

RESEARCH

Open Access



Integrated multi-omics analysis and functional validation uncovers *RPL26* roles in regulating growth traits of Asian water buffaloes (*Bubalus bubalis*)

Yangyang Shen^{1,2}, Zhenjiang An^{1,3}, Linna Gao^{1,4}, Mingfa Gu⁵, Shuwen Xia^{1,2}, Qiang Ding^{1,2}, Yinxia Li^{1,2}, Shaoxian Cao^{1,2}, Jianbin Li⁶, Jinming Huang⁶, Jifeng Zhong^{1,2}, Kunlin Chen^{1,2*}, Xiao Wang^{6,7*} and Huili Wang^{1,2*}

Abstract

Background Asian water buffaloes (*Bubalus bubalis*) in the Yangtze River Basin of China are the important meat provider for local residents because of its outstanding body size. Several previous studies have highlighted their genetic basis of growth characteristics, but the crucial genes regulating growth traits via multi-layer omics are still rarely investigated.

Results We conducted a comprehensive multi-omics analysis integrating blood and muscle transcriptome, plasma metabolome, rumen fluid metagenome, and genome of Haizi water buffaloes. Of note, ribosomal protein L26 (*RPL26*) located in the evolutionary selection regions associated with body sizes is the top differentially expressed gene (DEG) in both blood and muscle tissues. Further metabolomics and metagenomics identified growth-related molecular biomarkers (myristicin and Bacteroidales) and microbiological composition (Bacteroides and Prevotella). Leveraging cattle quantitative trait loci (QTLs) and genotype-tissue expression (CattleGTEx) databases, we found the significant correlations of QTL_180979 on *RPL26* and two identified cis-eQTLs in muscle tissue in the upstream of *RPL26* with weight gain. The follow-up cell assay validations confirmed the regulation roles of *RPL26* in cell cycle, apoptosis, and differentiation, where the low *RPL26* expressions enhanced the antiapoptotic ability and promoted the differentiation of myoblasts into myotubes markedly.

Conclusions Our study illustrates *RPL26* roles in regulating growth traits via both integrated multi-omics analysis and functional validations that suggests the further applications of *RPL26* for growth trait selection of water buffaloes.

Keywords Asian water buffalo, Multi-omics, Functional validation, *RPL26*, Growth trait

*Correspondence:

Kunlin Chen
chenkunlin@jaas.ac.cn
Xiao Wang
xiaowangzntc@163.com
Huili Wang
wanghuili318@163.com

¹Institute of Animal Science, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China

²Jiangsu Provincial Engineering Research Center of Precision Animal Breeding, Nanjing 210014, China

³College of Animal Science and Technology, Nanjing Agricultural University, Nanjing 210095, China

⁴School of Life Sciences and Food Engineering, Hebei University of Engineering, Handan 056038, China

⁵Sheyang Buffalo Farm, Sheyang 224300, China

⁶Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences, Jinan 250100, China

⁷Shandong OX Livestock Breeding Co., Ltd., Jinan 250100, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

In China, Asian water buffaloes (*Bubalus bubalis*) are mainly distributed in the rice-growing regions of the Yangtze River Basin as the meat provider for local residents because of its outstanding body size, such as Haizi (HZ) water buffaloes [1]. The genetic basis of growth characteristics has been investigated to understand the mechanisms underlying growth traits [2, 3], but the related genes were seldom identified. Therefore, it's valuable to explore the crucial genes regulating growth traits of HZ water buffaloes due to its more advantageous growth characteristics than other buffaloes.

The cattle genome has been extensively shaped by artificial selection [4–6] and the selection signature analysis facilitates the understanding of genetic mechanisms [7–9]. The frequencies of beneficial mutations selected over generations increase within populations and ultimately influence the final phenotypes [10]. In buffaloes, SNPs in exon and promoter regions of *HSP* family were associated with milk production, heat resistance, stress recovery, and disease susceptibility traits [1, 11]. Compared with quantitative trait locus (QTL) mapping, the current multi-omics methods enable deeper insights into biological mechanisms [12, 13]. The integration of differentially expressed genes (DEGs) with genome-wide association study (GWAS) increased the identification of 38 pleiotropic key regulators [14]. Utilizing the cattle Genotype-Tissue Expression (CattleGTEx), Liu et al. characterized the genetic regulatory variants on gene expressions and linked those gene expressions to 43 economic traits for different tissues [15].

Recent advances in multi-omics approaches have significantly enhanced our understanding of the genetic and molecular basis of complex traits in livestock. High-throughput sequencing technologies, combined with integrative omics analyses, have enabled researchers to dissect the intricate regulatory networks underlying growth traits, uncovering key genes, metabolic pathways, and microbiome interactions that contribute to phenotypic variation. In particular, studies on cattle and other ruminants have demonstrated that integrating transcriptomics, metabolomics, and metagenomics provides a more comprehensive view of the factors influencing growth performance and feed efficiency. A comprehensive investigation into the molecular mechanisms underlying complex growth traits has profound implications in genetic improvement and innovative utilization of the buffalo breeds [16, 17]. Therefore, we conducted a comprehensive multi-omics analysis of buffaloes using their integrated blood and muscle transcriptome, plasma metabolome, rumen metagenome, and genome, which aims to identify key genes regulating growth traits, the growth-related metabolites and rumen microbes, and their enriched pathways. Furthermore, we utilized cattle

QTLs and CattleGTEx databases to validate the QTLs and eQTLs on the identified key genes and functional experiments at the cellular level to confirm their regulation roles. This study will provide growth-related key genes and molecular biomarkers for the practical applications for selecting superior growth characteristics of Asian water buffaloes.

Materials and methods

Sample collection

HZ buffaloes were selected from a national conservation farm in Sheyang city, Jiangsu Province, under uniform feeding conditions and aged approximately 30 months. The buffalo population was individually weighed, and on the basis of weight differences, four high-weight (522 ± 23.87 kg) and four low-weight (444 ± 20.74 kg) individuals were chosen (Fig. S1A). The high-weight group averaged approximately 1.2 times the weight of the low-weight group. Before slaughter, live weights were recorded following a 24-hour fasting period and 8 h without water. Net meat weight was determined postslaughter by subtracting bones, inner block-shaped fat, and ligaments. Organ and bone weights were measured via standard procedures (Fig. S1B). Each buffalo provided two blood samples, one for RNA sequencing and the other for nontargeted metabolome sequencing. Additionally, samples of the longest back muscle were collected for transcriptome sequencing. Slaughter procedures adhered to GB/T19477-2018 guidelines, ensuring ethical handling. The specimens were taken from the longissimus dorsi muscle's identical position on the left half of the carcass within 15 min postslaughter, rinsed with cold saline, sanitized, and promptly frozen in 5 mL tubes in liquid nitrogen.

RNA extraction and RNA-seq analysis

Total RNA was extracted via TRIzol reagent, and its concentration, purity, and integrity were assessed with a Qubit 4, Nanodrop spectrophotometer, and Agilent 2100 Bioanalyzer kits. cDNA libraries were prepared following the protocols of the Illumina TruSeq™ RNA Sample Preparation Kit. After quality assessment of the libraries, PE150 sequencing was conducted on the Illumina HiSeq platform by Novogene Co., Ltd. (Beijing, China). HISAT2 (v2.0.5) [18] was employed to align the clean data against the buffalo reference genome (NDDB_SH_1, GenBank: GCA_019923935.1), followed by quantification via FeatureCounts (v1.5.0-p3) [19], resulting in counts and fragments per kilobase million (FPKM) values. Gene expression levels were ranked on the basis of scores derived from log₂-fold change ($\text{Log}_2\text{FC} > 1.5$) and adjusted P values ($P_{\text{adj}} < 0.05$). DESeq2 (v1.12.1) [20] was used to identify differentially expressed genes (DEGs) between high- and low-weight samples, with significance

determined at $P_{\text{adj}} \leq 0.05$. The significant DEGs were subsequently subjected to enrichment analysis via the web tool for Gene Ontology (GO) terms [21]. Functional enrichment analyses were conducted via the DAVID Bioinformatics Resources [22].

Metabolite detection and data analysis

To identify potential growth markers in water buffaloes, we employed a rigorous nontargeted LC-MS approach. The plasma samples (100 μL) were initially treated with cold methanol (200 μL), incubated at -20°C , and then centrifuged. The resulting supernatant (150 μL) was dried and reconstituted in H_2O : acetonitrile (50:50, v/v) with 0.1% formic acid for LC-MS analysis. Quality control (QC) samples were integrated throughout the system to ensure data reliability, monitor instrument performance and validate metabolite detection accuracy through segmental scanning. Chromatographic separation was performed with a Hypesil Gold C18 column with a gradient of 0.1% formic acid and methanol at a constant flow rate. The mass spectrometry conditions encompassed a wide mass range (m/z 100–1,500) in positive and negative ion modes, employing electrospray ionization. Compound Discoverer (v3.1) processed high-resolution MS data, facilitating peak alignment and extraction on the basis of defined parameters, and identification against spectral databases such as mzCloud, mzVault, and MassList. The identified metabolites were annotated via the KEGG and LIPIDMaps databases to link their molecular features with relevant biological pathways. metaX [23] conducted Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA), generating Variable Importance in Projection (VIP) scores. Differentially abundant metabolites between the high- and low-weight groups were identified on the basis of $\text{VIP} > 1$, $P < 0.05$, and fold change thresholds (≥ 2 or ≤ 0.5).

Microbial community and function profiling

Rumen metagenome DNA from HZ buffalo was extracted via the FastDNA[®] Spin Kit for Soil (MP Biomedicals) and fragmented via a Covaris ultrasonic crusher. Library preparation involved end repair, A-tailing, adapter ligation, PCR amplification, fragment screening, and purification. Sequencing was conducted on the BGISEQ DNBSEQ-T7 platform by Beijing Novogene Technology Co., Ltd. Quality trimming was performed via Fastp (v0.23.4) [24] to remove 3' and 5' end adapters, filter out reads shorter than 50 bp, reads with low average base quality scores (< 20), and those containing N bases. Reads were aligned to host DNA sequences via BWA-MEME (v1.0.6) [25] to filter out contaminant reads. MEGAHIT (v1.2.9) [26] was used for sequence assembly, generating contigs of ≥ 300 bp. Prodigal (v2.6.3) [27] was employed to predict open reading frames (ORFs) of ≥ 100 bp, which

were translated into amino acid sequences. The predicted genes were clustered into a nonredundant set via CD-HIT (v4.8.1) [28] with 90% identity and coverage, and the longest gene from each cluster was selected as the representative gene. SOAPaligner (v2.21) calculates gene abundance by aligning high-quality reads to the nonredundant gene set with a 95% similarity threshold. Amino acid sequences were aligned against the NR and KEGG databases via Diamond with an e-value threshold of $1e-5$ for species and functional annotation. Carbohydrate-active enzymes were annotated via hmmscan from the CAZy database. Analyses were conducted on the Majorbio Cloud Platform. Species alpha diversity was assessed via one-way ANOVA. Principal coordinate analysis (PCoA) was performed for intergroup analysis, and intergroup similarity was evaluated via ANOSIM with R QIIME software. Spearman correlation was employed to link differential species and functional genes with rumen carbohydrate-active enzyme activity.

