Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genome-wide association analysis of growth traits in Tianzhu white yak

Guowu Yang
Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Xuedong Qi
Tianzhu Autonomous County Animal Husbandry Technology Extension Station

Yongfu La
Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Xiaoming Ma
Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Yicheng Liu
Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Zhicheng Wang
Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Wenxue Luo
Tianzhu Autonomous County Animal Husbandry Technology Extension Station

Wenwen Ren
Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Zhenyu Zhang
Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Min Chu
Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Xiaoyun Wu
Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Xian Guo

Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

Wanzhen Qi

Tianzhu Autonomous County Animal Husbandry Technology Extension Station

Chunnian Liang

liangchunnian@caas.cn

Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Key Laboratory of Yak Breeding Engineering of Gansu Province

---

Research Article

Keywords: Tianzhu White Yak, GWAS, SNP, Growth characteristics

Posted Date: July 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9809492/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Genomics on October 14th, 2026. See the published version at https://doi.org/10.1186/s12864-026-13444-w.

Abstract

Background

Identifying key markers or genes associated with growth traits in Tianzhu White Yak is crucial for breeding. Precisely locating genetic factors related to body conformation and growth can improve breeding efficiency and economic value. In this study, low-coverage sequencing was used to resequencing 536 Tianzhu White Yak individuals at different growth stages (1 year, 2.5 years, and 4 years or older) for four growth trait indices: body weight (BW), body height (BH), body oblique length (BOL), and chest circumference (CC). Genome-wide association studies (GWAS) were conducted to identify candidate genes associated with these four body measurements, thereby laying the foundation for subsequent studies on the biological functions of these genes.

Results

After genotype imputation and quality control, a total of 24,383,269 SNPs were retained for subsequent analysis. A total of 366 significant SNP loci were identified, which were annotated to 32 candidate genes associated with body conformation traits, such as *PLCB1*, *SAR1B*, and *SRSF6* for body weight; *BRPF3*, *MAPK13*, and *ZNF76* for body height; *ABCD3* and *NTM* for body oblique length; and *LMNTD1* and *ACVR1* for chest circumference. Functional analysis revealed that multiple candidate genes are involved in key biological processes, including negative regulation of the G1/S transition of the mitotic cell cycle and fatty acid beta-oxidation. Notably, *TEAD3*, *ZNF76*, and *RBL1* were significantly enriched in the transcription regulator complex, a cellular component that may influence cell proliferation, differentiation, and tissue formation.

Conclusion

This study identified candidate loci and genes associated with growth traits in Tianzhu White Yak through low-coverage sequencing combined with GWAS analysis. These findings provide valuable genetic markers for genomic selection in Tianzhu White Yak, thereby offering resources to advance precision breeding programs.

Introduction

The Tianzhu white yak (*Bos grunniens*) is a rare, semi-wild yak population endemic to the Qinghai-Tibet Plateau, inhabiting the frigid alpine grasslands of Tianzhu County, Gansu Province, at elevations ranging from 2,500 to 3,400 meters. Located at the eastern end of the Qilian Mountains, where the Qinghai-Tibet, Loess, and Inner Mongolia plateaus converge, the region features complex terrain and significant vertical climatic variations. With an annual average temperature of just 0.5°C and lows reaching − 30°C, the dry season and frozen period last for about half a year, creating a harsh yet unique ecological environment.

Renowned as the “Ship of the Plateau” and the “Ship of the Snowlands” for its pure white coat, the Tianzhu White Yak is revered in Tibetan culture as the incarnation of the mountain god, embodying thousands of years of cultural memory and ethnic sentiment [1]. Through prolonged natural selection and artificial breeding, the Tianzhu White Yak has evolved into a distinctive and precious local breed among the yak populations of the Qinghai-Tibet Plateau. As a unique local breed in the Tianzhu region of Gansu Province, the Tianzhu White Yak population is relatively large and possesses stable genetic characteristics. Tianzhu White Yaks have a high meat yield and tender meat, characterized by being rich in nutrients, high in protein, and low in fat, making them highly favored by consumers. The Tianzhu White Yak is a local breed within the Qinghai-Tibet Plateau yak group. Having long lived under natural conditions of extreme cold, low oxygen levels, and a dry season lasting approximately six months, it has a smaller body size and weight, and its milk and meat production performance is lower than that of other yak breeds and ordinary cattle. Studies have shown that Qinghai-Tibet Plateau yak is characterized by low body weight, slow growth rates, and an early growth inflection point [2]. Therefore, improving the growth rate and meat yield of Tianzhu White Yak is a key priority that urgently needs to be addressed in the conservation and utilization of this breed.

Somatic traits (such as body weight, body height, body oblique length, and chest circumference) are important economic traits in animal production, directly influencing farming profitability and breed improvement strategies [3]. In livestock breeding, these somatic traits are typically complex traits controlled by polygenic factors with small effects and are significantly influenced by environmental factors, exhibiting high genetic complexity. In recent years, with the advancement of high-throughput genotyping technologies, genome-wide association studies (GWAS) have become a vital tool for elucidating the genetic basis of complex traits [4]. By scanning single-nucleotide polymorphism (SNP) markers across the entire genome, GWAS can identify candidate genes and genomic regions significantly associated with target traits, providing a theoretical basis for marker-assisted selection (MAS) and precision breeding [5, 6]. Initially, researchers used GWAS to identify variant loci associated with complex diseases [7], thereby pioneering the demonstration of GWAS’s effectiveness in identifying genetic factors underlying complex diseases. Across various animal populations, GWAS not only aids in deciphering the genetic architecture of body conformation traits but also reveals their potential association networks with traits such as growth, metabolism, and immunity. In major livestock species such as cattle [8, 9], sheep [10, 11], and pigs [12–14], GWAS has identified multiple key genes influencing growth rate and body conformation, providing a molecular basis for breed optimization. For example, researchers conducted a genome-wide association analysis of seven body conformation traits in Longling yellow goats, identifying multiple candidate genes—including *MMP16*, *MECOM*, *SCD5*, *LEF1*, *KAT6A*, *PDE4D*, *KCND2*, and *EPHA5*—associated with early growth and body conformation traits in Longling yellow goats [11]. In Nelore cattle, the *TMEM68* and *XKR4* genes were found to be associated with birth weight, and the *NPR2* gene was found to be associated with height at the withers [15].

To date, GWAS has been used to identify candidate genes associated with traits such as growth, immunity, and reproduction in various yak breeds. For example, whole-genome sequencing with 10· coverage was performed on 406 Maiwa yaks, identifying candidate genes such as *MFSD4*, *LRRC37B*, and

NCAM2 that are associated with body weight traits [16]. GWAS studies on yak immune markers have revealed that the *CASZ1* gene is associated with immune traits in yaks, while the *KMT2C* gene is linked to inflammatory responses [17]. By combining phenotypic analysis with genotyping chips, candidate genes such as *SNX9* and *PRKAA2* were identified as being associated with body size traits in yak [18]. Although candidate SNPs or genes associated with body conformation and body weight traits have been reported in the literature, our understanding of them remains limited. Therefore, further research is needed to validate their utility in marker-assisted selection. More research is needed to enhance the effectiveness of these genetic markers in breeding programs. Furthermore, as a rare and valuable local yak subpopulation in China, the Tianzhu White Yak was included in the “Compendium of Chinese Cattle Breeds” in 1998 and listed in the National Livestock and Poultry Breed Resource Conservation Catalog in 2000. It possesses high market value and economic significance, and its unique genetic resources serve as valuable material for studying the underlying mechanisms of phenotypic traits. We utilized low-coverage sequencing supplemented with high-coverage data to obtain high-quality genome-wide data, and conducted a GWAS on four body conformation traits (body weight, body height, body oblique length, and chest circumference) in Tianzhu White Yak. The identified SNP loci and candidate genes will provide valuable genetic resources for understanding the genetic architecture of body conformation traits and for advancing genomic selection and breeding programs for the Tianzhu White Yak.

