in endoplasmic reticulum, RNA transport, metabolic pathways, and mTOR signaling pathway were all involved in the regulation of HKGs (Fig. 2E).

To investigate the differences in HKGs between stall-fed and grazing black goats, we found that both groups have nearly the same number of HKGs (2939 and 2919 respectively), and they share 1002 common genes (Fig. 3A, Table S5, Table S6). Interestingly, the gene expression of stall-fed black goats is more concentrated, whereas grazing black goats display more outliers (Fig. 3B). Heatmap analysis of the shared and unique HKGs between stall-fed and grazing black goats reveals higher and more stable expression levels in grazing black goats (Fig. 3C, Figure S2). GO analysis indicates that the biological functions of HKGs in grazing black goats are mainly associated with the tricarboxylic acid cycle, regulation of cell shape, cytosolic ribosome, nuclear matrix, ER-Golgi intermediate compartment, ubiquitin-protein ligase binding, ATP binding, and ubiquitin binding (Fig. 3D). In contrast, the biological functions in stall-fed black goats are predominantly linked to tyrosine phosphorylation of peptides, regulation of cell migration, cytosolic ribosome, nuclear matrix, nucleolus, ATP binding, and protein serine/threonine/tyrosine kinase activity (Fig. 3E Figure S3).

Expression analysis of TSGs in black goats

To investigate TSGs in each tissue of black goats, a total of 9912 TSGs were identified after rigorous screening, including 4957 tissue-specific mRNAs and 4955 tissue-specific lncRNAs. Among them, the abomasum tissue had the fewest tissue-specific mRNAs (91), while the omasum tissue had the most (286) (Table S7). The colon tissue had the most identified tissue-specific lncRNAs (376), while no tissue-specific lncRNAs were identified in the abomasum, cecum, and cerebellum tissues (Fig. 4A, Table S8). By predicting mRNA target genes of tissue-specific lncRNAs, we found that all tissues had predicted mRNA target genes for their tissue-specific lncRNAs. The tissues with the most predicted mRNA target genes for tissue-specific lncRNAs were fat and ileum, while eyes and lungs had the fewest (Table S9). Additionally, some tissue-specific lncRNA genes targeted tissue-specific mRNA genes related to the physiological functions of the respective tissues, such as lncRNA *MSTRG.13002.2* in adipose tissue and lncRNA *MSTRG.49212.1* in the liver (Fig. 4B).
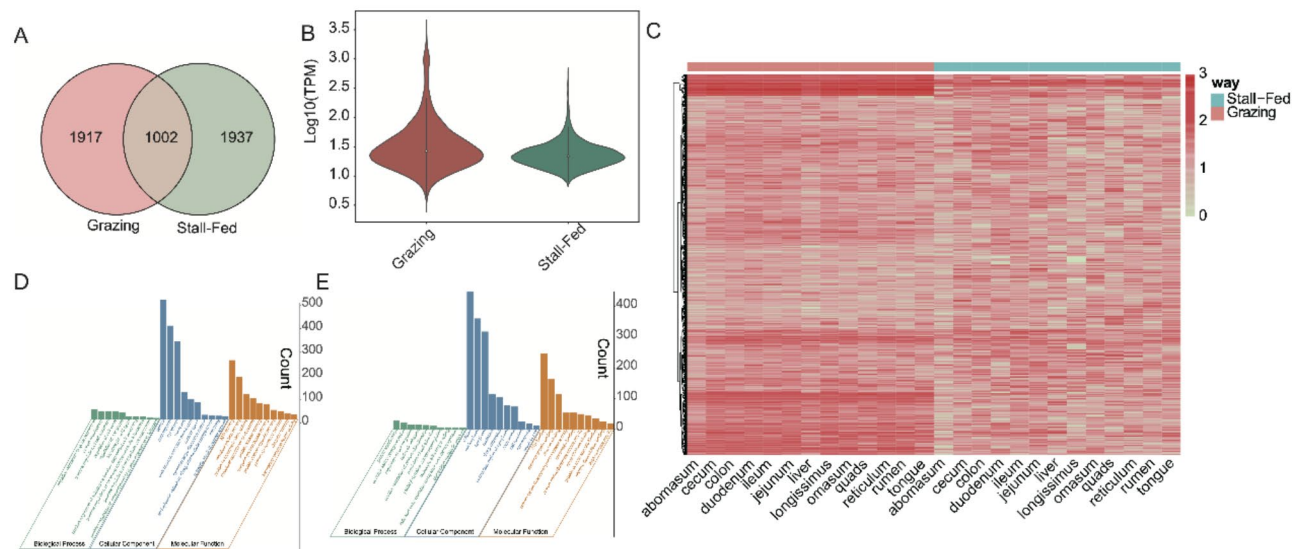
An analysis of the expression levels of tissue-specific mRNAs across all tissues revealed that the average expression level of tissue-specific mRNAs was lowest in the eyes, followed by fat (Fig. 4C). Although 9912 TSGs were identified, some of these TSGs had low expression levels in multiple tissues, leading to a stricter selection
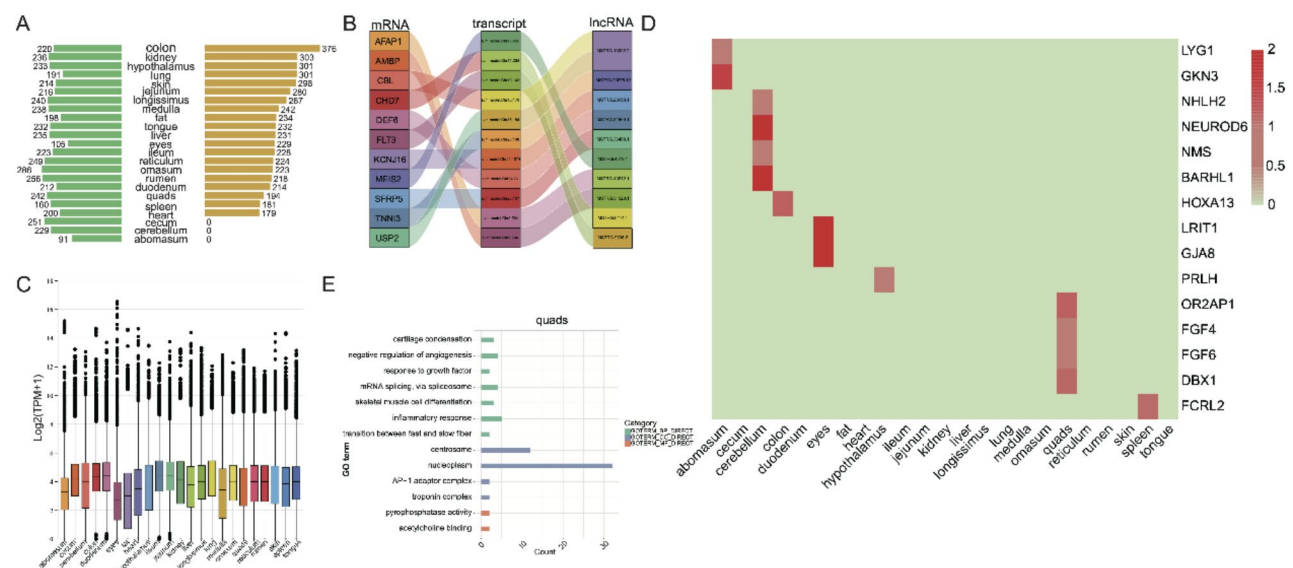
Fig. 2. HKGs in 23 tissues of black goats. (**A**) Preliminary heat map of HKGs expression. Red indicates high expression level, blue indicates low expression level. (**B**) Heat map of the preliminary screening using the average expression in each tissue. (**C**) Heat map of important HKGs expression. (**D**) Gene Ontology (GO) enrichment analysis of HKGs. E. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of HKGs.