Genome-wide alignment, variation detection and selection signature analysis

The large buffaloes consisted of 5 Haizi (HZ) and 5 Dehong (DH) buffaloes, whereas the small buffaloes included 6 Wenzhou (WZ) buffaloes and 5 Yibin (YB) buffaloes. The raw data used for selection signal analysis were downloaded from the China National Center for Bioinformatics website (<https://ngdc.cncb.ac.cn/gsa/browse/CRA001463>) [1]. After quality control, the sequenced reads were compared to the reference NDDDB_SH_1 (GCA_019923935.1) via BWA-MEM [29]. Genome-wide high-quality genetic variation was detected via GATK (v3.6) HaplotypeCaller and GenotypeGVCFs [30]. The GATK Variant Filtration command was employed with the parameter ‘-filter-expression “QD < 2.0 || MQ < 40.0 || FS > 60.0 || SOR > 3.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0” to enable efficient exploration of quality filters to exclude potential false-positive variant calls for mutation identification. Since we considered only single-nucleotide variation in this study, all called insertions and deletions (indels) were removed via VCFtools (0.1.17) [31] ‘-remove-indels’ parameter. The fixation index (Fst) was calculated via VCFtools [31] for both large buffaloes and small buffaloes. Fst was computed for the following pairs: HZ vs. WZ, HZ vs. YB, DH vs. WZ, and DH vs. YB. The sliding window size was set to 100 kb with a step size of 20 kb. The specific command is provided below: `--big_group--small_group --fst-window-size 100000 --fst-window-step 20000 --maf 0.05 --max-missing 0.90`. The genomic regions of selected sites were annotated via Snpeff (v3.9H).

Functional analysis of the *RPL26* gene

Homology comparisons of the *RPL26* gene across diverse species were performed via UniProt [32]. The gene name “*RPL26*” was entered, and the relevant species were selected by clicking “Align.” Custom colors were applied to the “percent identity matrix” based on the alignment results to visualize similarities and differences among the species. To construct the evolutionary tree of the *RPL26* gene, gene sequences from various species were downloaded from NCBI. These sequences were imported into MEGA (v11.0) [33], and alignment was performed via ClustalW with default parameters. The aligned sequences were then used to construct a phylogenetic tree via the “phylogeny-construct/text maximum likelihood tree” function in MEGA, with default settings generally applied. This method provides systematic insights into the homology and evolutionary relationships of the *RPL26* gene across different species. The promoter sequence of the *RPL26* gene was retrieved from NCBI, and predicted transcription factor-binding sites were identified through ConTra (v3) [34]. Linkage disequilibrium analysis within the promoter region and selected sites across the gene was conducted via LDblockshow (v1.40) [35]. LDBlockShow -InVCF buffalo_snp_filtered.Recode.vcf -OutPut 7k_Id -ShowNum -Region NC_059159.1:35121000:35128000 -OutPng -SeleVar 1, covering a region with a total of 56 selected SNPs.

Cell culture

C2C12 myoblasts, which exhibit vigorous growth after 48 ~ 72 h of passaging, were digested with 2.5 g/L trypsin solution to generate cell suspensions. The cell density was adjusted to 4×10^4 cells per milliliter, and the cells were seeded into 6-well culture plates. In a 6-well plate, each well received 2 mL of cell suspension, and the cells were cultured at 37 °C in a constant-temperature incubator for 24 h until they adhered to the surface. The cells were divided into two groups: the control group was subjected to normal culture conditions (DMEM, 10% fetal bovine serum and 1% penicillin–streptomycin solution (Gibco, Shanghai, China)), while the second group was subjected to siRNA-mediated knockdown. siRNAs targeting *RPL26* (si*RPL26*-1, si*RPL26*-2) (Table S6-1) were obtained from GenePharma Co., Ltd. (Shanghai, China). Each siRNA was transfected into the cells at a concentration of 5 nmol/L via CALNP™ RNAi in vitro (D-Nano Therapeutics, Beijing, China) following the manufacturer’s protocol. C2C12 cells were collected 48 h after transfection for further experiments. For the differentiation experiments, after 24 h of transfection, the medium was replaced with normal growth medium. When the cells reached approximately 90% confluence, the medium was changed to differentiation medium (DMEM + 2% HS), and the medium

was changed every two days. Experiments were conducted after 3 days of differentiation.

RT-qPCR

Total RNA was extracted from C2C12 cells via a high-purity total RNA extraction kit (Proteinssci, AF505B, Shanghai, China), and cDNA was synthesized via HiScript® III RT SuperMix for qPCR (+gDNA wiper) (Vazyme, R323-01, Nanjing, China). RT-qPCR amplification was performed via specific primers (Tsingke, Nanjing, China) and SYBR Green I dye (Vazyme, Q711, Nanjing, China) on a QuantStudio 5 system. The relative transcript abundance was determined via the $2^{-\Delta\Delta C_t}$ method and normalized to that of *GAPDH* (Table S6-2).

Western blot

Protein was collected after 3 days of differentiation in differentiation medium. To each well of a 6-well plate, 200 μ L of RIPA lysis buffer (with 10 μ L of PMSF) was added. The cells were lysed on ice for 10 min and then scraped off (if viscous, brief sonication was used). After denaturation at 98 °C for 15 min and cooling to room temperature, the lysate was centrifuged at 12,000 rpm for 10 min at 4 °C. The supernatant was aliquoted and stored at -80 °C (for long-term storage) or -20 °C. The collected protein samples were loaded onto polyacrylamide gels for electrophoresis, followed by transfer to membranes. The membranes were blocked with 5% skim milk for 2 h. The following primary antibodies were prepared in TBST at a 1:1000 dilution: RPL26 (Proteintech, Wuhan, China), MyoD1 (Proteintech, Wuhan, China), MyoG (Proteintech, Wuhan, China), and MYH1 (ABclonal, Wuhan, China). The membranes were incubated with the appropriate primary antibodies overnight at 4 °C. The membranes were then washed three times with TBST, each lasting 5 min. HRP-conjugated secondary antibodies (rabbit, mouse) were diluted 1:5000 in TBST and incubated at room temperature for 2 h. After the membranes were washed, enhanced chemiluminescence (ECL) reagent (1:1 mixture of solutions A and B) was applied for detection, and chemiluminescent imaging was performed. Protein bands were qualitatively assessed on the basis of marker size and quantitatively analyzed via ImageJ software.

Immunofluorescence

C2C12 cells were transfected with siRNA, and after 24 h, the medium was replaced with growth medium for recovery. The cells were digested with trypsin to form a cell suspension, which was evenly plated into 24-well plates. When the cells reached approximately 90% confluence, the medium was replaced with differentiation medium. After 4 days of differentiation, the medium was removed, and the cells were washed twice with PBS, each

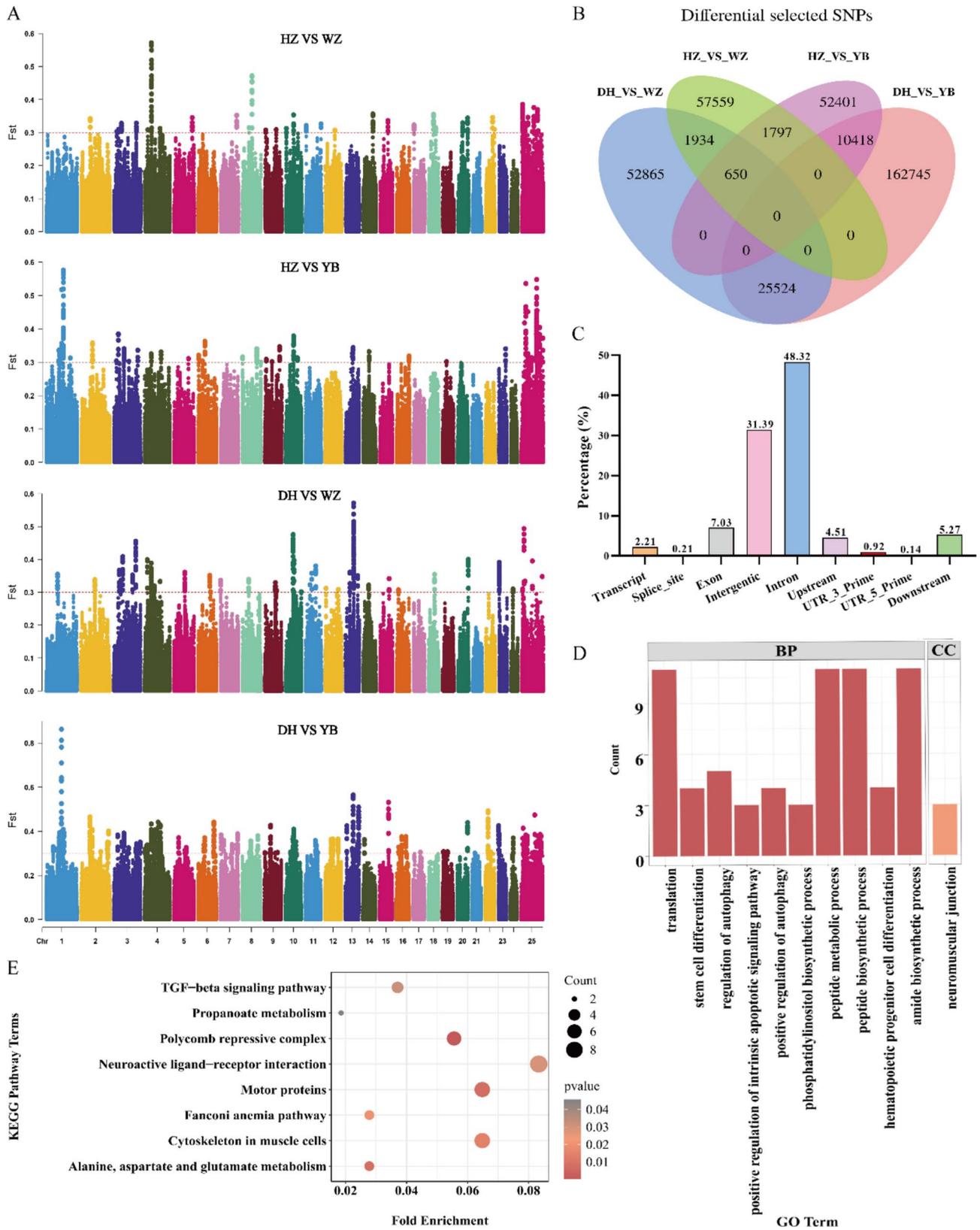


Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Visualization of whole-genome selection signal analysis in water buffaloes of different body sizes via public databases. **(A)** Manhattan plot comparing the genome-wide selected regions between water buffaloes of different body sizes. **(B)** Venn diagram of selected regions between water buffaloes of different body sizes. **(C)** Classification statistics bar chart of annotated genome positions for the selected sites after annotation. The x-axis represents different genomic position names, and the y-axis represents the percentage of sites. GO **(D)** and KEGG **(E)** enrichment analysis of genes annotated to the selected regions

lasting 5 min. After the PBS was removed, the cells were fixed with 4% paraformaldehyde at room temperature for 30 min and washed three times with PBS, each lasting 5 min. The cells were permeabilized with 0.5% Triton X-100 at room temperature for 20 min and washed three times with PBS. Blocking was performed with 5% BSA at room temperature for 60 min. The blocking solution was removed, and primary antibody dilution (MYH1, 1:100) was added overnight incubation at 4 °C. After the primary antibody was removed, the cells were washed three times with PBST, each lasting 5 min. The secondary antibody diluted with green fluorescence (1:200) was added and incubated at room temperature for 1 h in the dark. Hoechst 33,342 (1:1000) was used to stain the cell nuclei for 10 min, followed by washing with PBST. Fluorescence microscopy was used for observation and imaging.