Materials and methods

Experimental Animals and Body Measurements

The experimental yaks were sourced from pastoral areas in Tianzhu Tibetan Autonomous County, Wuwei City, Gansu Province, China. The average elevation of the region is 2,500 meters. In this study, 536 Tianzhu white yaks were sampled, and their body weight, height at withers, body length, and chest circumference were measured. All measurements were taken in the morning on an empty stomach in early November 2025. Blood samples were collected from the jugular vein and stored in vacuum blood collection tubes pre-filled with anticoagulant. After thorough mixing, the samples were transported to the laboratory at -20°C for DNA extraction.

DNA Library Construction and Sequencing

DNA was extracted from the blood of Tianzhu white yak using the TIANamp Genomic DNA Kit (TianGen Biotechnology Co., Ltd., Beijing), and DNA concentration and sample quality were assessed using a NanoDrop™ 2000 spectrophotometer and 1.5% agarose gel electrophoresis. DNA samples that passed quality control underwent library preparation according to the following steps: restriction enzyme digestion, random fragmentation, end repair, A-tailing, sequencing adapter ligation, purification, and ligation. The constructed libraries were sequenced using a DNBSEQ-T7 high-throughput sequencer to obtain raw FASTQ data.

SNP Extraction, Genotype Completion, and Data Quality Control

After data acquisition, to ensure the quality of subsequent analysis, the data were evaluated and filtered using SOAPnuke (v.2.2.1) [19]. The high-quality sequences obtained after filtering were then aligned to the reference genome (LU_Bosgru_v3.0) using BWA (v.0.7.17) [20]. The resulting alignment data was converted into sorted BAM format files using Samtools (v.1.9) [21]. Subsequently, statistical analysis and quality assessment of the BAM files were performed using the bamqc module of the Qualimap2 (v.2.2.2-dev) software [22]. Based on this, GATK (v.4.2.6.1) [23] was used with default parameters to detect single-nucleotide polymorphisms (SNPs), generate a VCF file, and extract SNP variants using the SelectVariants tool. The VCF file containing the extracted SNP variants was then used for genotype imputation with GLIMPSE2 software (v.5.4). The reference panel for genotype imputation was constructed using 950 yak genomes with 20× coverage from the research group's previous studies, serving to assist in the inference and correction of missing genotypes in low-coverage samples. To improve data quality, Plink (v.1.90) was used to perform strict quality control according to the following criteria: --geno 0.2 (filtering out variant sites with a genotype missing rate exceeding 20%), --hwe 1×10^{-5} (p-value > 1×10^{-5} for the Hardy-Weinberg equilibrium test), --maf 0.01 (filtering out variants with a minimum allele frequency below 0.01), and --mind 0.3 (excluding individuals with a genotype missing rate exceeding 30%) [24]. This process ultimately yielded 24,383,269 high-quality SNP loci for subsequent analysis. Finally, SnpEff (v.5.0) [25] was used to perform functional annotation and classification of the obtained SNPs.

Population Structure and Phylogenetic Analysis

Population genotype data were used to calculate the genomic relationship matrix (GRM) among samples. The GRM was computed using the GCTA (v.1.94.0) software based on the VanRaden method, using only autosomal SNPs [10]. The complete GRM matrix was imported into R (v.4.2.2), and a relationship heatmap was generated using the heatmap package (v.1.0.13) to visualize genetic similarity among individuals. To determine the population structure of the experimental population, analysis was performed using the Admixture software (v.1.3.0).

Genome-wide association studies and candidate gene annotation

We used a linear mixed model (MLM) to conduct a genome-wide association study (GWAS) of SNPs and body conformation traits. Traditionally, Bonferroni correction results in a high false-negative rate and may overlook some SNPs that are truly associated with growth traits [26]. In this study, the threshold for genome-wide significance was set at $P = 1 \cdot 10^{-6}$, consistent with previous studies [10, 11]. The proportion of phenotypic variation explained (PVE) by a single SNP was estimated using the formula reported in previous studies [27].

$$PVE = \frac{2\beta^2 \cdot MAF \cdot (1 - MAF)}{2\beta^2 \cdot MAF \cdot (1 - MAF) + 2N \cdot (SE[\beta])^2 \cdot MAF \cdot (1 - MAF)}$$

Here, β is the estimated effect of the SNP variant, SE(β) is the standard error of the β estimate, MAF is the minor allele frequency, and N is the sample size. Manhattan plots and QQ plots were generated using R. The first two principal components were extracted using R, and principal component analysis (PCA) scatter plots were plotted using the ggplot2 package to examine whether stratification existed within the cohort. The first five principal components were included as covariates in a mixed linear model to correct for false associations in GWAS caused by population stratification. Linkage disequilibrium (LD) decay analysis was performed using PopLDdecay (v.3.31), with a maximum distance threshold set at 500 kb. The resulting pairwise r^2 values were binned and averaged to generate an LD decay plot. To identify candidate genes associated with the target trait, this study utilized the Ensembl database. Based on the yak gene annotation (LU_Bosgru_v3.0), significant SNP loci identified through association analysis were used to locate potential candidate genes within a 240 kb upstream and downstream region.

Functional enrichment analysis

Functional Enrichment Analysis

The selected candidate genes were annotated for GO functions using the DAVID database (https://david.ncifcrf.gov), and KEGG pathway analysis was performed using KOBAS 3.0 (http://bioinfo.org/kobas) [24]. To accurately elucidate the biological functions of the annotated candidate genes and changes in pathways, the species *Bos grunniens* (domestic yak) was specifically selected for enrichment analysis. In the enrichment analysis, $p < 0.05$ was considered statistically significant.

Statistical Analysis

Excel was used to summarize the growth trait data, and then SPSS25.0 software (IBM, Armonk, NY, USA) was used to perform One-Way ANOVA analysis of variance and LSD multiple comparisons on all data. The results were expressed as mean ± SD. $p < 0.05$ indicated significant difference.

Results

Statistical Analysis of the Phenotypic Data of Tianzhu White Yak

The 536 Tianzhu White Yak animals included in this study were primarily divided into three age groups: 1 year old, 2.5 years old, and 4 years old and older. Descriptive statistics were performed on body weight (BW), body height (BH), body oblique length (BOL), and chest circumference (CC) for the Tianzhu White

Yak in these three age groups (Table 1). The results indicate that significant phenotypic variation exists for these growth traits within the Tianzhu White Yak population, with the coefficient of variation for body weight at 1 year of age being the highest at 31.28%. Furthermore, normality tests confirmed that most traits approximate a normal distribution (Fig. 1A–C), thereby ensuring the applicability of parametric methods in subsequent analyses.

Table 1
Statistical data on growth traits of sequenced Tianzhu white yak individuals

Items	1 year (n = 98)		2.5 year (n = 201)		4 ~ year (n = 237)	
	mean ± SD	CV (%)	mean ± SD	CV (%)	mean ± SD	CV (%)
BW (kg)	85.01 ± 26.59	31.28	153.55 ± 33.80	22.01	211.06 ± 33.00	15.64
BH (cm)	83.76 ± 6.17	7.37	100.17 ± 7.36	7.34	106.85 ± 5.69	5.33
BOL (cm)	83.43 ± 8.85	10.61	106.26 ± 9.93	9.34	118.51 ± 10.30	8.69
CC (cm)	125.52 ± 15.73	12.53	147.85 ± 14.06	9.51	161.13 ± 11.26	6.99
Note: BW = Body weight; BH = Body height; BOL = Body oblique length; CC = Chest circumference.						