of 127 TSGs that were expressed in only one tissue (Table S10). A heatmap showing 15 selected TSGs related to the physiological functions of specific tissues indicated high expression levels for *GKN3* in the abomasum, *NEUROD6* and *BARHL1* in the cerebellum, *LRIT1* and *GJA8* in the eyes, and *DBX1* in the quadriceps (Fig. 4D). Furthermore, KEGG and GO analyses of TSGs in each tissue (Figure S4) showed that TSGs were significantly enriched in functional pathways corresponding to their biological functions. For example, TSGs in the biceps femoris were enriched in pathways related to growth factor response, skeletal muscle cell differentiation, fast and slow fiber transition, and the troponin complex (Fig. 4E). In the heart, TSGs were enriched in pathways related to myocardial contraction, ventricular myocardium morphogenesis, regulation of cardiac contraction force, cardiac troponin complex, and adrenergic signaling in cardiomyocytes. In the liver, TSGs were enriched

Fig. 3. HKGs in Stall-fed and Grazing black goats. (**A**) Venn diagram showing the number of HKGs in stall-fed and grazing black goats. (**B**) Violin plots depicting the expression levels of HKGs in stall-fed and grazing black goats. (**C**) Heatmap illustrating the expression patterns of key HKGs in stall-fed and grazing black goats. (**D**) GO analysis of HKGs in grazing black goats. (E) GO analysis of HKGs in stall-fed black goats.

Fig. 4. TSGs in 23 Tissues of black goats. (**A**) Bar graph of the number of tissue-specific mRNAs and lncRNAs. (**B**) Sankey diagram of tissue-specific lncRNA target genes. (**C**) Box plot of tissue-specific gene expression. (**D**) Heatmap of the expression of some TSGs. (**E**) GO analysis of tissue-specific mRNAs in the biceps femoris tissue.

in pathways related to complement and coagulation cascades, platelet activation, plasminogen activation, fibrinolysis, blood coagulation, and fibrin clot formation.