Results

Genome-wide selection signatures

The genome-wide selection signature analysis was conducted for two buffalo species with large body size (HZ and Dehong (DH)) and two buffalo species with small body size (Wenzhou (WZ) and Yibin (YB)). The selected windows were primarily located on chromosome 3, 4, 8, and 25 (Fig. 1A). Under the selected windows for each buffalo species, 365,893 selected SNP positions were subsequently extracted to yield 40,323 shared SNP positions after pairwise comparisons (Fig. 1B). A total of 438 genes were identified within the selected regions and they were predominantly located in intron and intergenic regions (Fig. 1C). GO (Fig. 1D) and KEGG (Fig. 1E) enrichment analysis of these selected genes revealed their involvement in molecular functions such as peptide metabolic and biosynthetic process, motor proteins and cytoskeleton in muscle cells.

Identified differentially expressed genes and key metabolites

On average, 94.88% of clean reads were filtered from raw reads with 90.99% of Q30 base quality metrics (Table S1) and 23,703 and 23,216 genes were identified in blood and muscle, respectively. Using thresholds of $P_{\text{adj}} < 0.05$ and $|\log_2(\text{FoldChange})| > 2$, we identified 60 differentially expressed genes (DEGs) in blood with 4 upregulated and 56 downregulated genes (Fig. 2A and B, Table S2) and 18 DEGs in muscle with 11 upregulated and 7 downregulated genes (Fig. 2A and C, Table S2). The top three DEGs were *RPL26*, *CRISP2*, and *P2RX5* (Fig. 2D) and *RPL26*, *GON7*,

and *NPB* (Fig. 2E) of blood and muscle, respectively. Interestingly, *RPL26* was significantly downregulated in both blood and muscle tissues of the high-weight group, suggesting the crucial role of *RPL26* in regulating body growth traits.

A total of 851 metabolites were detected in positive ($n=563$) and negative ($n=288$) ion modes (Fig. 3A, Table S3). Using thresholds of P value < 0.05 , fold change < 1.5 , and VIP > 1 , 11 key metabolites were identified including 3 upregulated (myristicin, TKK, progesterone) and 8 downregulated (glycodeoxycholic acid, adenosine 5'-monophosphate, PC, LPE, L-ascorbate) metabolites in high-weight group compared with those in low-weight group (Figs. 2C and 3B). The DEG-metabolite correlation analysis found that *RPL26* was significantly negatively correlated with the metabolites TTK and progesterone but significantly positively correlated with the metabolites glycodeoxycholic acid, PC, and 8,15-dihete (Fig. 3D).

Pathway enrichments of the genes, metabolites and rumen microbiome

In blood, KEGG enrichment analysis of DEGs revealed nine significantly enriched pathways, such as the PPAR signaling, the nucleotide and purine metabolism, and the thiamine metabolism (Fig. 4A). In muscle, the significant pathways included the phospholipase D signaling, the pantothenate and CoA biosynthesis, and the amino sugar and nucleotide sugar metabolism (Fig. 4B). Pathway analysis of the metabolites further emphasized their key roles in lipid, nucleotide, and amino acid metabolism (Fig. 4C), closely mirroring the pathways enriched in the blood transcriptome.

The different microbes between two groups at the species level included *Kiritimatiellae* bacterium, *Butyricimonas virosa*, and *Clostridium* sp. (Fig. 5A). The microbial metabolism, the secondary metabolite biosynthesis, and the amino acid and cofactor biosynthesis were involved in the major KEGG enrichment of rumen microbiome pathways (Fig. 5B). Notably, the only differentially enriched pathway between the groups was nitrotoluene degradation (Fig. 5C), which implied the specific environmental or dietary influences on rumen microbial activity. Subsequent correlation analysis between species and functional abundance revealed that *Bacteroides* contributed mostly to metabolic pathways such as biosynthesis of amino acids, biosynthesis of cofactors, and carbon metabolism (Fig. 5D).

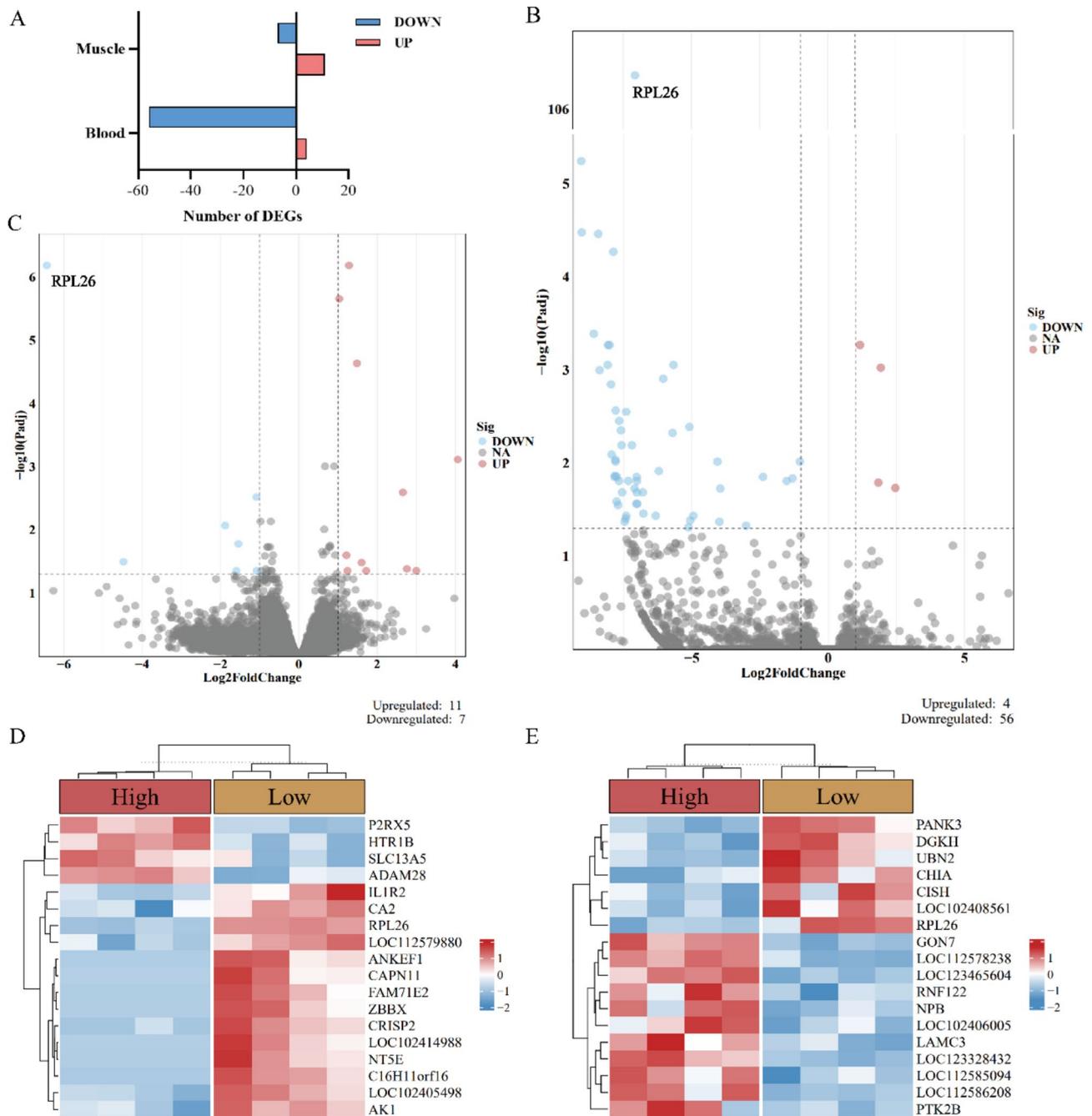


Fig. 2 DEGs in blood and muscle samples between high- and low-weight groups. **(A)** Number of identified differentially expressed genes (DEGs) in blood and muscle. Volcano plot of DEGs in blood **(B)** and muscle **(C)**. Gray (ns) indicates genes whose expression did not significantly differ. Clustering heat map of top 18 DEGs in blood **(D)** and muscle **(E)**. Red indicates high gene expression and blue indicates low gene expression

Integrated functional analysis of key gene RPL26

Analysis with ExPASy online tools (<https://web.expasy.org/protparam/>) revealed that *RPL26* in water buffaloes encodes 88 amino acids, with lysine (Lys) being the most abundant at 19.3%, followed by valine (Val) at 12.5% (Fig. 6A). On the basis of the amino acid sequence of *RPL26*, a phylogenetic tree was constructed via MEGA11, which revealed the closest evolutionary relationship of

water buffaloes with cattle and sheep (Fig. 6B). Homology analysis revealed approximately 98% similarity for *RPL26* among different species, with 98.86% homology with mammals such as bovines, sheep, goats, pigs, and mice (Fig. 6C), demonstrating its high degree of conservation across species.