Population structure and kinship analysis

In this study, whole-genome resequencing was performed on 536 Tianzhu White yak individuals to evaluate their growth traits, including BW, BH, BOL, and CC. The whole-genome sequencing was conducted at a low depth of approximately 1.49×. Following genotype filling and quality control, the average DR^2 of the filled variants was 0.944. The GC content ranged from 42.19% to 46.87%. The average Q20 score after quality control was 98.88%, and the average Q30 score was 97.17%, indicating high sequencing quality and no significant bias in the library preparation and sequencing processes (Table S1). Before data filtering, the raw data contained 53,602,519 SNP loci. After quality control, a total of 24,383,269 SNPs were retained for subsequent analysis. Analysis of the population structure of the Tianzhu White Yak revealed homogeneity at K = 1. As K increased to 2–5, distinct ancestral components emerged among individuals, indicating a slight subpopulation structure within the Tianzhu White Yak population (Fig. 2A). The optimal K value for the Tianzhu White Yak was evaluated using six-fold cross-validation, with the smallest cross-validation error observed at K = 4 (Fig. 2B). Principal component analysis (PCA) results showed that most individuals clustered together, with only a few outliers, indicating a relatively uniform genetic structure within the population and no obvious stratification (Fig. 2C). LD decay curves were plotted using the ggplot2 package in R (Fig. 2D). The results indicate that when the LD coefficient r^2 declines to 0.1, the corresponding physical distance is approximately 240 kb, and the rate of decline gradually flattens. Therefore, this study will identify potential key genes within a 240 kb range upstream and downstream of significant loci. Phylogenetic heatmap analysis revealed a low overall degree of relatedness among individuals, with only a few localized clusters suggesting a possible family structure (Fig. 2E).

Genome-wide association study of growth traits

We used a linear mixed model (MLM) to conduct a GWAS of SNPs and body conformation traits. Based on a P-value threshold of 1×10^{-6}, we identified 366 significantly associated SNP loci. These SNPs were annotated to 32 candidate genes associated with growth traits, and the PVE for each SNP was calculated based on β values and MAF (Table S2). A GWAS analysis of growth traits in 1-year-old Tianzhu white yak (Fig. 3, Table S2) identified two significant SNPs only for BW, leading to the identification of one candidate gene, such as *PLCB1*. A total of 331 significantly associated SNP loci were identified in the GWAS analysis of growth traits in Tianzhu white yak at 2.5 years of age (Fig. 4, Table S2). For the BW trait, 21 significant SNPs were identified, and 11 candidate genes were characterized, such as *SERGEF*, *SAR1B*, and *SRSF6*. The highest number of SNPs, 308 in total, was significantly associated with the BH trait, annotated to 12 candidate genes, including *BRPF3*, *MAPK13*, and *ZNF76*. Only one SNP was significantly associated with BOL, annotated to one candidate gene, such as *ABCD3*. Similarly, only one significant SNP was associated with CC, annotated to one candidate gene, such as *LMNTD1*. A total of 33 SNPs showing significant associations were identified in the GWAS analysis of growth traits in Tianzhu White Yak aged 4 years and older (Fig. 5, Table S2). No significant SNPs were detected for body weight. A total of 22 SNPs were significantly associated with BH, and these were annotated to three candidate genes, including *B4GALT4*, *FAM98B*, and *HPGD*. A total of 3 SNPs were significantly associated with BOL, annotated to 1 candidate gene, such as *NTM*. A total of 8 significant SNPs were associated with CC, annotated to 2 candidate genes, such as *ACVR1* and *ACSS3*. QQ plots of the GWAS results across the three stages showed that λ values were all close to 1, with no obvious signs of population stratification.

Functional enrichment analysis

Functional enrichment analysis was performed on the 32 candidate genes identified through screening for various traits at different ages. The results of the GO enrichment analysis (Fig. 6A) showed enrichment in three biological processes: negative regulation of the G1/S transition of the mitotic cell cycle, fatty acid beta-oxidation, and positive regulation of transcription by RNA polymerase II. Four were enriched in cellular components: chromatin, transcription regulator complex, nucleoplasm, and Golgi cisterna membrane. One was enriched in molecular function: phosphatidylinositol-4,5-bisphosphate binding. Among these, *RBL1*, which is associated with BW, was significantly enriched in GO terms such as the transcription regulator complex. KEGG pathway enrichment analysis revealed (Fig. 6B) that the candidate genes were significantly enriched in multiple pathways, including growth and development-related signaling pathways such as Hippo, TGF-β, and sphingolipid pathways, involving genes such as *PLCB1*, *PPP2R2B*, and *MAPK13*.

Genotyping results for key SNPs loci

Six significant SNP loci were selected for validation to determine the direction of their effects (positive or negative) on phenotypes associated with the Tianzhu White Yak. The candidate genes identified at these

loci were *ADGRL1*, *SLC16A7*, *MAPK13*, *ANKS1A*, *ACVR1*, and *NTM*. Based on genotyping results, all six SNP loci exhibited a low frequency of mutant individuals, and there was consistently a highly significant difference between the wild-type and mutant homozygous genotypes (Fig. 7).

Discussion

As a rare and precious yak genetic resource in the world, Tianzhu white yak has formed a local variety of yak with both meat and hair after long-term natural selection and artificial breeding. Its series of products such as meat, milk, skin, hair, cashmere, tail and bone have unique quality and market value. Although significant achievements have been made in the conservation of the Tianzhu White Yak's genetic resources and the development and utilization of its industry, the breed still faces a series of challenges and difficulties. With the advancement of genome-enhanced breeding and the improvement of genomic selection efficiency, the conservation and development of the Tianzhu White Yak's breeding resources will become feasible, thereby promoting the modernization and industrialization of the Tianzhu White Yak industry [28]. Body measurement indices are important economic traits for the genetic improvement of large livestock [29]. GWAS are an effective method for identifying major genes associated with complex traits, providing new avenues for studying the genetic mechanisms underlying these traits. This study utilized resequencing and GWAS to identify SNPs and genetic variations influencing growth traits. The goal is to provide scientific evidence and technical support for the conservation, genetic improvement, and sustainable utilization of the Tianzhu White Yak's genetic resources.

This study utilized low-coverage sequencing data with an average sequencing depth of 1.49×. The data alignment rate reached 99.27%, and the average genome coverage of ≥ 1× was 59.39%, which is consistent with expectations for low-coverage sequencing. To compensate for the low coverage, we used a population-matched reference panel comprising 950 Tianzhu white yak individuals (sequencing depth approximately 20×) to perform genotype imputation for the 536 low-coverage individuals, achieving an average DR^2 of 0.944 for the imputed variants. Low-coverage sequencing combined with a high-density haplotype reference panel restores uncovered or low-confidence loci through genotype imputation, thereby obtaining high-density genome-wide markers on a large-scale basis [30]. Studies have shown that when the sample size is sufficiently large and the sequencing depth is ≥ 1×, the accuracy of the imputed genotypes can approach that of high-depth sequencing [31]. This strategy has been widely applied in large-scale population genetics, genome-wide association studies, molecular breeding, prenatal aneuploidy screening, and circulating tumor DNA detection, serving as an effective approach that balances cost and data breadth in large-scale genomic research.

Sample size is critical to the reliability of genome-wide association study results [32]. This study collected phenotypic data from a total of 536 Tianzhu White yak individuals at three different developmental stages: 98 at 1 year of age, 201 at 2.5 years of age, and 237 at 4 years of age or older. Although this sample size may be considered moderate for a typical genome-wide association study, it is important to note that such analyses have demonstrated robust performance in populations of similar or

smaller sizes. This has been confirmed by several previous studies, including a GWAS of early growth traits in 212 Inner Mongolian cashmere goats, which identified 21 SNPs and 117 candidate genes associated with early growth traits [33]. Another study involving 100 Qira Black sheep and 84 German Merino sheep identified 84 candidate genes associated with production traits [34]. A significantly smaller study involving 96 Baluchi sheep was also conducted [35]. Similarly, small-sample GWAS studies have been performed in large livestock; for example, a GWAS of a 202 head Huaxi cattle population identified 99 variants significantly associated with growth traits and annotated 83 candidate genes [36]. These studies have all successfully identified significant genetic variations associated with body weight and body size traits. Furthermore, our rigorous phenotypic measurements across multiple growth stages further reinforce our confidence that our sample size is sufficient for meaningful genetic discoveries.