Co-expression gene network analysis of black goat tissues

To investigate the biological relationships and potential functions of HUBGs in different tissues, we constructed a network based on WGCNA. We performed WGCNA on 16,885 genes that had at least one sample with TPM ≥ 1 in previous studies. When selecting $R^2 = 0.8$, the soft threshold power β was 18, which was used to effectively
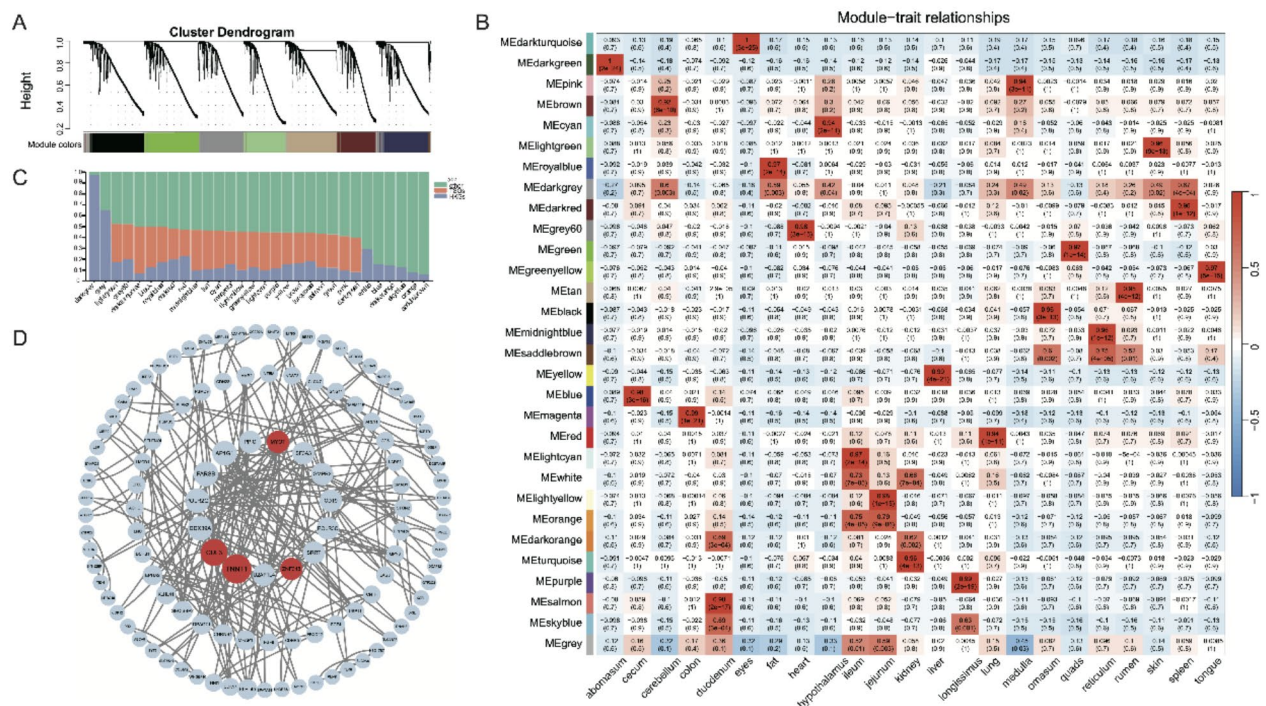
construct a scale-free network (Figure S5). Using hierarchical clustering based on the topological overlap matrix (Fig. 5A), we obtained 30 modules (Table S11). Furthermore, correlating these modules with tissues, we found that the darkgrey module had a high correlation with each tissue, while the darkgreen, blue, brown, magenta, salmon, darkturquoise, royalblue, grey60, cyan, lightcyan, lightyellow, turquoise, yellow, purple, red, pink, black, green, midnightblue, tan, lightgreen, darkred, and greenyellow modules (23 modules in total) showed extremely high correlations with the abomasum, cecum, cerebellum, colon, duodenum, eyes, fat, heart, hypothalamus, ileum, jejunum, kidney, liver, longissimus, lung, medulla, omasum, quadriceps, reticulum, rumen, skin, spleen, and tongue, respectively (Fig. 5B). We also analyzed the proportions of HKGs and TSGs within each module. The darkgrey module had the highest proportion of HKGs, followed by the grey module (Fig. 5C), while TSGs were predominantly found in the modules most highly correlated with their respective tissues (Figure S6).

In addition, by identifying HUBGs in the aforementioned 23 modules, we obtained a total of 279 HUBGs (Figure S7). We observed that in these network diagrams, these HUBGs are connected to most genes, indicating that HUBGs have close associations with genes in the related tissue-modules. For example, *ESR1*, *POLG*, *PARP1*, *FGF12* in the longest muscle-purple module (Fig. 5D), and *TNNT1*, *MYOT* in the biceps femoris-green module.

Discussion

Goats are among the earliest animals domesticated by humans, providing high-quality food and skins for human society since the agricultural period, and playing an important role in human history. With the large-scale application and rapid development of RNA-seq and analysis algorithms in animals, research on gene regulation in various tissues of many animals has made significant progress. However, studies on gene expression patterns in goats using large-scale multi-tissue data are still limited. A comprehensive study of gene expression patterns across different tissues can provide valuable insights into complex biological functions and regulatory mechanisms[42]. In our study, we aimed to explore HKGs, TSGs, and HUBGs in black goats through large-scale RNA-seq analysis of 23 different goat tissues. We preliminarily identified 2,968 low-expression variability HKGs, including six highly expressed and conserved HKGs. The number of HKGs in the Stall-fed and grazing black goats was 2,929 and 2,939, respectively, with 1,002 common HKGs. Interestingly, the HKGs in grazing black goats had a higher average expression level. In addition, we identified 9,912 TSGs, including 4,957 mRNA TSGs and 4,955 lncRNA TSGs, as well as 30 modules and 279 HUBGs.

In this study, we constructed a transcriptome atlas of 23 different tissues from black goats. These tissues were derived from seven systems: the nervous system, digestive system, musculoskeletal system, respiratory system, circulatory system, endocrine system, and urinary system. We observed that while the quantities of mRNA and lncRNA in these 23 tissues were nearly identical, their expression levels varied significantly. As

Fig. 5. Co-expression gene network analysis of 23 tissues in black goats. (**A**) Cluster dendrogram of TPM > 1. (**B**) Heatmap of the relationship between the 30 modules and 23 tissues. Red represents high correlation, blue represents low correlation. (**C**) The distribution of HKGs, TSGs, and other genes across various modules. (**D**) protein-protein interaction network (PPI) interaction diagram of HUBGs in the biceps femoris. The innermost genes representing the HUBGs.

a ruminant, the goat possesses a more complex digestive system compared to most mammals. Our analysis showed strong correlation in gene expression among certain tissues within the digestive system (duodenum, jejunum, ileum, cecum, colon) and among the fore-stomach tissues (rumen, reticulum, omasum, abomasum). However, we noted poor correlation between the fore-stomach tissues and the abomasum. Additionally, in the nervous system, the cerebellum, medulla oblongata, and hypothalamus exhibited good correlation, likely due to similarities in their physiological functions. These findings suggest that tissues with similar physiological functions share certain gene expression profiles, potentially linked to tissue differentiation. This observation aligns with multi-tissue gene expression patterns previously reported in cattle studies[43].