We used the animal QTL database (Cattle QTL db) to explore whether the regulatory growth-related QTL

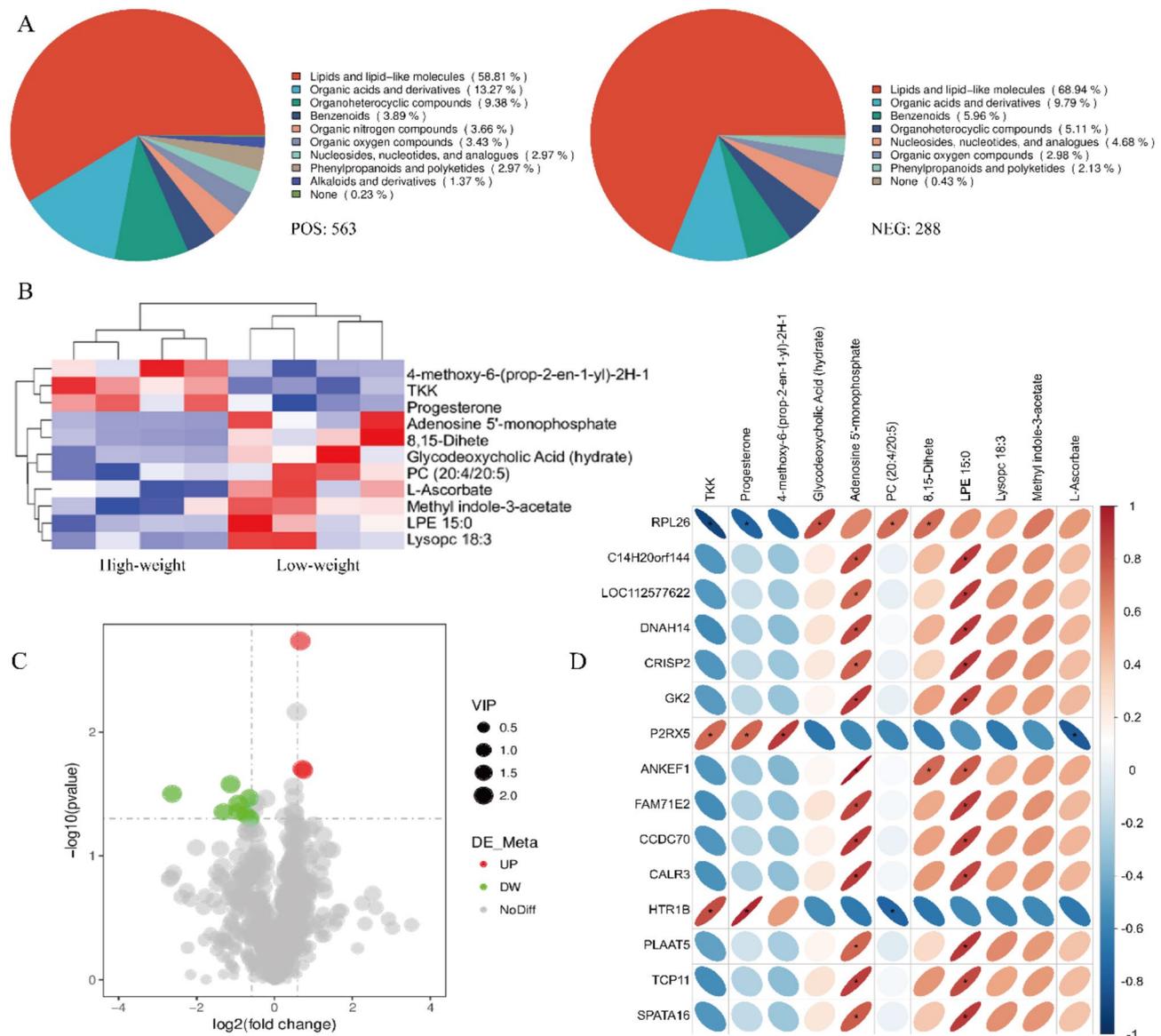


Fig. 3 Metabolites in plasma between high- and low-weight groups. **(A)** Summary pie chart of detected metabolites in positive and negative ion mode. **(B)** Cluster heatmap of differentially regulated metabolites. **(C)** Volcano plot of differentially regulated metabolites identified on the basis of set thresholds. Red represents upregulated metabolites in the high-weight group, green represents downregulated metabolites, and gray represents nonsignificant metabolites. **(D)** Pearson correlation analysis between differentially expressed genes and differentially abundant metabolites. The vertical axis represents genes, the horizontal axis represents metabolites, and red to blue represents positive to negative correlations

intervals are on *RPL26*. Considering the high amino acid sequence homology (98.86%) between bovine and water buffalo for *RPL26*, we aligned each selected SNP position on *RPL26* from buffalo to the reference genome of cattle (ARS-UCD1.2) (Table S5). The results revealed that the selected region on *RPL26* (primary assembly 19:27974414–27979325 reverse strand) overlapped with QTL_180979 in cattle, which was significantly correlated with daily weight gain ($P = 2.93 \times 10^{-9}$). The QTL peak is located at 28 Mb (50.38 cM, Chr19:27977049) on cattle chromosome 19 (Fig. S2). One QTL peak was identified in the significant block of the promoter region and gene

body of *RPL26* in buffalo (Fig. 6D). To further explore *RPL26*, we downloaded the promoter sequence of this gene from NCBI and predicted transcription factor-binding sites via ConTra (<http://bioit2.irc.ugent.be/contra/v3/>) and predicted the most related transcription factors like *MAX* (Fig. 6E), *MYCN* (Fig. 6F), *KLF4* (Fig. 6G), and *MEF2A* (Fig. 6H).

An extensive analysis utilizing the CattleGTEx database (<https://cgtex.roslin.ed.ac.uk/search/>) identified two cis-eQTLs across multiple tissues, situated 377.32 kb and 378.05 kb upstream of the gene body, which specifically modulate *RPL26* expression levels. (Fig. 7A, Fig. S3).

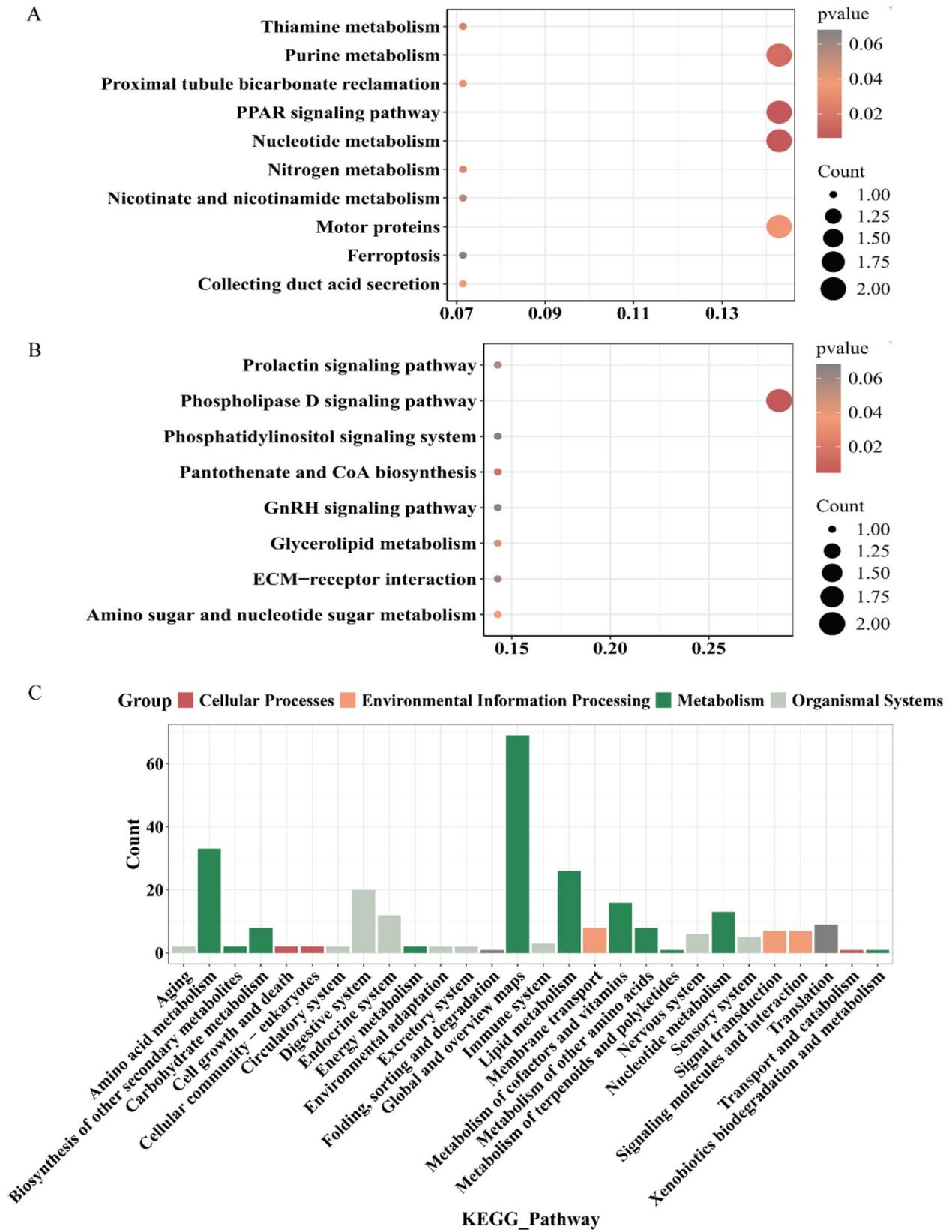
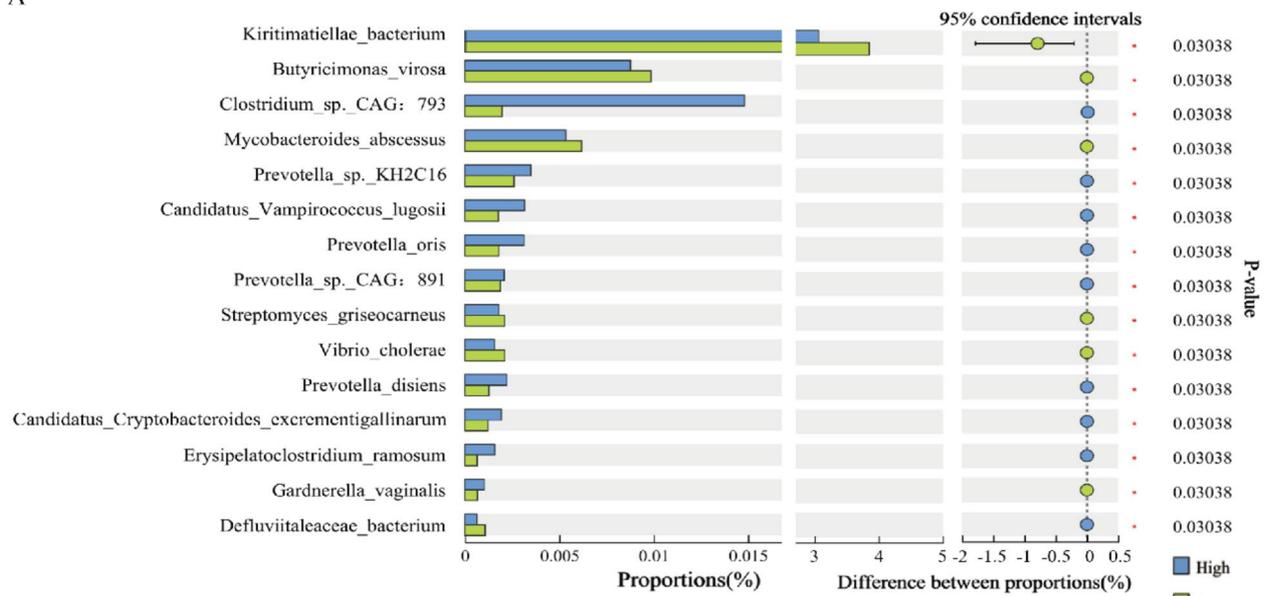
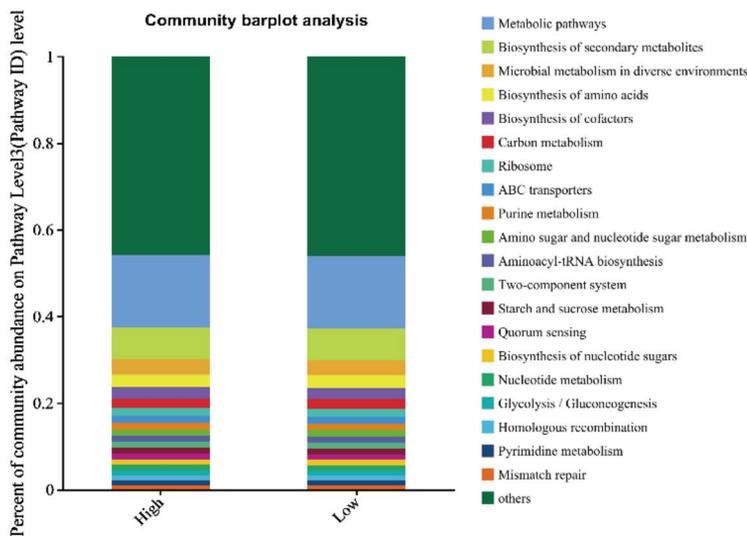