The coefficients of variation (CV) for BW, BH, BOL, and CC across the three age groups in this study ranged from 5.33% to 31.28%, with the CV for BW in 1-year-old Tianzhu White Yak being the highest at 31.28%. In contrast, the CVs for BH, BOL, and CC remained at low to moderate levels (5.33%–12.53%) across all age groups, indicating good consistency of these morphological traits within the population. Overall, this CV range is considered acceptable for growth trait GWAS studies and is suitable for subsequent association analysis. To obtain better GWAS analysis results, we analyzed the population structure and kinship of the Tianzhu White Yak population. Population structure analysis showed that when $K = 2$, individuals were clearly divided into two major subpopulations, whereas when $K \geq 3$, the population partitioning tended to become mixed, indicating the presence of multiple minor subpopulations within the population. This result is consistent with the PCA analysis, where PC1 and PC2 revealed distinct clustering among individuals, yet some genetic admixture still existed between different subpopulations. The kinship heatmap showed that most individuals were not closely related, indicating that the overall sample was suitable for GWAS analysis. Additionally, the lighter-colored blocks corresponding to subpopulation differentiation corroborated the results of the PCA and population structure analyses.

Using genomic sequencing and GWAS studies, we demonstrated that 32 candidate genes are significantly associated with growth traits in Tianzhu White Yak. For BW traits, two significant SNPs were identified in an association analysis of 1-year-old Tianzhu White Yak, leading to the identification of the *PLCB1* (Phospholipase C Beta 1) gene. The study found that *PLCB1* participates in the phosphatidylinositol signaling pathway, which is involved in the regulation of glucose metabolism, such as gluconeogenesis and glycogen metabolism [37]. This may be related to the fact that 1-year-old yaks are in the stage of cell proliferation and early differentiation, providing them with the energy and metabolic regulation necessary for growth. *ADGRL1* (Adhesion G-Protein Coupled Receptor L1) is an adhesion G protein-coupled receptor (adhesion GPCR) with a long extracellular domain that participates in cell-cell and cell-matrix interactions and may play a role in cell adhesion and signal transduction [38]. Studies have shown that adhesion GPCRs (including *ADGRL1*) are involved in cytoskeletal remodeling, which is closely associated with cell migration, shape changes, and dynamics [39]. *ADGRL1* not only mediates cell adhesion but also activates downstream G proteins and small GTPase signaling [40]; therefore, the *ADGRL1* gene may influence processes such as cell proliferation and migration in Tianzhu

white yak cells. *PPP2R2B* encodes the regulatory subunit B55b of protein phosphatase 2A (PP2A), and this gene may serve as a key regulator in the AKT/mTOR signaling pathway [41]. Studies have found that the *PPP2R2B* gene, as a regulatory subunit of PP2A, plays a central role in regulating protein dephosphorylation, cell signaling inhibition, and cell cycle control [42]. In Chinese Holstein cattle, the *PPP2R2B* gene has been identified as being associated with milk production traits [43], and it has also been found to be potentially associated with sperm motility in Duroc pigs [44]. In this study, the gene was found to be associated with body weight traits in Tianzhu White Yak.

Regarding body conformation, *SLC16A7* (Monocarboxylate Transporter 7) encodes a transmembrane protein whose primary function is to transport monocarboxylic acids, such as lactate, pyruvate, and ketone bodies [45]. In studies of growth traits in Barki sheep, including birth weight and weaning weight, *SLC16A7* was identified as a candidate gene for birth weight [46]. Egami R et al. [47] used a mouse model to genetically engineer the monocarboxylate transporter Slc16a7 into skeletal muscle. The results showed that *SLC16A7* expression in skeletal muscle decreased, and its ability to transport ketone bodies was impaired. This indicates that it participates in the metabolic regulation of skeletal muscle under normal physiological conditions. Interestingly, we annotated this gene in the BH trait of 2.5-year-old Tianzhu White yak. The protein encoded by the *ANKS1A* gene is a member of the Anks1 protein family, featuring two SAM domains and one PTB domain at its C-terminus [48]. Based on structural differences in PTB adaptors, the Anks1 family of adaptors belongs to the Dab class of PTB adaptors, which play a role in regulating the transport of transmembrane proteins [49]. Studies have found that the *ANKS1A* gene is associated with eye muscle area and eye muscle depth in pigs [50]. *ACVR1* (Activin A receptor type I) encodes the type I osteomorphogenic protein receptor of the TGF-β receptor superfamily. It is involved in various biological processes, including the development and regulation of the skeletal, cardiac, cartilaginous, nervous, and reproductive systems [51]. *NTM* (Neurotrimin) belongs to the IgLON family of glycosylphosphatidylinositol (GPI)-anchored cell adhesion molecules and has been shown to promote neurite growth and adhesion [52]. Studies indicate that NTM acts in concert with other cell adhesion proteins to regulate neurite outgrowth, axon fasciculation, and synapse formation—processes that are critical for the proper positioning and development of the nervous system [53].

A functional enrichment analysis was performed on the annotated genes. The GO enrichment analysis revealed that the annotated genes are primarily involved in membrane-related functions (such as phosphatidylinositol-4,5-bisphosphate binding and Golgi cisterna membrane), transcriptional regulation (Transcription regulator complex), and cell cycle regulation (Negative regulation of G1/S transition of mitotic cell cycle). The “negative regulation of G1/S transition of the mitotic cell cycle” entry primarily involves inhibiting the transition of the cell cycle from the G1 phase to the S phase, regulating cell proliferation rates, and maintaining cell cycle homeostasis [54]. Interestingly, the *ACVR1* gene annotated in this study is enriched in this pathway. As a BMP/TGF-β receptor, *ACVR1* can downregulate Cyclin/CDK activity associated with the G1/S transition via the Smad signaling pathway, thereby regulating the rate of cell proliferation [55–57]. KEGG enrichment analysis revealed that differentially expressed genes were significantly enriched in signaling pathways such as Sphingolipid, TGF-beta, Hippo, and Wnt, which are extensively involved in the regulation of cell proliferation, differentiation, and tissue development [58, 59].

Furthermore, the enrichment of the Calcium signaling pathway and Rap1 signaling pathway suggests that these genes may be involved in cell migration and morphological changes, providing a molecular basis for the growth and development of the Tianzhu white yak.

Conclusion

In summary, this study conducted genome-wide association analyses for four body conformation traits in a total of 536 Tianzhu White yak at ages 1, 2.5, and 4 years or older. A total of 32 candidate genes were identified, including *ADGRL1*, *SLC16A7*, *MAPK13*, *ANKS1A*, *ACVR1*, and *NTM*, which were associated with body weight and conformation traits in Tianzhu White Yak. These results provide useful molecular markers for breeding programs and offer a pathway for genetically improving the growth performance and productivity of this breed. Furthermore, these loci and genes provide foundational information for elucidating the complex genetic mechanisms underlying growth in Tianzhu White Yak.

Declarations

Acknowledgements

The authors would like to express their gratitude to all reviewers who participated in the review of the paper, as well as to Glbizzia Biosciences Co. Ltd, which provided sequencing-related services for the trial. Furthermore, thank you to the ranch staff for their assistance during the sample collection process.

Author contributions

GWY performed the experiments and the manuscript. XDQ, YFL and XMM performed the sample collection. GWY, YCL and ZCW analysed the data. WXL, WWR, ZYZ and MC suggested revisions to the manuscript. XYW and XG was involved in the design of the experiments. WZQ and CNL designed the study and provided overall guidance. All authors read and approved the final version of the manuscript.