HKGs are widely and stably expressed across different tissues, serving as internal controls in many biotechnological applications and genomic studies[21,44]. To better understand the HKGs in black goats, we screened 11,872 HKGs and observed that their expression levels varied across different tissues. Notably, we found lower expression levels of HKGs in the abomasum and eyes. After further screening, we identified 2,968 highly conserved HKGs with high expression levels. From these, we selected six important HKGs and found that compared to commonly used reference genes in black goats, such as *ACTB*, *UBC*, *PPIB*, *RPS13*, *EIF3K*, and *SDHA*, the genes *EEF1A1*, *RPS6*, *RPL7*, *RPL23*, *RPS15*, and *RPL19* exhibited higher expression levels and greater conservation. Research suggests that these important HKGs can serve as reference genes in other species. For instance, *EEF1A1* is the most stable reference gene during different states in bats and during the differentiation of human vascular smooth muscle cells into fat cells[45,46]. *RPS6* is the most suitable reference gene in the goat rumen[47]. *RPL7* is considered the most reliable reference gene in Ferrisia gilli Gullan and cotton aphids[48,49]. *RPL23* is the best candidate reference gene in chicken thigh muscles[50]. *RPS15* is one of the most stable reference genes during the developmental cycle of the yak rumen[51]. *RPL19* is the most suitable reference gene for the temporomandibular joint of sheep[52]. This indicates that the important HKGs we screened can serve as new reference genes for researchers studying black goats. However, under different physiological conditions, some HKGs may not be stably expressed[53]. Therefore, we conducted a large-scale analysis of HKGs in black goats under different feeding methods and found significant differences. Although the number of HKGs identified in both feeding methods was nearly the same, only one-third of the HKGs were common to both. Furthermore, the average expression level of HKGs in grazing black goats was higher, with greater variability, possibly due to the complex environment in which grazing black goats are located. Additionally, the screened HKGs in black goats are associated with ribosomes, oxidative phosphorylation, mitochondrial respiratory chain complex I, protein processing in the endoplasmic reticulum, and translation initiation factor activity, all of which are related to maintaining basic cell functions and energy metabolism, consistent with previous reports[24,54].

Understanding TSGs in various tissues enhances our grasp of the complex genetic traits and biological processes specific to those tissues[55]. In this study, We screened a total of 4957 mRNA TSGs and 4955 lncRNA TSGs from 23 tissues of Guangxi black goats. Notably, some mRNAs interacting with lncRNAs were linked to the physiological functions of their respective tissues. For instance, in the liver, the lncRNA *MSTRG.49212.1* was predicted to target the *AFAP1*, with previous studies indicating that the lncRNA *AFAP1-AS1* promotes liver cancer development[56]. In adipose tissue, the lncRNA *MSTRG.13002.2* was predicted to target the mRNA *USP2*, which research by Saito Natsuko et al. showed alters insulin sensitivity in obese mice through adipocyte-dependent mechanisms[57]. Furthermore, rigorous screening identified 127 TSGs expressed exclusively in a single tissue. Analysis of these highly expressed genes revealed their involvement in the physiological functions of the corresponding tissues. For example, in the rumen, the TSG *AQP12B* is implicated in the secretion of digestive enzymes[58], while *GKN3* is a specific marker for chronic gastritis[59]. In the hypothalamus, the TSG *PRLH* plays a key role in estrogen processes[60]. In the eyes, the TSG *LRIT1* regulates synaptic connections between cone photoreceptor cells and bipolar cells in the retina[61], and *GJA8*, a major gap junction protein in vertebrate lenses, is linked to human cataracts when mutated[62]. Additionally, these TSGs are enriched in pathways closely related to the physiological functions of their respective tissues. For example, the regulation of the actin cytoskeleton may be involved in muscle development[63], while sarcomere organization and cardiac muscle contraction are crucial for heart function[64,65].

Co-expression network analysis is a common method for exploring gene relationships based on transcriptome expression. In this study, network co-expression analysis was used to construct 30 modules from 23 tissues. Through correlation analysis between modules and tissues, we found that each TSG was included in the corresponding most correlated module. This suggests that the expression patterns of TSGs in each tissue are relatively similar. Interestingly, the dark grey module was found to consist almost entirely of HKGs, though some HKGs were also present in other modules. This indicates that the expression patterns of certain HKGs are not necessarily similar, while some HKGs share more similar expression patterns with TSGs. Furthermore, some of the HUBGs identified are closely related to the physiological functions of their respective tissues. For example, mutations in HUBGs *TNNT1* and *CUL3* in the biceps femoris can cause nemaline myopathy[66,67]; mutations in the liver HUBG *ALB*, which encodes albumin, can lead to liver cancer[68], and overexpression of *HRG* can promote liver cancer cell metastasis[69]. Additionally, the HUBG *FASN* in adipose tissue can influence the width and weight of goat tail fat when mutated[70].

In conclusion, we identified HKGs, TSGs, and HUBGs in Guangxi black goats. Additionally, we observed significant differences in the overall expression levels of HKGs between stall-fed and grazing black goats. These findings improve our understanding of gene expression regulation across different tissues and rearing environments in black goats, laying a foundation for further research into biological development and processes within various goat tissues.

Materials and methods

Collection of samples

In this experiment, all animals received humane care as outlined in the NIH Guide for the Care and Use of Laboratory Animals. The animal experiments were approved by the Animal Experiment Ethics Review Committee of Guangxi University, Nanning, China (Grant NO.: Gxu-2020-1001). Also, all the experiments in the manuscript follows the recommendations in the Animal Research Reporting in Vivo Experiments (ARRIVE) guidelines. In a commercial slaughterhouse in Nanning, Guangxi Zhuang Autonomous Region, six 12-month-old Castrated rams black goats were euthanized using the instantaneous high-voltage electric shock method, and a total of 108 samples were collected from them. These rams included three stall-fed black goats and grazing black goats. The samples were taken from 23 different tissues, including the rumen, reticulum, omasum, abomasum, duodenum, jejunum, ileum, cecum, colon, liver, longissimus dorsi muscle, biceps femoris muscle, tongue, cerebellum, heart, spleen, kidney, lung, eye, medulla oblongata, hypothalamus, fat, and skin. The samples were rapidly placed in labeled centrifuge tubes, immediately frozen in liquid nitrogen, and then transferred to a -80 °C freezer upon returning to the laboratory for subsequent transcriptome sequencing.