Fig. 4 Functional enrichment of the differential genes and metabolites. KEGG pathway enrichment of DEGs in blood (A) and muscles (B). The dot size represents the number of differentially expressed genes in that pathway, and the dot color represents the p value. (C) Functional enrichment bar chart of detected metabolites in positive and negative ion mode categorized into cellular process, environmental information processing, genetic information processing, metabolism, and organismal systems

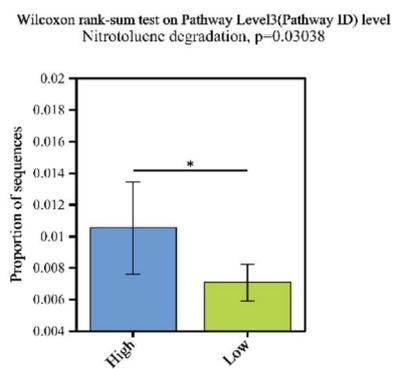
A



B



C



D

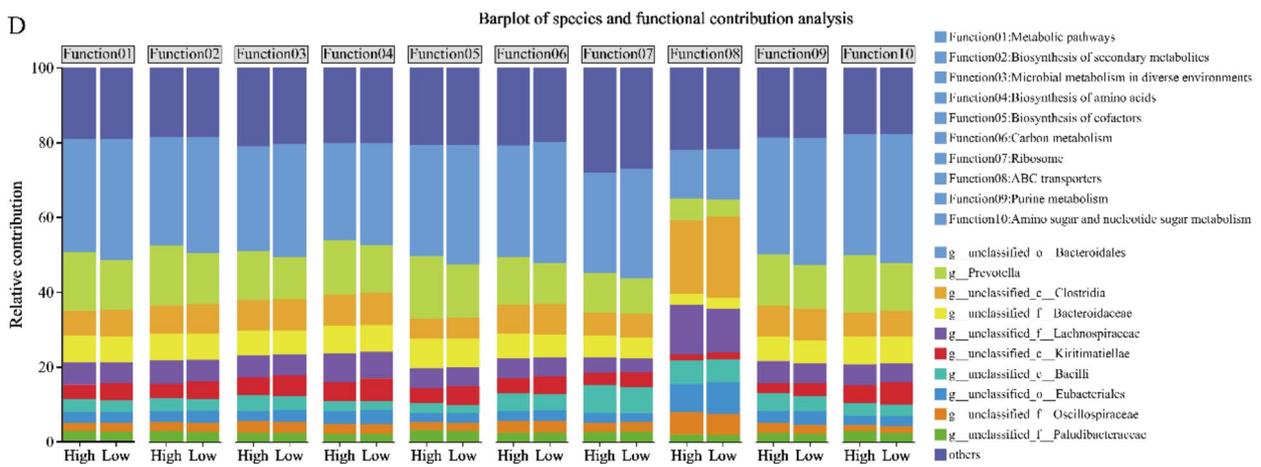


Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Rumen microbiota composition of the high- and low-weight groups. **(A)** Bar plot of the mean relative abundance differences of the same species between groups. The y-axis indicates the species names at different taxonomic levels, and the x-axis shows the percentage abundance of a species in the sample. **(B)** KEGG pathway functional composition and changes in the composition of the dominant pathways. **(C)** Analysis of the differential pathways associated with the microbiota/function composition. * $P < 0.05$ indicates a significant difference. ** $P < 0.01$ and *** $P < 0.001$ indicate very significant differences. **(D)** The bar chart shows the contributions of species to different functions at the species level. 1–10 represent 10 functional classifications, and 11 different colors represent the degree of species contribution under each functional classification

Based on UK Biobank release 2 data, the PheWAS results showed that *RPL26* gene was significantly correlated with height and standing height (Fig. 7B). These remarkable findings enhance our understanding of the crucial role of *RPL26* in regulating bovine growth (Fig. 7C).

Low expression of *RPL26* regulates cell cycle progression and enhances antiapoptotic ability

To further verify the functions of *RPL26* in cells, we introduced sequence-specific siRNAs against *RPL26* and conducted follow-up experiments in the C2C12 mouse myoblast cell line. Two siRNAs for *RPL26* were successfully transfected into C2C12 cells, resulting in high knockdown efficiency (Fig. 8A and B). The expression of *Bax* and *Bcl-2* was significantly downregulated after the knockdown of *RPL26*, as shown by the detection of apoptosis-related genes. The increase of *Bcl-2*/*Bax* ratio indicated the enhanced antiapoptotic ability (Fig. 8C). Moreover, the expression of the cell cycle-related gene *Cyclin D1* was significantly upregulated after knockdown, whereas *Cyclin D2* expression remained unchanged (Fig. 8D). These results suggest that *RPL26* knockdown can promote cell proliferation and increase the antiapoptotic ability of cells.

Low expression of *RPL26* promotes myogenic differentiation and fusion to form muscle ducts

We examined the protein expression of *RPL26* and found that the siRNAs significantly inhibited *RPL26* expression (Fig. 8E). We also investigated the expression levels of the key myogenic differentiation factors at both mRNA (Fig. 8F) and protein (Fig. 8G) levels. The expression levels of the myogenesis markers *MyoD1* and *MyoG* were significantly lower in the knockdown group than in the control group, whereas *MYH1* expression was significantly upregulated. Additionally, we employed immunofluorescence staining to examine the morphological changes in the myocytes after 3 days of differentiation (Fig. 8H). Immunofluorescence staining for *MYH1* revealed that in the MOCK and NC groups, the myocytes were round or spindle shaped, and no fusion events or myotube formation was observed. After 3 days of differentiation, a significant number of myocytes began to fuse, and immunofluorescence staining revealed the presence of *MYH1*-positive multinucleated myotubes, although most were binucleated at this stage (Fig. 8H, red arrows). Collectively, these results suggest that *RPL26*

knockdown promotes myogenic differentiation and myotube hypertrophy.

Discussion

Currently, many countries are transforming water buffaloes for labor into for dairy and meat products [36]. The body size of cattle plays a crucial role in their production, health, breeding selection, and environmental adaptation [37]. Body size encompasses a range of complex quantitative traits, such as body weight, body length, and height [38]. In this study, 438 genes were identified in the selection regions including *RPL26* and the expressions of *RPL26* in blood and muscle were significantly lower, providing crucial evidence for the significant role of *RPL26* in regulating buffalo body size. This finding aligns with a previous study on sheep body size, where Antonios et al. identified the top 5 genes associated with sheep body size, namely, *TP53*, *BMPRIA*, *PIK3R5*, *RPL26*, and *PRKDC*, through a combined GWAS and ‘guided by association’-based prioritization analysis [39]. The higher number of differentially expressed genes (DEGs) observed in blood compared to muscle can be attributed to several biological factors. Blood is a dynamic and systemic tissue that reflects metabolic, immune, and hormonal fluctuations, whereas muscle is a more structurally stable tissue with relatively conserved gene expression. Blood cells, particularly leukocytes, have a high turnover rate and are highly responsive to physiological changes, making it easier to detect transcriptional differences. Additionally, blood is a heterogeneous tissue comprising various cell types, leading to greater transcriptomic variability, while muscle tissue is more homogeneous, primarily consisting of terminally differentiated myofibers with fewer transient transcriptional changes. These factors collectively contribute to the greater number of DEGs identified in blood than in muscle.

By integrating publicly available animal QTL databases, we identified a QTL interval associated with daily weight gain in cattle located on the *RPL26* gene. Interestingly, this region showed high linkage with the gene’s promoter region, suggesting that SNP sites in the promoter region may exert vigorous promoter activity on the QTL interval. Aniek et al. [40] conducted a meta-analysis of GWAS that compared the genomes of cattle, humans, and dogs and reported a significant overlap in the loci controlling body size. Using CattleGTEx database [15], we discovered two cis-eQTLs on the *RPL26* gene in bovine

muscle tissue that influence gene expression. Interestingly, these two cis-eQTLs are located approximately 377 kb upstream of the *RPL26* gene (Fig. S3), specifically on another gene, *PIK3R5*, which is known to play crucial roles in cell growth, proliferation, differentiation, movement, survival, and intracellular transport [41, 42]. An interaction between *RPL26* and *PIK3R5* potentially influencing the growth and development of cattle could ultimately result in differences in body size (Fig. S4).

Currently, research on *RPL26* has mainly focused on endoplasmic reticulum homeostasis, cellular autophagy, and apoptosis [43]. The ubiquitin-like modification of *RPL26* regulates the translocation of nascent proteins to the endoplasmic reticulum, promoting the targeted degradation of stalled endoplasmic reticulum proteins to lysosomes [44, 45]. By maintaining the stability of ER proteins, *RPL26* plays a crucial role in ER protein biosynthesis, facilitating protein synthesis and participating in posttranslational modifications [45]. In our study, the knockdown of the *RPL26* gene in C2C12 cells via siRNA resulted in a significant decrease in the mRNA expression of *p53*, *bcl2*, and *Bax*, whereas *Cyclin D2* expression was significantly increased. Therefore, we speculate that *RPL26* inhibits cell apoptosis by affecting p53 transcriptional activation and reducing p53 protein expression, although the specific regulatory mechanisms require further investigation.