Funding

This research was funded by Central Guidance Funds for Local Science and Technology Development Projects (25ZYJA008); the Project for Collecting and Evaluating Sample Data of Livestock and Poultry Genetic Resources of the Department of Agriculture and Rural Affairs of Gansu (GSZYTC-ZX-260480); the Modern Beef Yak Industry Technology System, grant number MATS-Beef Cattle System, CARS-37; Gansu Province Talent Program(2026RCXM096).

Data availability

The sequenced datasets generated and analyzed during the current study are publicly available from GenBank with the Bio project (SRA accession: PRJNA1472866).

Ethics approval and consent to participate

All the animal experiments were approved by Lanzhou Institute of Husbandry and Pharmaceutical Sciences of the Chinese Academy of Agricultural Sciences (CAAS) with the grant number: No. 2026-51. All the slaughter as well as sampling procedures strictly complied with the Guidelines for Ethical Treatment of Experimental Animals of China.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

1. Zhang Q, Wang Q, Zhang Y, Cheng S, Hu J, Ma Y, Zhao X. Comprehensive Analysis of MicroRNA–Messenger RNA from White Yak Testis Reveals the Differentially Expressed Molecules Involved in Development and Reproduction. Int J Mol Sci. 2018;19(10):3083.
2. Ma J, Zhu Y, Wang Z, Yu X, Hu R, Wang X, Cao G, Zou H, Shah AM, Peng Q, et al. Comparing the Bacterial Community in the Gastrointestinal Tracts Between Growth-Retarded and Normal Yaks on the Qinghai-Tibetan Plateau. Front Microbiol. 2020;11:600516.
3. Li X, He SG, Li WR, Luo LY, Yan Z, Mo DX, Wan X, Lv FH, Yang J, Xu YX, et al. Genomic analyses of wild argali, domestic sheep, and their hybrids provide insights into chromosome evolution, phenotypic variation, and germplasm innovation. Genome Res. 2022;32(9):1669–84.
4. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019;20(8):467–84.
5. Stevens A, Meyer S, Hanson D, Clayton P, Donn RP. Network analysis identifies protein clusters of functional importance in juvenile idiopathic arthritis. Arthritis Res Ther. 2014;16(3):R109.
6. Zhang J, Zha X, Yang G, Ma X, La Y, Wu X, Guo X, Chu M, Bao P, Yan P et al. Polymorphisms of TXK and PLCE1 Genes and Their Correlation Analysis with Growth Traits in Ashidan Yaks. *Animals (Basel)* 2024, 14(23).
7. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, et al. Complement factor H polymorphism in age-related macular degeneration. Science. 2005;308(5720):385–9.
8. Bejarano DH, Martínez RA, Rocha JF. Genome-wide association study for growth traits in Blanco Orejinegro and Romosinuano cattle. Trop Anim Health Prod. 2023;55(6):358.
9. Edea Z, Jeoung YH, Shin SS, Ku J, Seo S, Kim IH, Kim SW, Kim KS. Genome-wide association study of carcass weight in commercial Hanwoo cattle. Asian-Australas J Anim Sci. 2018;31(3):327–34.
10. Han M, Wang X, Du H, Cao Y, Zhao Z, Niu S, Bao X, Rong Y, Ao X, Guo F, et al. Genome-wide association study identifies candidate genes affecting body conformation traits of Zhongwei goat.

BMC Genomics. 2025;26(1):37.

11. He XY, Wei YS, Kuang JC, Li ZF, Yu ZX, Wang XY, Di R, Zhu CY, Chu MX. Genome-wide association study identifies candidate genes regulating multi-stage body size traits in Longling yellow goats. Animal. 2026;20(3):101766.
12. Tang Z, Xu J, Yin L, Yin D, Zhu M, Yu M, Li X, Zhao S, Liu X. Genome-Wide Association Study Reveals Candidate Genes for Growth Relevant Traits in Pigs. Front Genet. 2019;10:302.
13. Wang H, Wang X, Li M, Sun H, Chen Q, Yan D, Dong X, Pan Y, Lu S. Genome-Wide Association Study of Growth Traits in a Four-Way Crossbred Pig Population. *Genes (Basel)* 2022, 13(11).
14. Ye H, Li W, Tian F, Wang Q, Ma Z, Guan J, Ding Y, Zheng X, Yin Z, Zhang X. Genome-Wide Association Study Identifies Candidate Genes for Body Size Traits in Wanyue Black Pigs. Anim (Basel) 2025, 16(1).
15. Terakado APN, Costa RB, de Camargo GMF, Irano N, Bresolin T, Takada L, Carvalho CVD, Oliveira HN, Carvalheiro R, Baldi F, et al. Genome-wide association study for growth traits in Nelore cattle. Animal. 2018;12(7):1358–62.
16. Wang J, Li X, Peng W, Zhong J, Jiang M. Genome-Wide Association Study of Body Weight Trait in Yaks. Anim (Basel) 2022, 12(14).
17. Yu D, Ma X, Huang C, Wang T, Zhang M, Feng F, Wu X, La Y, Guo X, Yan P et al. Genome-Wide Association Study of Immune Indices in Yaks. Anim (Basel) 2025, 15(14).
18. Zhao J, Li Z, Liu X, Liu Y, Yue B, Wang H, Zhang M, Peng W, Shu S, Wang G, et al. Large-scale genome-wide association analysis reveals candidate genes in yak body size traits. BMC Genomics. 2025;26(1):844.
19. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009.
20. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.
21. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics. 2016;32(2):292–4.
22. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–303.
23. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012;6(2):80–92.
24. Liu Y, Qi X, La Y, Ma X, Ren W, Yang G, Zhang Z, Chu M, Wu X, Guo X et al. Genome-Wide Association Study of Gross Hair Weight and Hair Length Traits in Different Body Regions of Tianzhu White Yak. Biomolecules 2026, 16(2).

25. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
26. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ. 1995;310(6973):170.
27. Shim H, Chasman DI, Smith JD, Mora S, Ridker PM, Nickerson DA, Krauss RM, Stephens M. A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. PLoS ONE. 2015;10(4):e0120758.
28. Ma J, Cao Y. Genetic Dissection of Grain Yield of Maize and Yield-Related Traits Through Association Mapping and Genomic Prediction. Front Plant Sci. 2021;12:690059.
29. Li G, Liu X, Ma Y, Wang B, Zheng L, Wang M. Body size measurement and live body weight estimation for pigs based on back surface point clouds. Biosyst Eng. 2022;218:10–22.
30. Wang C, Zhan X, Liang L, Abecasis GR, Lin X. Improved ancestry estimation for both genotyping and sequencing data using projection procrustes analysis and genotype imputation. Am J Hum Genet. 2015;96(6):926–37.
31. Fei C, Zhang S, Chen X, Liu J, Peng W, Zhang G, You W, Wu F. Evaluation of Low-Coverage Sequencing Strategies for Whole-Genome Imputation in Pacific Abalone Haliotis discus hannai. Int J Mol Sci 2025, 26(10).
32. Grotzinger AD, Fuente J, Privé F, Nivard MG, Tucker-Drob EM. Pervasive Downward Bias in Estimates of Liability-Scale Heritability in Genome-wide Association Study Meta-analysis: A Simple Solution. Biol Psychiatry. 2023;93(1):29–36.
33. Rong Y, Ao X, Guo F, Wang X, Han M, Zhang L, Xia Q, Shang F, Lv Q, Wang Z et al. Genome-Wide Association Analysis Revealed Candidate Genes Related to Early Growth Traits in Inner Mongolia Cashmere Goats. Vet Sci 2025, 12(3).
34. Tuersuntuoheti M, Zhang J, Zhou W, Zhang CL, Liu C, Chang Q, Liu S. Exploring the growth trait molecular markers in two sheep breeds based on Genome-wide association analysis. PLoS ONE. 2023;18(3):e0283383.
35. Pasandideh M, Gholizadeh M, Rahimi-Mianji G. A genome-wide association study revealed five SNPs affecting 8-month weight in sheep. Anim Genet. 2020;51(6):973–6.
36. Liu CL, Ren T, Ruan PC, Huang YF, Ceccobelli S, Huang DJ, Zhang LP. E GX: Genome-Wide Association Integrating a Transcriptomic Meta-Analysis Suggests That Genes Related to Fat Deposition and Muscle Development Are Closely Associated with Growth in Huaxi Cattle. Vet Sci 2025, 12(2).
37. Lacerda JT, Gomes PRL, Zanetti G, Mezzalira N, Lima OG, de Assis LVM, Guler A, Castrucci AM, Moraes MN. Lack of TRPV1 Channel Modulates Mouse Gene Expression and Liver Proteome with Glucose Metabolism Changes. Int J Mol Sci 2022, 23(13).
38. Perdigones N, Perin JC, Schiano I, Nicholas P, Biegel JA, Mason PJ, Babushok DV, Bessler M. Clonal hematopoiesis in patients with dyskeratosis congenita. Am J Hematol. 2016;91(12):1227–33.
39. Cruz-Ortega JS, Boucard AA. Actin cytoskeleton remodeling defines a distinct cellular function for adhesion G protein-coupled receptors ADGRL/latrophilins 1, 2 and 3. Biol Open 2019, 8(4).