Library construction and transcriptome sequencing

Total RNA from 23 different tissues of 6 Guangxi black goats was extracted using Trizol reagent. Genomic DNA was removed by treating the samples with DNase I. The RNA quality and concentration were assessed using a NanoDrop 2000 spectrophotometer. To eliminate ribosomal RNA (rRNA), specific probes were employed. Subsequently, sequencing libraries were constructed using the NEBNext Ultra RNA Library Prep Kit for Illumina, provided by NEB (USA). The libraries were then subjected to high-throughput sequencing on the Illumina HiSeq 4000 platform[71].

Data quality control, alignment, and processing

A uniform approach was employed for analyzing each sample. The raw RNA-seq data were processed into clean data using fastp (v0.23.4) software, with default parameters for quality control[72]. FastQC (v0.12.1) was used to assess the quality of the sequencing data. The reference genome index was constructed using the hisat2-build program from Hisat2 software (v2.2.1) (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/001/704/415/GCF_001704415.2_ARS1.2/)[73]. Subsequently, clean data from each sample were mapped to the reference genome using Hisat2 software, and the resulting SAM files were converted to BAM format using Samtools (v1.17)[74]. Gene assembly for each sample was conducted using Stringtie (v2.2.1)[75], and the merge parameter of Stringtie was used to combine all files for standardization. Gene expression in transcripts per million (TPM) was obtained using Stringtie (v2.2.1). GffCompare software (v0.12.6) was utilized to select transcripts with class codes "i," "x," "u," "o," or "e." The protein-coding potential of lncRNA transcripts was predicted using four methods: CPC2, CPAT, PLEK, and Pfam. By analyzing combinations of these four filtering criteria, final candidate lncRNAs were determined, excluding lncRNA transcripts possessing protein-coding potential. To elucidate the function of lncRNAs, the differentially regulated target mRNAs of lncRNAs were predicted. Python scripts were employed to select upstream and downstream coding genes within 100 kb, considered potential cis-regulatory target mRNAs. The Pearson correlation coefficient (r) in the corresponding relationship was calculated using the Hmisc package (v5.1) in R to examine the correlation between differential lncRNAs and mRNAs.

Gene expression pattern among tissues

In this study, to investigate the gene expression patterns in the 23 tissues of Guangxi black goats, we first tallied the number of expressed genes (both mRNA and lncRNA) in each tissue. Subsequently, we conducted sample clustering based on genes with TPM > 1 in at least one sample across all samples, followed by further analysis. Then, we generated a PCA clustering plot using genes with TPM > 1 in at least one sample across all samples. We computed the average gene expression values within replicate samples (samples from the same tissue) and employed these averages for inter-tissue correlation analysis.

Detection of tissue-specifc gene in black goats

Firstly, genes that are lowly expressed (TPM < 1) across all samples are excluded, and then the average TPM value for all biological replicates is calculated. If a gene's TPM value in a particular tissue is at least three times higher than its TPM value in other tissues, it is defined as a Tissue-Specifc Gene. The analysis of TSGs for lncRNA is consistent with mRNA. Bar graphs are plotted to depict the number of TSGs and tissue-specific lncRNAs. Box plots are generated to illustrate the expression of TSGs across various tissues. We employed DAVID for GO and KEGG enrichment analysis of TSGs in black goats[76,77] and visualized the results using R software (v4.2.3) and ggplot2 (v3.4.2) package[78]. TSGs expressed only in one tissue are defined as specific TSGs, and heatmaps are plotted for important specific TSGs. Using the STRING software, a PPI is constructed for TSGs, and significant gene networks are built[79]. Visualization is conducted using Cytoscape software (v3.10.0)[80].

Detection of housing-keeping genes in black goats

First, genes with low expression (TPM < 1) are removed as preliminary HKGs. In addition, the CV is used to assess the variability of each preliminary HKG[44]. Specifically, the CV is defined as the ratio of the standard deviation to the mean (CV = σ/μ, where σ represents the standard deviation and μ represents the mean)[43]. Subsequently, the CVs are categorized into low, medium, and high variability groups based on the quartiles of the overall distribution. The analysis of HKGs in grazing black goats and stall-fed black goats is consistent with that of the black goats. Low-variability HKGs are visualized in a heatmap using the pheatmap (v1.0.12) package[81], showing the different expression patterns of the genes in each tissue. The low-variability HKGs group is further divided into low, medium, and high expression groups. Sort all housekeeping genes by their CV in

ascending order, select the top 100 HKGs, and create a heatmap. Identify the highly expressed and highly stable genes as important HKGs. Important HKGs and commonly used HKGs in black goats are also visualized using the pheatmap (v1.0.12) package. The HKGs of black goats are subjected to GO and KEGG enrichment analysis using DAVID, and the results are visualized using the ggplot2 (v3.4.2) package in R software (v4.2.3). The low-variability HKGs in grazing and stall-fed black goats are analyzed, and Venn diagrams and violin plots are generated. Heatmaps of common and unique low-variability HKGs in both grazing and stall-fed black goats are created using the pheatmap (v1.0.12) package. The GO and KEGG enrichment analysis of the HKGs in grazing and stall-fed black goats is consistent with that of black goats. The tissues analyzed include 13 types: rumen, reticulum, omasum, abomasum, duodenum, jejunum, ileum, cecum, colon, liver, longissimus dorsi muscle, biceps femoris, and tongue.