The p53 protein is a key regulator of cell growth, proliferation, and damage repair and plays a pivotal role in important biological processes of cell cycle regulation and apoptosis [46]. *RPL26* is also a key mediator of p53 signal responsive to ribosomal stress induced by ribosome biogenesis. When overexpressed or subjected to ribosomal stress, *RPL26* binds to MDM2 and inhibits MDM2-mediated p53 ubiquitination and proteasomal degradation, thereby stabilizing p53 and inducing cell cycle arrest. Alternatively, *RPL26* enhances the binding of *p53* mRNA to ribosomes and increases *p53* translation, inhibiting cell proliferation and inducing p53-dependent cell cycle arrest at the G1 phase, thereby increasing radiation-induced cell apoptosis [47].

This study revealed that the mRNA and protein levels of both *MyoD* and *MyoG* were significantly lower in the *RPL26*-knockdown group than in the control group. These findings suggest that *RPL26* knockdown inhibits the expression of these factors in the early stages of differentiation, promoting cell cycle exit and accelerating differentiation. *MyoD* is a muscle-specific transcription factor expressed in myoblasts during the late proliferation and early differentiation stages. It is generally believed that *MyoD* is upstream of *MyoG* in the differentiation process and that *MyoD* expression levels can reflect the early differentiation status of myoblasts [48, 49]. The myosin heavy chain (MYHC) subtypes, which

are composed of *MYH7*, *MYH2*, *MYH1*, and *MYH4*, are critical markers of myotube formation. In this study, *MYH1* was significantly upregulated in the MYH1-knockdown group, and the immunofluorescence results also visually confirmed the occurrence of fusion events and the formation of binucleated myotubes. Jin et al. [50] found that *MYH1* transgenic mice demonstrated greater endurance, running longer and farther on a treadmill than did wild-type mice. Given that, it is suggested that low *RPL26* expression may enhance endurance by upregulating *MYH1* expression.

Our study revealed that *RPL26* expression exhibits distinct correlations with specific metabolites—significantly negative correlations with TTK and progesterone, and positive correlations with glycodeoxycholic acid (GDCA) and phosphatidylcholine (PC). Specifically, studies on PGC-1 α —a transcriptional coactivator regulating muscle metabolism—have shown that steroid hormones like progesterone can suppress anabolic pathways, leading to muscle atrophy [51, 52]. Conversely, the positive correlation with GDCA, a bile acid involved in lipid metabolism, supports its role in enhancing energy availability for myocyte proliferation [53]. The link between PC and *RPL26* highlights the importance of membrane dynamics in myogenesis. PC, a major phospholipid component, contributes to membrane integrity during myoblast fusion and satellite cell activation [54]. Notably, PGC-1 α -mediated pathways have been shown to upregulate phospholipid biosynthesis [55], which supports membrane expansion in differentiating muscle cells. *RPL26* may enhance PC availability to stabilize mTOR signaling—a key pathway driving myogenic differentiation [56]. In summary, *RPL26* likely serves as a nexus integrating metabolic and transcriptional networks in muscle biology. Its dual role in suppressing catabolic signals (progesterone) while enhancing anabolic and regenerative pathways (via GDCA, PC) positions it as a potential therapeutic target for muscle-wasting disorders. Future studies should validate these mechanisms using genetic models (e.g., muscle-specific *RPL26* knockout) and explore cross-talk with established regulators like PGC-1 α [53].

Conclusion

In summary, our study identified top DEG gene *RPL26*, molecular biomarkers (myristicin and Bacteroidales) and microbiological composition (Bacteroides and Prevotella) associated with growth. Besides, One QTL (QTL_180979, Chr19: 27974414–27979325) and two cis-eQTLs related to growth traits are also located in the upstream of *RPL26*. The follow-up cell experiments validated the key role of *RPL26* in promoting muscle differentiation. Our results suggest the further applications of *RPL26* for growth trait selection of water buffaloes.

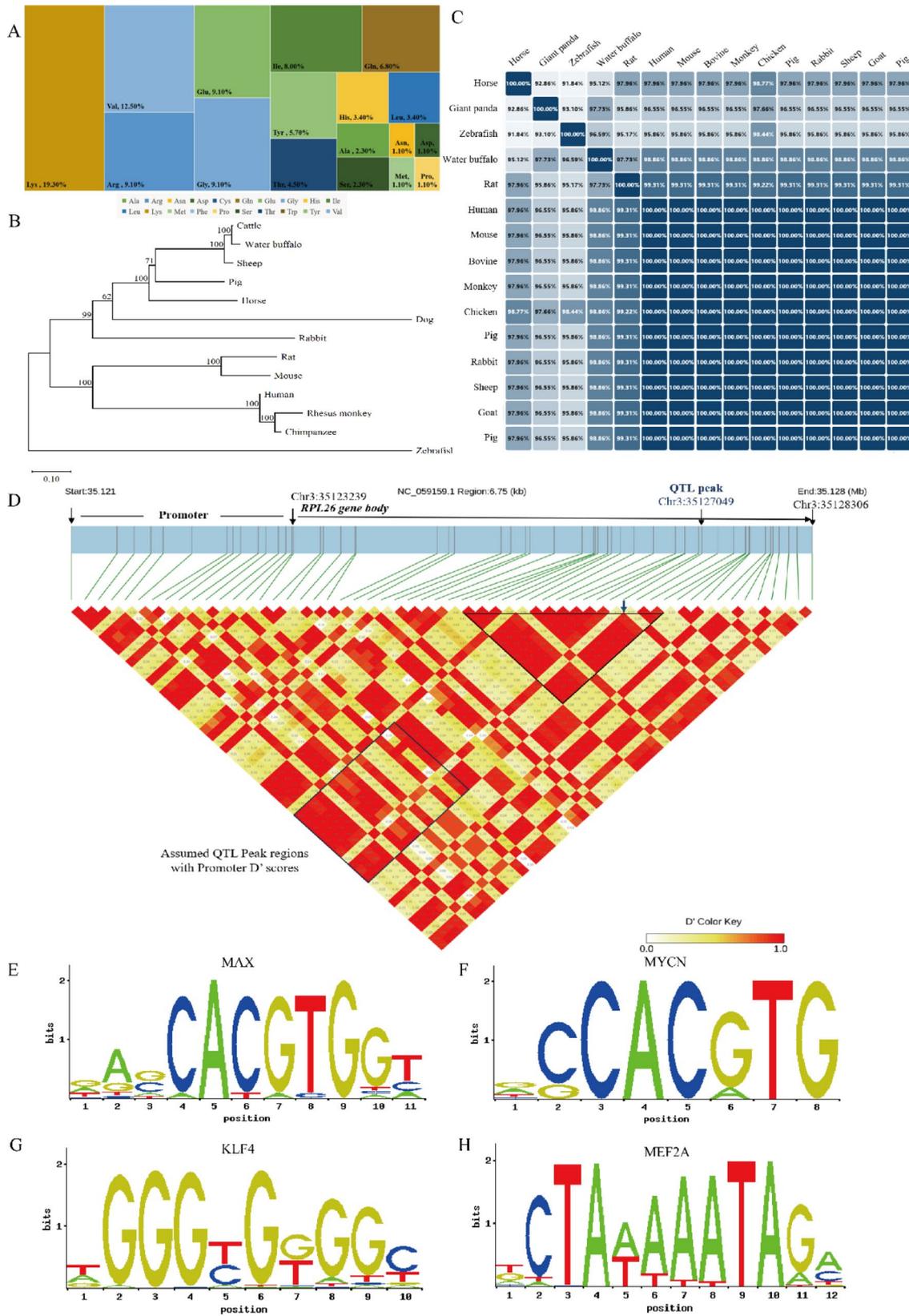


Fig. 6 Bioinformatics analysis of *RPL26*. **(A)** Amino acid composition of *RPL26*. **(B)** Evolutionary analysis of the amino acid sequences of *RPL26* in different species. **(C)** Homology analysis of *RPL26* in different species. **(D)** Linkage disequilibrium blocks detected in the *RPL26* promoter and gene body regions on buffalo. The SNPs in red boxes have the highest P values. The black boxes represent hypothesized highly linked blocks. **(E, F, G, H)** Predicted potential transcription factor-binding sites via the upstream 2 kb promoter sequence of *RPL26*

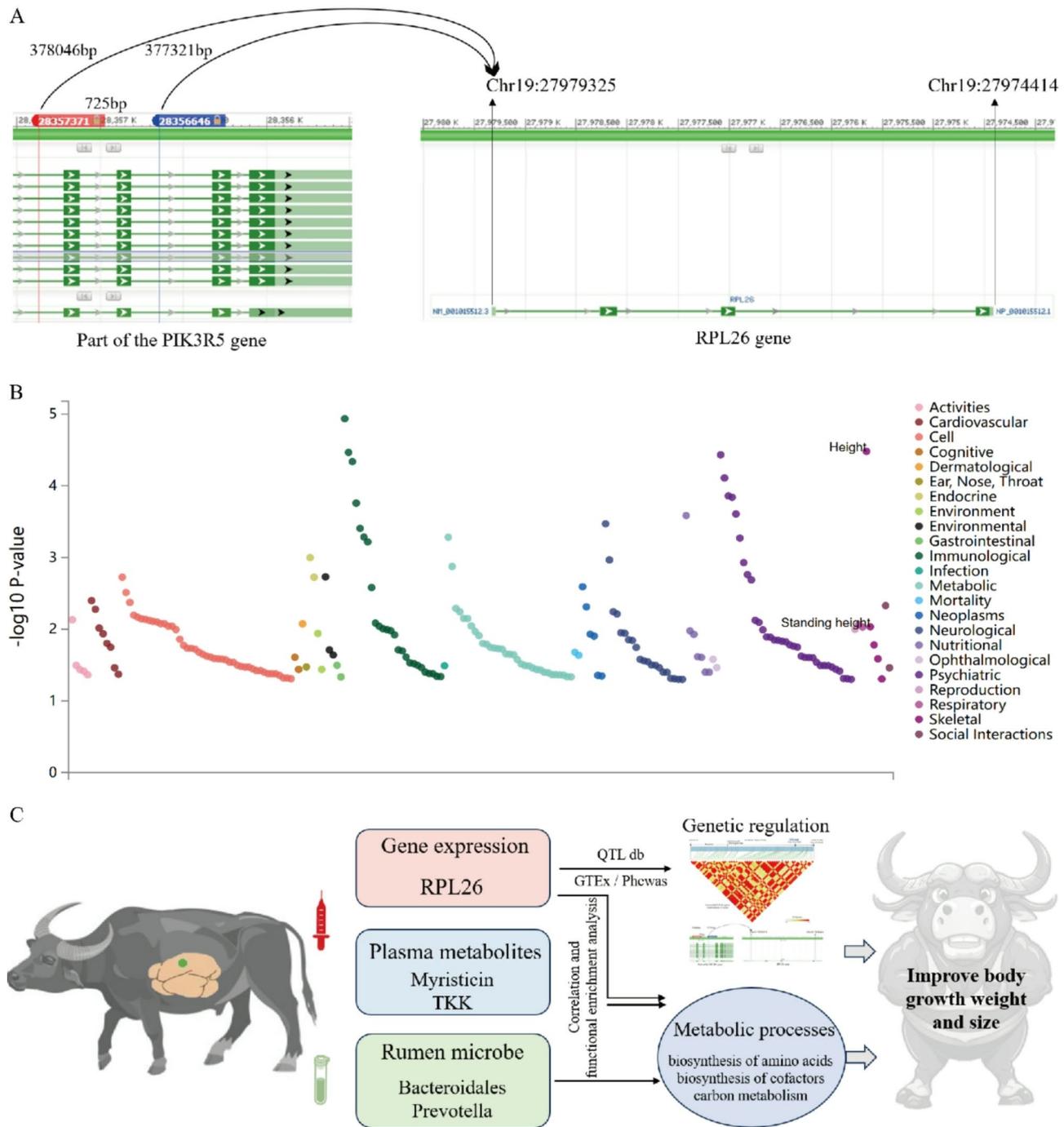


Fig. 7 Functional analysis of *RPL26*. **(A)** Location diagram of two cis-eQTLs on *RPL26*, which are located 377.32 kb and 378.05 kb upstream of the gene body. **(B)** PheWAS results of *RPL26* gene. Screenshot of query results from Genome wide association study ATLAS (<https://atlas.ctglab.nl/PheWAS>). **(C)** Integrating transcriptome-metabolome-metagenomic data revealed the central role of the key gene *RPL26* and molecular markers in regulating the growth and size of Haizi buffalo