40. Morgan RK, Anderson GR, Araç D, Aust G, Balenga N, Boucard A, Bridges JP, Engel FB, Formstone CJ, Glitsch MD et al. The expanding functional roles and signaling mechanisms of adhesion G protein–coupled receptors. *Annals of the New York Academy of Sciences* 2019.
41. Hu Q, Zhou S, Hu X, Zhang H, Huang S, Wang Y. Systematic screening identifies a 2-gene signature as a high-potential prognostic marker of undifferentiated pleomorphic sarcoma/myxofibrosarcoma. J Cell Mol Med. 2020;24(1):1010–21.
42. Shlomai G, Zelenko Z, Antoniou IM, Stasinopoulos M, Tobin-Hess A, Vitek MP, LeRoith D, Gallagher EJ. OP449 inhibits breast cancer growth without adverse metabolic effects. Endocr Relat Cancer. 2017;24(10):519–29.
43. Jia R, Fu Y, Xu L, Li H, Li Y, Liu L, Ma Z, Sun D, Han B. Associations between polymorphisms of SLC22A7, NGFR, ARNTL and PPP2R2B genes and Milk production traits in Chinese Holstein. BMC Genom Data. 2021;22(1):47.
44. Gao N, Chen Y, Liu X, Zhao Y, Zhu L, Liu A, Jiang W, Peng X, Zhang C, Tang Z, et al. Weighted single-step GWAS identified candidate genes associated with semen traits in a Duroc boar population. BMC Genomics. 2019;20(1):797.
45. Gu X, Xiao Q, Ruan Q, Shu Y, Dongre A, Iyer R, Humphreys WG, Lai Y. Comparative untargeted proteomic analysis of ADME proteins and tumor antigens for tumor cell lines. Acta Pharm Sin B. 2018;8(2):252–60.
46. Abousoliman I, Reyer H, Oster M, Murani E, Mohamed I, Wimmers K. Genome-Wide Analysis for Early Growth-Related Traits of the Locally Adapted Egyptian Barki Sheep. Genes (Basel) 2021, 12(8).
47. Egami R, Kokaji T, Hatano A, Yugi K, Eto M, Morita K, Ohno S, Fujii M, Hironaka KI, Uematsu S et al. Trans-omic analysis reveals obesity-associated dysregulation of inter-organ metabolic cycles between the liver and skeletal muscle. *iScience* 2021, 24(3):102217.
48. Park S. Defective Anks1a disrupts the export of receptor tyrosine kinases from the endoplasmic reticulum. BMB Rep. 2016;49(12):651–2.
49. Shan Z, Tu Y, Yang Y, Liu Z, Zeng M, Xu H, Long J, Zhang M, Cai Y, Wen W. Basal condensation of Numb and Pon complex via phase transition during Drosophila neuroblast asymmetric division. Nat Commun. 2018;9(1):737.
50. Zhuang Z, Li S, Ding R, Yang M, Zheng E, Yang H, Gu T, Xu Z, Cai G, Wu Z, et al. Meta-analysis of genome-wide association studies for loin muscle area and loin muscle depth in two Duroc pig populations. PLoS ONE. 2019;14(6):e0218263.
51. Valer JA, Sánchez-de-Diego C, Pimenta-Lopes C, Rosa JL, Ventura F. ACVR1 Function in Health and Disease. Cells 2019, 8(11).
52. Sellar GC, Watt KP, Rabiasz GJ, Stronach EA, Li L, Miller EP, Massie CE, Miller J, Contreras-Moreira B, Scott D, et al. OPCML at 11q25 is epigenetically inactivated and has tumor-suppressor function in epithelial ovarian cancer. Nat Genet. 2003;34(3):337–43.
53. Vinci M, Treccarichi S, Failla P, Musumeci A, Gloria A, Virgillito M, Federico C, Saccone S, Calì F. Unravelling the role of Neurotrimin (NTM), a member of the IgLON family, in mild intellectual

disability and anxiety-like behaviors. Gene. 2026;976:149876.

54. Song D, He H, Indukuri R, Huang Z, Stepanauskaite L, Sinha I, Haldosén LA, Zhao C, Williams C. ERα and ERβ Homodimers in the Same Cellular Context Regulate Distinct Transcriptomes and Functions. Front Endocrinol (Lausanne). 2022;13:930227.
55. Androutsellis-Theotokis A, Leker RR, Soldner F, Hoeppner DJ, Ravin R, Poser SW, Rueger MA, Bae SK, Kittappa R, McKay RD. Notch signalling regulates stem cell numbers in vitro and in vivo. Nature. 2006;442(7104):823–6.
56. Massagué J. TGFbeta in Cancer. Cell. 2008;134(2):215–30.
57. Shi Y, Massagué J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell. 2003;113(6):685–700.
58. Petri V, Hayman GT, Tutaj M, Smith JR, Laulederkind S, Wang SJ, Nigam R, De Pons J, Shimoyama M, Dwinell MR. Disease, Models, Variants and Altered Pathways-Journeying RGD Through the Magnifying Glass. Comput Struct Biotechnol J. 2016;14:35–48.
59. Phillips JE, Santos M, Konchwala M, Xing C, Pan D. Genome editing in the unicellular holozoan Capsaspora owczarzaki suggests a premetazoan role for the Hippo pathway in multicellular morphogenesis. Elife 2022, 11.