Co-expression network analysis

The co-expression network analysis of gene expression is performed using the WGCNA package (v.1.72-1) in R programming[35]. Briefly, we first constructed an expression matrix using tissues from 23 black goats. Cluster analysis of all tissues was carried out using clustering functions and the method of longest distance. The soft-thresholding power (β) was determined based on a scale-free topology. The softConnectivity function was used to calculate the connectivity between each gene based on the expression matrix datExpr. The adjacency function was employed to calculate the adjacency between genes in the expression data datExpr. The TOMdist function was used to compute the dissimilarity 1-TOM, i.e., the dissimilarity matrix dissTOM. Based on the TOM dissimilarity, hierarchical clustering analysis was conducted using the hclust function, constructing a hierarchical clustering tree of genes. Module identification and gene module division based on the hierarchical clustering tree were visualized. PCA was used to calculate the eigengenes (ME) of each module, visualizing the entire gene co-expression network, and analyzing the correlation between samples and gene modules. The gene composition of different modules was depicted in bar charts. The distribution of TSGs in 23 modules across 23 tissues was visualized using R software (v4.2.3) and the ggplot2 package (v3.4.2).

Data availability

References

1. Zeder, M. A. & Hesse, B. The initial domestication of goats (Capra hircus) in the Zagros Mountains 10,000 years ago. *Science*. **287**, 2254–2257 (2000).
2. Saleh, A. A., Rashad, A. M. A., Hassanine, N. N. A. M., Sharaby, M. A. & Zhao, Y. Evaluation of morphological traits and physiological variables of several Chinese goat breeds and their crosses. *Trop. Anim. Health Prod.* **53**, 74 (2021).
3. Naderi, S. et al. The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proc. Natl. Acad. Sci. USA*. **105**, 17659–17664 (2008).
4. Sacarrão-Birrento, L. & de Almeida, A. M. The Portuguese Serrana goat breed: a review. *Trop. Anim. Health Prod.* **53**, 114 (2021).
5. van Leeuwen, S. S. et al. Goat milk oligosaccharides: their diversity, quantity, and functional properties in comparison to human milk oligosaccharides. *J. Agric. Food Chem.* **68**, 13469–13485 (2020).
6. Bertolini, F. et al. Signatures of selection and environmental adaptation across the goat genome post-domestication. *Genet. Sel. Evol. GSE.* **50**, 57 (2018).
7. Saleh, A. A., Rashad, A. M. A., Hassanine, N. N. A. M., Sharaby, M. A. & Zhao, Y. Assessment of hair and cashmere properties and their genetic background of several goat breeds in Southwest China. *Sci. Rep.* **12**, 11135 (2022).
8. Gawat, M., Boland, M., Singh, J. & Kaur, L. Goat meat: production and quality attributes. *Foods Basel Switz.* **12**, 3130 (2023).
9. MacHugh, D. E. & Bradley, D. G. Livestock genetic origins: goats buck the trend. *Proc. Natl. Acad. Sci. USA.* **98**, 5382–5384 (2001).
10. Stark, R., Grzelak, M. & Hadfield, J. RNA sequencing: the teenage years. *Nat. Rev. Genet.* **20**, 631–656 (2019).
11. Wu, C., Qin, C., Fu, X., Huang, X. & Tian, K. Integrated analysis of lncRNAs and mRNAs by RNA-Seq in secondary hair follicle development and cycling (anagen, catagen and telogen) of Jiangnan cashmere goat (Capra hircus). *BMC Vet. Res.* **18**, 167 (2022).
12. Shen, J. et al. Deep small RNA sequencing reveals important miRNAs related to muscle development and intramuscular fat deposition in longissimus dorsi muscle from different goat breeds. *Front. Vet. Sci.* **9**, 911166 (2022).
13. Poklukar, K. et al. Adipose tissue gene expression of entire male, immunocastrated and surgically castrated pigs. *Int. J. Mol. Sci.* **22**, 1768 (2021).
14. La, Y. et al. Identification and characterization of Piwi-interacting RNAs for early testicular development in Yak. *Int. J. Mol. Sci.* **23**, 12320 (2022).
15. Zhang, M. et al. Transcriptome analysis reveals the molecular regulatory network of muscle development and meat quality in Sunit lamb supplemented with dietary probiotic. *Meat Sci.* **194**, 108996 (2022).
16. Yao, Y. C. et al. Transcriptome analysis of sheep follicular development during prerecruitment, dominant, and mature stages after FSH superstimulation. *Domest. Anim. Endocrinol.* **74**, 106563 (2021).
17. Zhen, H. et al. Characteristics and expression of circ_003628 and its promoted effect on proliferation and differentiation of skeletal muscle satellite cells in goats. *Anim. Open. Access. J. MDPI.* **12**, 2524 (2022).
18. Hou, B. et al. Transcriptome analysis reveals mRNAs and long non-coding RNAs associated with fecundity in the hypothalamus of high-and low-fecundity goat. *Front. Vet. Sci.* **10**, 1145594 (2023).
19. Li, L. et al. The novel RNA-RNA activation of H19 on MyoD transcripts promoting myogenic differentiation of goat muscle satellite cells. *Int. J. Biol. Macromol.* **253**, 127341 (2023).
20. Levy, B. & Dixon, G. H. Diversity of sequences of polyadenylated cytoplasmic RNA from rainbow trout (Salmo gairdnerii) testis and liver. *Biochemistry*. **16**, 958–964 (1977).
21. Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. *Trends Genet. TIG.* **29**, 569–574 (2013).
22. Chang, C. W. et al. Identification of human housekeeping genes and tissue-selective genes by microarray meta-analysis. *PLoS ONE.* **6**, e22859 (2011).
23. Pseudogenes as weaknesses of ACTB (Actb). and GAPDH (gapdh) used as reference genes in reverse transcription and polymerase chain reactions - PubMed. https://pubmed.ncbi.nlm.nih.gov/22927912/

24. Pan, X. et al. Expression profile of housekeeping genes and tissue-specific genes in multiple tissues of pigs. *Animals.* **12**, 3539 (2022).