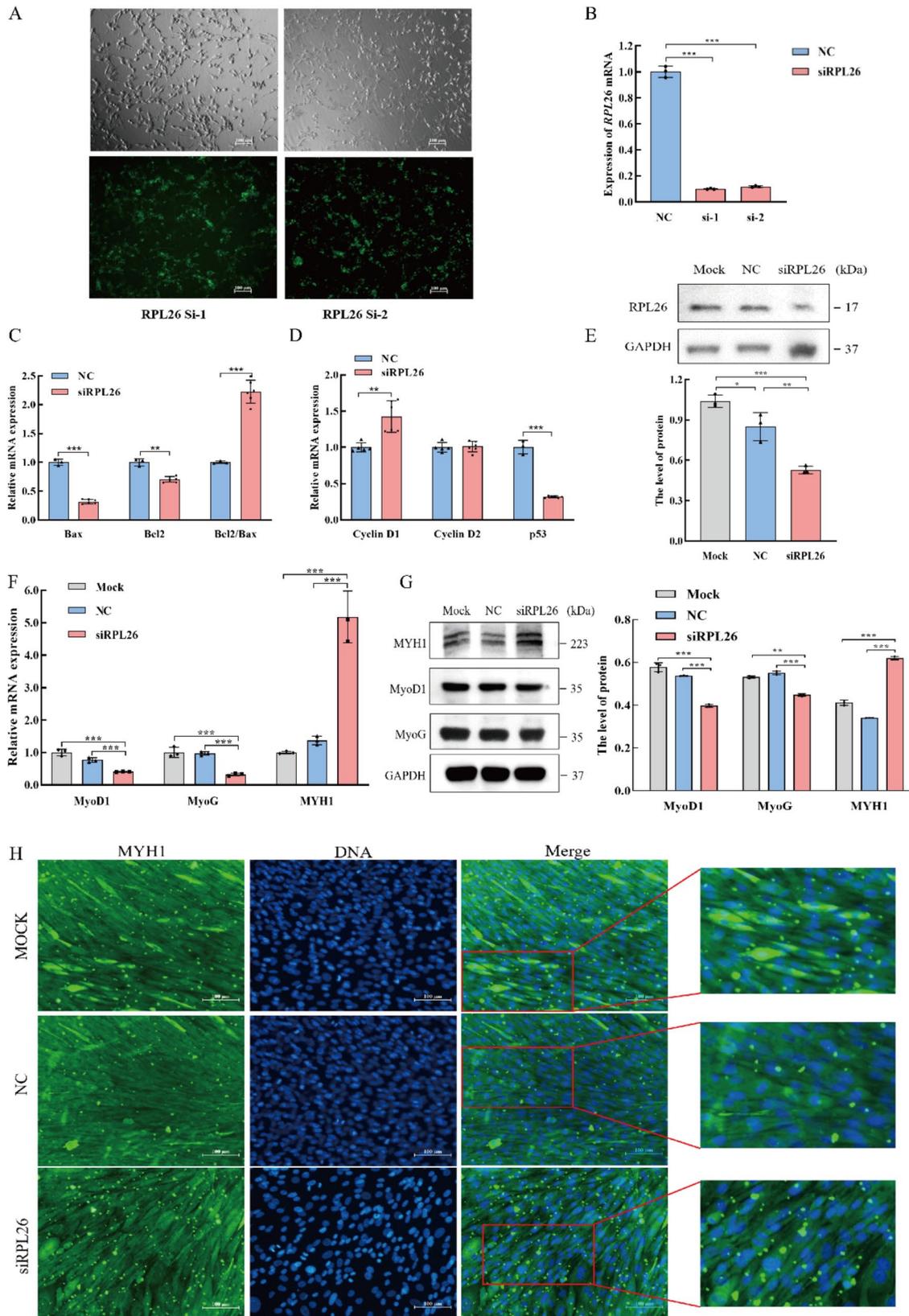


Fig. 8 (See legend on next page.)

(See figure on previous page.)

Fig. 8 Experimental verification of the *RPL26* gene in C2C12 cells. **(A)** Fluorescence images of C2C12 cells treated with different siRNAs. **(B)** Low-efficiency screening of *RPL26* gene siRNA-mediated knockdown and determination of the mRNA level of *RPL26*. **(C)** mRNA expression detection of Bax and Bcl2 after *RPL26* knockdown during the cell proliferation period. **(D)** mRNA expression detection of Cyclin D1, Cyclin D2 and p53 after *RPL26* knockdown during the cell proliferation period. **(E)** The protein level of *RPL26* knockdown. mRNA **(F)** and protein **(G)** expression detection of *MyoD1*, *MyoG* and *MYH1* after *RPL26* knockdown during the cell differentiation period. **(H)** Immunofluorescence image of the myotubal formation status after C2C12 cell differentiation for 3 days under *RPL26* knockdown. DNA was used to stain the nucleus, and the nucleus was blue. *MYH1* is shown in green. Merge is a synthesis diagram of the cytoplasm and nucleus. The scale bar “100 μm ” is indicated by a white straight line. $^{**}P < 0.01$, $^{***}P < 0.001$, both indicate very significant differences

Abbreviations

RPL26	Ribosomal protein L26
DEGs	Differentially expressed genes
QTL	Quantitative trait loci
eQTL	Expression quantitative trait loci
CattleGTEx	Genotype-tissue expression
HZ	Haizi water buffaloes
DH	Dehong water buffaloes
WZ	Wenzhou water buffaloes
YB	Yibin water buffaloes
GWAS	Genome-wide association study
FPKM	Fragments per kilobase million
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
PCA	Principal component analysis
PLS-DA	Partial least squares discriminant analysis
VIP	Variable importance in projection
MYHC	Myosin heavy chain

ngdc.cncb.ac.cn/gsub/submit/gsa/subCRA026044. Muscle transcriptome: <https://ngdc.cncb.ac.cn/gsub/submit/gsa/subCRA026411>. The metagenome sequencing data are under accession code CRA016362 (<https://ngdc.cncb.ac.cn/gsub/submit/gsa/subCRA026231>). A dataset of the raw plasma metabolome of HZ buffaloes was deposited into Figshare with the accession number 26161633. The WGS data of DH, YB, HZ and WZ buffaloes were downloaded from the genome sequence archive (GSA) under accession code CRA001463 (<https://ngdc.cncb.ac.cn/gsa/browse/CRA001463>).

Declarations

Ethics approval and consent to participate

All procedures involving tissue sample collection and animal care were performed according to the approved protocols and ARRIVE (Animal Research: Reporting In Vivo Experiments) guidelines and were approved by the Ethics Committee of Jiangsu Academy of Agriculture Science. The experimental procedures were approved by the Research Committee of the Jiangsu Academy of Agricultural Sciences and conducted in accordance with the Regulations for the Administration of Affairs Concerning Experimental Animals (Decree No. 63 of the Jiangsu Academy of Agricultural Science on 8 July 2014).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 19 November 2024 / Accepted: 21 April 2025

Published online: 08 May 2025

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11618-6>.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5
Supplementary Material 6
Supplementary Material 7

Acknowledgements

We thank Dr. Kai Shi (Ningbo Academy of Agricultural Sciences) for the helpful discussion of this manuscript. We thank Cassie Zhao (GemPharmatech Co., Ltd.) for improving the English of this manuscript. We acknowledge the China National Center for Bioinformatics for data sharing and GitHub for the platform used.

Author contributions

Y.Y.S. and Z.J.A. analyzed and interpreted the population genetic evolution part. S.W.X. performed the selective sweep analysis. Y.Y.S. and Z.J.A. prepared Figs. and L.N.G. in charge of the validation part. Q.D., Y.X.L., S.X.C., J.B.L., J.M.H., J.F.Z., K.L.C., X.W. and H.L.W. participated in the analysis of the results. S.W.X., Q.D., K.L.C., and H.L.W. collected the samples. M.F.G., X.W. and H.L.W. revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by grants from the National Key R & D Program of China (2021YFD1200404), Shandong Provincial Natural Science Foundation (ZR2023QC252) and Jiangsu Funding Program for Excellent Postdoctoral Talent (2024ZB420).