Figures

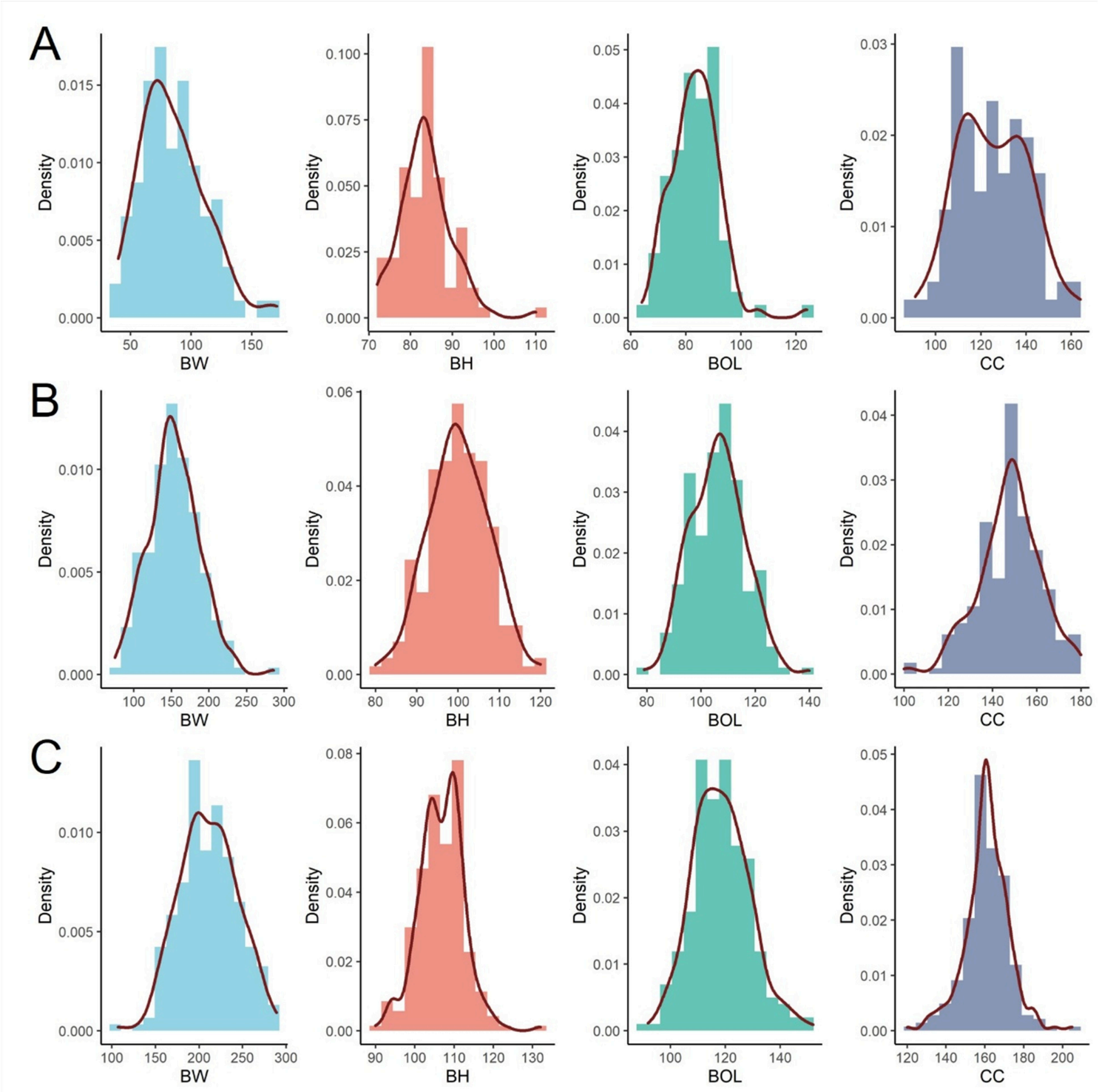


Figure 1

Frequency distribution histograms of phenotypic values for growth traits and heatmap of phenotypic correlations in Tianzhu white yak. A: Histogram showing the frequency distribution of phenotypic values for various body conformation traits at 1 year of age; B: Histogram showing the frequency distribution of phenotypic values for various body conformation traits at 2.5 year of age; C: Histograms showing the frequency distributions of phenotypic values for various body measurements in children aged 4 and

older. Abbreviations: BW = Body weight; BH = Body height; BOL = Body oblique length; CC = Chest circumference.

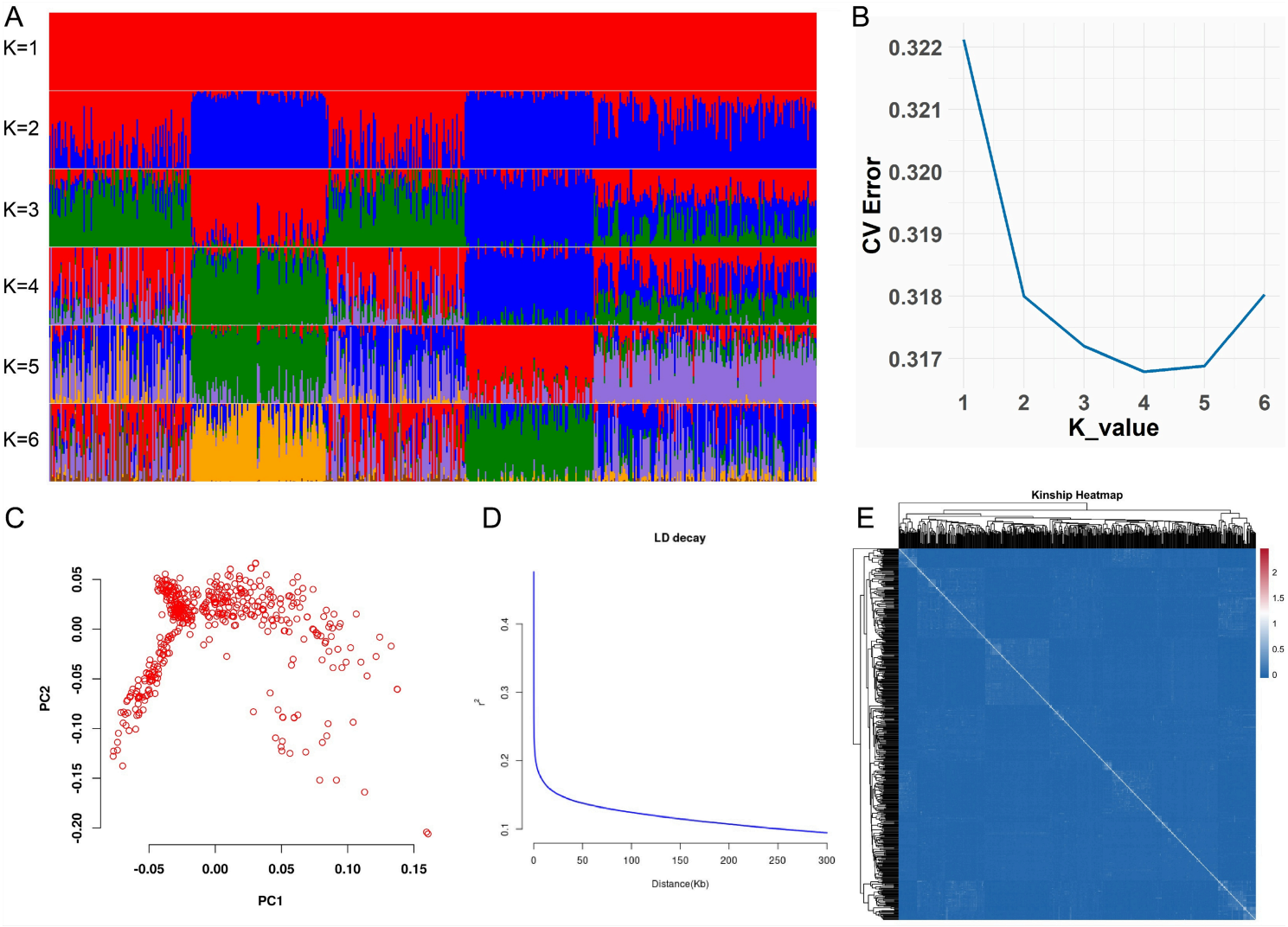


Figure 2

Population structure and kinship analysis in Tianzhu white yak. A: Population structure map; B: Line cross-validation error calculated by the Admixture; C: Principal component analysis results; D: Linkage disequilibrium (LD) decay analysis. E: Heatmap of kinship analysis. Distant and close affinities are shown by a blue-to-red colour change.