25. Joshi, C. J., Ke, W., Drangowska-Way, A., O'Rourke, E. J. & Lewis, N. E. What are housekeeping genes? *PLoS Comput. Biol.* **18**, e1010295 (2022).

26. Gong, H. et al. Evaluation of candidate reference genes for RT-qPCR studies in three metabolism related tissues of mice after caloric restriction. *Sci. Rep.* **6**, 38513 (2016).

27. Zhao, X. et al. The selection and identification of compound housekeeping genes for quantitative real-time polymerase chain reaction analysis in rat fetal kidney. *J. Appl. Toxicol. JAT.* **42**, 360–370 (2022).

28. Sullivan-Gunn, M., Hinch, E., Vaughan, V. & Lewandowski, P. Choosing a stable housekeeping gene and protein is essential in generating valid gene and protein expression results. *Br. J. Cancer.* **104**, 1055 (2011). author reply 1056.

29. Blotas, C., Férec, C. & Moisan, S. Tissue-specific regulation of CFTR gene expression. *Int. J. Mol. Sci.* **24**, 10678 (2023).

30. Ryaboshapkina, M. & Hammar, M. Tissue-specific genes as an underutilized resource in drug discovery. *Sci. Rep.* **9**, 7233 (2019).

31. Sonawane, A. R. et al. Understanding tissue-specific gene regulation. *Cell. Rep.* **21**, 1077–1088 (2017).

32. Zhang, L. & Li, W. H. Mammalian housekeeping genes evolve more slowly than tissue-specific genes. *Mol. Biol. Evol.* **21**, 236–239 (2004).

33. Zhang, S. et al. Screening of bovine tissue-specific expressed genes and identification of genetic variation within an adipose tissue-specific lncRNA gene. *Front. Vet. Sci.* **9**, 887520 (2022).

34. Rogala, S. et al. The lncRNA Sweetheart regulates compensatory cardiac hypertrophy after myocardial injury in murine males. *Nat. Commun.* **14**, 7024 (2023).

35. Langfelder, P. & Horvath, S. W. G. C. N. A. An R package for weighted correlation network analysis. *BMC Bioinf.* **9**, 559 (2008).

36. Nangraj, A. S. et al. Integrated PPI- and WGCNA-retrieval of hub gene signatures shared between barrett's esophagus and esophageal adenocarcinoma. *Front. Pharmacol.* **11**, 881 (2020).

37. Feng, S. et al. Integrative analysis from multicenter studies identifies a WGCNA-derived cancer-associated fibroblast signature for ovarian cancer. *Front. Immunol.* **13**, 951582 (2022).

38. Luo, D. et al. Long RNA profiles of human brain extracellular vesicles provide new insights into the pathogenesis of alzheimer's disease. *Aging Dis.* **14**, 229–244 (2023).

39. Yu, T. et al. Transcriptome, proteome and metabolome analysis provide insights on fat deposition and meat quality in pig. *Food Res. Int. Ott. Ont.* **166**, 112550 (2023).

40. Silva-Vignato, B. et al. Gene co-expression networks associated with carcass traits reveal new pathways for muscle and fat deposition in Nelore cattle. *BMC Genom.* **20**, 32 (2019).

41. Wang, J., Chen, H. & Zeng, X. Identification of hub genes associated with follicle development in multiple births sheep by WGCNA. *Front. Vet. Sci.* **9**, 1057282 (2022).

42. Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).

43. Zhang, T. et al. Transcriptional atlas analysis from multiple tissues reveals the expression specificity patterns in beef cattle. *BMC Biol.* **20**, 79 (2022).

44. de Jonge, H. J. M. et al. Evidence based selection of housekeeping genes. *PLoS ONE.* **2**, e898 (2007).

45. Friedrichs, V., Balkema-Buschmann, A., Dorhoi, A. & Pei, G. Selection and stability validation of reference gene candidates for transcriptional analysis in Rousettus aegyptiacus. *Sci. Rep.* **11**, 21662 (2021).

46. Gentile, A. M. et al. RPL13A and EEF1A1 are suitable reference genes for qPCR during adipocyte differentiation of vascular stromal cells from patients with different BMI and HOMA-IR. *PLoS ONE.* **11**, e0157002 (2016).

47. Zhao, J. et al. Genome-wide identification of reference genes for reverse-transcription quantitative PCR in goat rumen. *Anim. Open. Access. J. MDPI.* **11**, 3137 (2021).

48. Bansal, R., Haviland, D. R. & Hunter, W. B. Selection and validation of reference genes for quantifying gene expression in the Gill's mealybug. *J. Econ. Entomol.* **116**, 2166–2172 (2023).

49. Ma, K. S. et al. Identification and validation of reference genes for the normalization of gene expression data in qRT-PCR analysis in Aphis gossypii (Hemiptera: Aphididae). *J. Insect Sci. Online.* **16**, 17 (2016).

50. Mogilicherla, K., Athe, R. P., Chatterjee, R. N. & Bhattacharya, T. K. Identification of suitable reference genes for normalization of quantitative real-time PCR-based gene expression in chicken (Gallus gallus). *Anim. Genet.* **53**, 881–887 (2022).

51. Min, Q., Yang, L., Wang, Y., Liu, Y. & Jiang, M. Transcriptome-based evaluation of optimal reference genes for quantitative real-time PCR in yak stomach throughout the growth cycle. *Anim. Open. Access. J. MDPI.* **13**, 925 (2023).

52. Jiang, X. et al. Evaluation of reference gene suitability for quantitative expression analysis by quantitative polymerase chain reaction in the mandibular condyle of sheep. *Mol. Med. Rep.* **12**, 5633–5640 (2015).