Data availability

The RNA-sequencing reads of HZ buffaloes have been deposited into the GSA in the China National Center for Bioinformatics (CNCB) under accession codes CRA016358 (Blood) and CRA016473 (Muscle). Blood transcriptome: <https://>

References

- Luo X, Zhou Y, Zhang B, Zhang Y, Wang X, Feng T, Li Z, Cui K, Wang Z, Luo C, et al. Understanding divergent domestication traits from the whole-genome sequencing of swamp- and river-buffalo populations. *Natl Sci Rev*. 2020;7(3):686–701.
- Joshi P, Gowane GR, Alex R, Gupta ID, Worku D, George L, Ranjan A, Verma A. Estimation of genetic parameters of growth traits for direct and maternal effects in Murrah Buffalo. *Trop Anim Health Prod*. 2022;54(6):352.
- George L, Alex R, Sukhija N, Jaglan K, Vohra V, Kumar R, Verma A. Genetic improvement of economic traits in Murrah Buffalo using significant SNPs from genome-wide association study. *Trop Anim Health Prod*. 2023;55(3):199.
- Silva-Vignato B, Coutinho LL, Cesar ASM, Poleti MD, Regitano LCA, Balieiro JCC. Comparative muscle transcriptome associated with carcass traits of Nel-lore cattle. *BMC Genomics*. 2017;18(1):506.
- Purfield DC, Evans RD, Berry DP. Reaffirmation of known major genes and the identification of novel candidate genes associated with carcass-related metrics based on whole genome sequence within a large multi-breed cattle population. *BMC Genomics*. 2019;20(1):720.
- Bhuiyan MSA, Kim HJ, Lee DH, Lee SH, Cho SH, Yang BS, Kim SD, Lee SH. Genetic parameters of carcass and meat quality traits in different muscles (longissimus dorsi and semimembranosus) of Hanwoo (Korean cattle). *J Anim Sci*. 2017;95(8):3359–69.
- Qanbari S, Pausch H, Jansen S, Somel M, Strom TM, Fries R, Nielsen R, Simianer H. Classic selective sweeps revealed by massive sequencing in cattle. *PLoS Genet*. 2014;10(2):e1004148.
- Petersen JL, Mickelson JR, Rendahl AK, Valberg SJ, Andersson LS, Axelsson J, Bailey E, Bannasch D, Binns MM, Borges AS, et al. Genome-wide analysis reveals selection for important traits in domestic horse breeds. *PLoS Genet*. 2013;9(1):e1003211.

9. Xu L, Bickhart DM, Cole JB, Schroeder SG, Song J, Tassell CP, Sonstegard TS, Liu GE. Genomic signatures reveal new evidences for selection of important traits in domestic cattle. *Mol Biol Evol.* 2015;32(3):711–25.
10. Vy HM, Kim Y. A Composite-Likelihood method for detecting incomplete selective sweep from population genomic data. *Genetics.* 2015;200(2):633–49.
11. Hassan FU, Nawaz A, Rehman MS, Ali MA, Dilshad SMR, Yang C. Prospects of HSP70 as a genetic marker for thermo-tolerance and immuno-modulation in animals under climate change scenario. *Anim Nutr.* 2019;5(4):340–50.
12. Miao Y, Zhao Y, Wan S, Mei Q, Wang H, Fu C, Li X, Zhao S, Xu X, Xiang T. Integrated analysis of genome-wide association studies and 3D epigenomic characteristics reveal the BMP2 gene regulating loin muscle depth in Yorkshire pigs. *PLoS Genet.* 2023;19(6):e1010820.
13. Zhang H, Mi S, Brito LF, Hu L, Wang L, Ma L, Xu Q, Guo G, Yu Y, Wang Y. Genomic and transcriptomic analyses enable the identification of important genes associated with subcutaneous fat deposition in Holstein cows. *J Genet Genomics.* 2023;50(6):385–97.
14. Fonseca PAS, Id-Lahoucine S, Reverter A, Medrano JF, Fortes MS, Casellas J, Miglior F, Brito L, Carvalho MRS, Schenkel FS, et al. Combining multi-OMICs information to identify key-regulator genes for pleiotropic effect on fertility and production traits in beef cattle. *PLoS ONE.* 2018;13(10):e0205295.
15. Liu S, Gao Y, Canela-Xandri O, Wang S, Yu Y, Cai W, Li B, Xiang R, Chamberlain AJ, Pairo-Castineira E, et al. A multi-tissue atlas of regulatory variants in cattle. *Nat Genet.* 2022;54(9):1438–47.
16. Jiang J, Cole JB, Freebern E, Da Y, VanRaden PM, Ma L. Functional annotation and bayesian fine-mapping reveals candidate genes for important agronomic traits in Holstein bulls. *Commun Biology.* 2019;2:212.
17. Fang L, Liu S, Liu M, Kang X, Lin S, Li B, Connor EE, Baldwin RLT, Tenesa A, Ma L, et al. Functional annotation of the cattle genome through systematic discovery and characterization of chromatin States and butyrate-induced variations. *BMC Biol.* 2019;17(1):68.
18. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37(8):907–15.
19. Liao Y, Smyth GK, Shi W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923–30.
20. Love MI, Huber W, Anders S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
21. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.* 2007;35(Web Server issue):W193–200.
22. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* 2003;4(5):P3.
23. Wen B, Mei Z, Zeng C, Liu S. MetaX: a flexible and comprehensive software for processing metabolomics data. *BMC Bioinformatics.* 2017;18(1):183.
24. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34(17):i884–90.
25. Jung Y, Han D. BWA-MEME: BWA-MEM emulated with a machine learning approach. *Bioinformatics.* 2022;38(9):2404–13.
26. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics.* 2015;31(10):1674–6.
27. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics.* 2010;11:119.
28. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics.* 2012;28(23):3150–2.
29. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754–60.
30. Nekrutenko A, Taylor J. Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. *Nat Rev Genet.* 2012;13(9):667–72.
31. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. The variant call format and vcfutils. *Bioinformatics.* 2011;27(15):2156–8.
32. Pundir S, Martin MJ, O'Donovan C, UniProt C. UniProt tools. *Curr Protocols Bioinf* 2016, 53:1 29 21–21 29 15.
33. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 2018;35(6):1547–9.
34. Kreft L, Soete A, Hulpiu P, Botzki A, Saeys Y, De Bleser P. ConTra v3: a tool to identify transcription factor binding sites across species, update 2017. *Nucleic Acids Res.* 2017;45(W1):W490–4.
35. Dong SS, He WM, Ji JJ, Zhang C, Guo Y, Yang TL. LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Brief Bioinf.* 2021;22(4):bbaa227.
36. Greenwood PL. Review: an overview of beef production from pasture and feedlot globally, as demand for beef and the need for sustainable practices increase. *Animal.* 2021;15(Suppl 1):100295.
37. Chen Q, Huang B, Zhan J, Wang J, Qu K, Zhang F, Shen J, Jia P, Ning Q, Zhang J, et al. Whole-genome analyses identify loci and selective signals associated with body size in cattle. *J Anim Sci.* 2020;98(3):skaa068.
38. Maulana MI, Riksen JAG, Snoek BL, Kammenga JE, Sterken MG. The genetic architecture underlying body-size traits plasticity over different temperatures and developmental stages in *Caenorhabditis elegans*. *Heredity (Edinb).* 2022;128(5):313–24.
39. Kominakis A, Hager-Theodorides AL, Zoidis E, Saridakis A, Antonakos G, Tsiamis G. Combined GWAS and 'guilt by association'-based prioritization analysis identifies functional candidate genes for body size in sheep. *Genet Selection Evol.* 2017;49(1):41.
40. Bouwman AC, Daetwyler HD, Chamberlain AJ, Ponce CH, Sargolzaei M, Schenkel FS, Sahana G, Gogvignon-Gion A, Boitard S, Dolezal M, et al. Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. *Nat Genet.* 2018;50(3):362–7.
41. Quintanilha JCF, Racioppi A, Wang J, Etheridge AS, Denning S, Pena CE, Skol AD, Crona DJ, Lin D, Innocenti F. PIK3R5 genetic predictors of hypertension induced by VEGF-pathway inhibitors. *Pharmacogenomics J.* 2022;22(1):82–8.
42. Wang X, Liang Z, Liu Q, Ye X, Wu X, Deng C, Zhao L, Lu C, Qiu Z, Yao Y, et al. Identification of PIK3R5 as a hub in septic myocardial injury and the cardioprotective effects of psoralidin. *Phytomedicine.* 2024;122:155146.
43. Scavone F, Gumbin SC, Da Rosa PA, Kopito RR. RPL26/uL24 ufmylation is essential for ribosome-associated quality control at the Endoplasmic reticulum. *Proc Natl Acad Sci USA.* 2023;120(16):e2220340120.
44. Liang JR, Lingeman E, Luong T, Ahmed S, Muhar M, Nguyen T, Olzmann JA, Corn JE. A Genome-wide ER-phagy screen highlights key roles of mitochondrial metabolism and ER-Resident ufmylation. *Cell.* 2020;180(6):1160–77.
45. Gerakis Y, Quintero M, Li H, Hetz C. The ufmylation system in proteostasis and beyond. *Trends Cell Biol.* 2019;29(12):974–86.
46. Levine AJ. p53: 800 million years of evolution and 40 years of discovery. *Nat Rev Cancer.* 2020;20(8):471–80.
47. Towers CG, Guarnieri AL, Micalizzi DS, Harrell JC, Gillen AE, Kim J, Wang CA, Oliphant MUJ, Drasin DJ, Guney MA, et al. The Six1 oncoprotein downregulates p53 via concomitant regulation of RPL26 and microRNA-27a-3p. *Nat Commun.* 2015;6:10077.
48. Xie K, Fu Z, Li H, Gu X, Cai Z, Xu P, Cui X, You L, Wang X, Zhu L, et al. High folate intake contributes to the risk of large for gestational age birth and obesity in male offspring. *J Cell Physiol.* 2018;233(12):9383–9.
49. Blais A, Tsikitis M, Acosta-Alvear D, Sharan R, Kluger Y, Dynlacht BD. An initial blueprint for myogenic differentiation. *Genes Dev.* 2005;19(5):553–69.
50. Ahn JS, Kim DH, Park HB, Han SH, Hwang S, Cho IC, Lee JW. Ectopic over-expression of Porcine Myh1 increased in slow muscle fibers and enhanced endurance exercise in Transgenic mice. *Int J Mol Sci.* 2018;19(10):2959.
51. Halling JF, Pilegaard H. PGC-1alpha-mediated regulation of mitochondrial function and physiological implications. *Appl Physiol Nutr Metab.* 2020;45(9):927–36.
52. Duan H, Chen S, Mai X, Fu L, Huang L, Xiao L, Liao M, Chen H, Liu G, Xie L. Low-intensity pulsed ultrasound (LIPUS) promotes skeletal muscle regeneration by regulating PGC-1alpha/AMPK/GLUT4 pathways in satellite cells/myoblasts. *Cell Signal.* 2024;117:111097.
53. Yao X, Mai X, Tian Y, Liu Y, Jin G, Li Z, Chen S, Dai X, Huang L, Fan Z et al. Skeletal muscle-specific Bambi deletion induces hypertrophy and oxidative switching coupling with adipocyte thermogenesis against metabolic disorders. *Sci China Life Sci* 2025.
54. Ding K, Zhang L, Fan X, Guo X, Liu X, Yang H. The effect of pedal Peptide-Type neuropeptide on locomotor behavior and muscle physiology in the sea cucumber *Apostichopus japonicus*. *Front Physiol.* 2020;11:559348.

55. Senoo N, Miyoshi N, Goto-Inoue N, Minami K, Yoshimura R, Morita A, Sawada N, Matsuda J, Ogawa Y, Setou M, et al. PGC-1alpha-mediated changes in phospholipid profiles of exercise-trained skeletal muscle. *J Lipid Res.* 2015;56(12):2286–96.
56. Saxton RA, Sabatini DM. mTOR signaling in growth, metabolism, and disease. *Cell.* 2017;168(6):960–76.

Publisher's note

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