53. Arukwe, A. Toxicological housekeeping genes: do they really keep the house? *Environ. Sci. Technol.* **40**, 7944–7949 (2006).

54. Li, B. et al. A comprehensive mouse transcriptomic bodymap across 17 tissues by RNA-seq. *Sci. Rep.* **7**, 4200 (2017).

55. Fang, L. et al. Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. *Genome Res.* **30**, 790–801 (2020).

56. Zhang, J. Y. et al. Long noncoding RNA AFAP1-AS1 indicates a poor prognosis of hepatocellular carcinoma and promotes cell proliferation and invasion via upregulation of the RhoA/Rac2 signaling. *Int. J. Oncol.* **48**, 1590–1598 (2016).

57. Saito, N. et al. Macrophage ubiquitin-specific protease 2 modifies insulin sensitivity in obese mice. *Biochem. Biophys. Rep.* **9**, 322–329 (2017).

58. Itoh, T. et al. Identification of a novel aquaporin, AQP12, expressed in pancreatic acinar cells. *Biochem. Biophys. Res. Commun.* **330**, 832–838 (2005).

59. Bockerstett, K. A. et al. Single cell transcriptional analyses identify lineage-specific epithelial responses to inflammation and metaplastic development in the gastric corpus. *Gastroenterology.* **159**, 2116–2129e4 (2020).

60. Zhong, Y., Di, R., Yang, Y., Liu, Q. & Chu, M. Transcriptome analysis of neuroendocrine regulation of ovine hypothalamus-pituitary-ovary axis during ovine anestrus and the breeding season. *Genes.* **12**, 1861 (2021).

61. Ueno, A. et al. Lrit1, a retinal transmembrane protein, regulates selective synapse formation in cone photoreceptor cells and visual acuity. *Cell. Rep.* **22**, 3548–3561 (2018).

62. Ping, X. et al. Rapamycin relieves the cataract caused by ablation of Gja8b through stimulating autophagy in zebrafish. *Autophagy.* **17**, 3323–3337 (2021).

63. Liu, Y., Jia, Y., Liu, C., Ding, L. & Xia, Z. RNA-Seq transcriptome analysis of breast muscle in Pekin ducks supplemented with the dietary probiotic Clostridium butyricum. *BMC Genom.* **19**, 844 (2018).

64. Zhu, J. et al. Identification of tissue-specific protein-coding and noncoding transcripts across 14 human tissues using RNA-seq. *Sci. Rep.* **6**, 28400 (2016).

65. Perreault, L. R., Le, T. T., Oudin, M. J. & Black, L. D. RNA sequencing indicates age-dependent shifts in the cardiac fibroblast transcriptome between fetal, neonatal, and adult developmental ages. *Physiol. Genom.* **53**, 414–429 (2021).

66. Wang, G., Zhao, D., Yan, C. & Lin, P. Exon skipping caused by splicing mutation in TNNT1 nemaline myopathy. *J. Hum. Genet.* **68**, 97–101 (2023).

67. Mansur, A. et al. Dynamic regulation of inter-organelle communication by ubiquitylation controls skeletal muscle development and disease onset. *eLife* **12**, e81966 (2023).

68. Dietlein, F. et al. Genome-wide analysis of somatic noncoding mutation patterns in cancer. *Science.* **376**, eabg5601 (2022).

69. Yin, Y. et al. HRG inhibits liver cancer lung metastasis by suppressing neutrophil extracellular trap formation. *Clin. Transl Med.* **13**, e1283 (2023).
70. Zhao, L. et al. Ovine ELOVL5 and FASN genes polymorphisms and their correlations with sheep tail fat deposition. *Gene.* **807**, 145954 (2022).
71. Li, H. et al. Comparison of long non-coding RNA expression profiles of cattle and buffalo differing in muscle characteristics. *Front. Genet.* **11**, 98 (2020).
72. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinforma Oxf. Engl.* **34**, i884–i890 (2018).
73. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods.* **12**, 357–360 (2015).
74. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinforma Oxf. Engl.* **25**, 2078–2079 (2009).
75. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
76. Sherman, B. T. et al. A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* **50**, W216–W221 (2022).
77. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
78. Dennis, G. et al. Database for annotation, visualization, and integrated discovery. *Genome Biol.* **4**, P3 (2003).
79. Szklarczyk, D. et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **51**, D638–D646 (2023).
80. Juengel, J. L., Meberg, B. M., Turzillo, A. M., Nett, T. M. & Niswender, G. D. Hormonal regulation of messenger ribonucleic acid encoding steroidogenic acute regulatory protein in ovine corpora lutea. *Endocrinology.* **136**, 5423–5429 (1995).
81. Zhao, Y. et al. A compendium and comparative epigenomics analysis of cis-regulatory elements in the pig genome. *Nat. Commun.* **12**, 2217 (2021).

Author contributions
Q.L. designed and led the project. Z.X. and Y.C. assisted in sample collection. C.Q. collected and processed the samples, and wrote the manuscript. D.W. conducted data analysis. M.W. and H.H. provided funding. X.W. and Z.L. provided suggestions and assisted in resolving proofreading issues. All authors participated in reading and approving the submitted version.

Funding

Declarations

Competing interests
The authors declare no competing interests.

Ethics approval
This animal study was reviewed and approved by the Guangxi University Animal Care and Use Committee (Nanning, China), in strict compliance with the Regulations for the Administration of Affairs Concerning Experimental Animals (No. 588 Document of the State Council of the People's Republic of China, 2011), and informed consent was obtained from the company's directors to allow their animals to participate in this study.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-72844-8.

Correspondence and requests for materials should be addressed to M.W., Z.L. or Q.L.

Reprints and permissions information is available at www.nature.com/reprints.

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