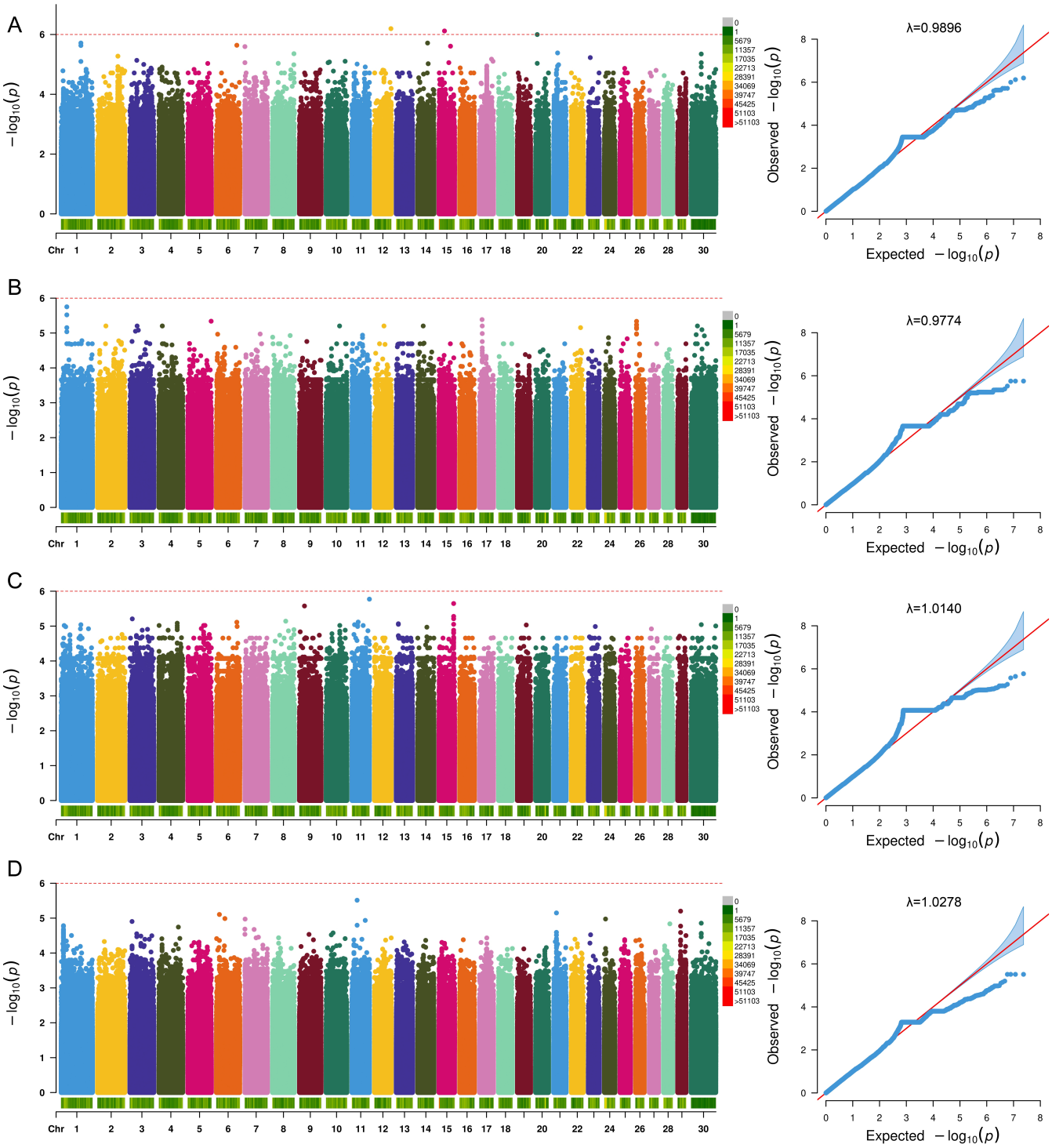


Figure 3

Manhattan and QQ plot of growth traits in 1-year-old Tianzhu white yak. A: Manhattan and QQ plots of BW traits; B: Manhattan and QQ plots of BH traits; C: Manhattan and QQ plots of BOL traits; D: Manhattan and QQ plots of CC traits. The dashed line indicates the genome-wide significance threshold, which is set at 1.0×10^{-6}.

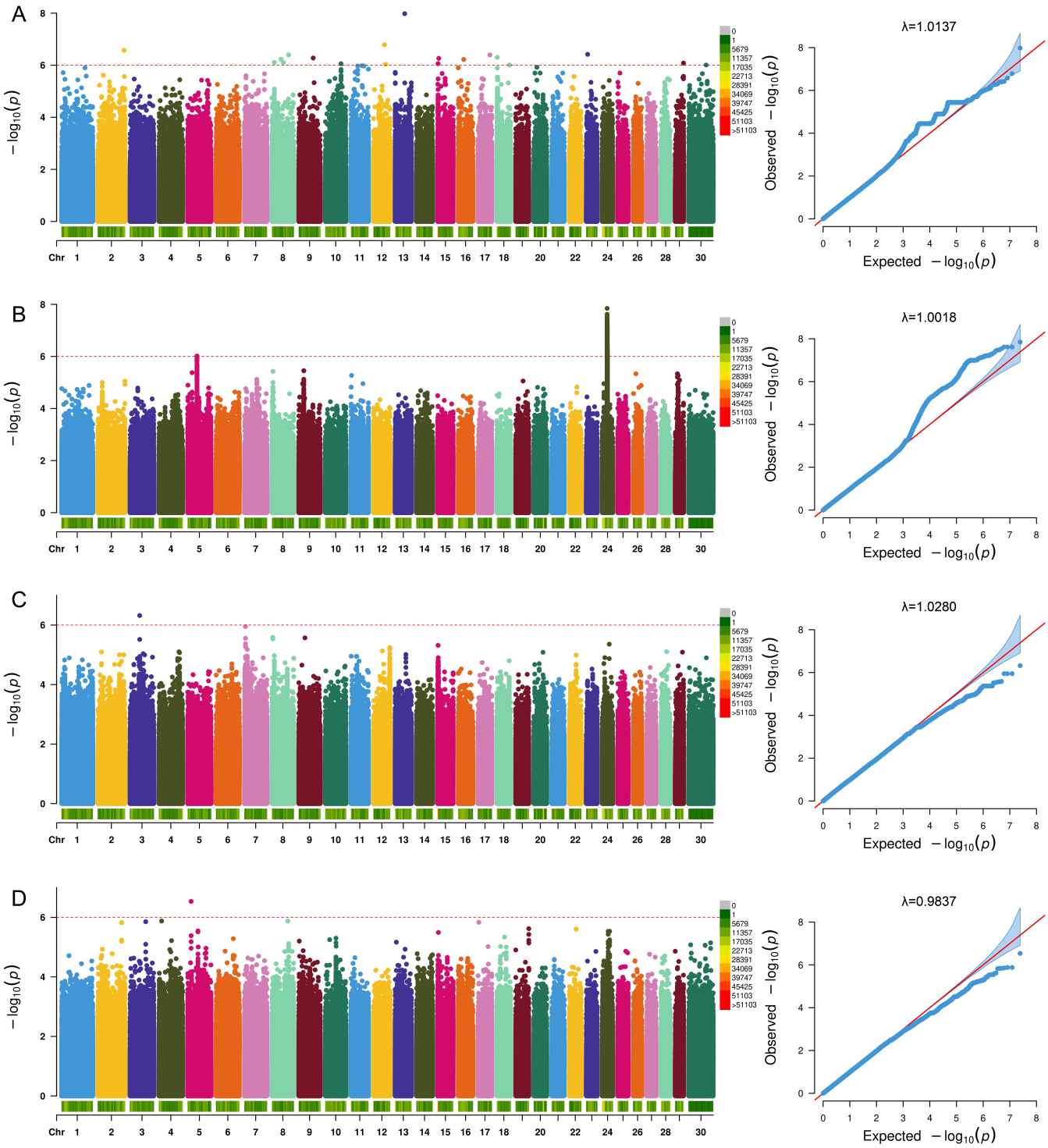


Figure 4

Manhattan and QQ plot of growth traits in 2.5-year-old Tianzhu white yak. A: Manhattan and QQ plots of BW traits; B: Manhattan and QQ plots of BH traits; C: Manhattan and QQ plots of BOL traits; D: Manhattan and QQ plots of CC traits. The dashed line indicates the genome-wide significance threshold, which is set at 1.0×10^{-6}.

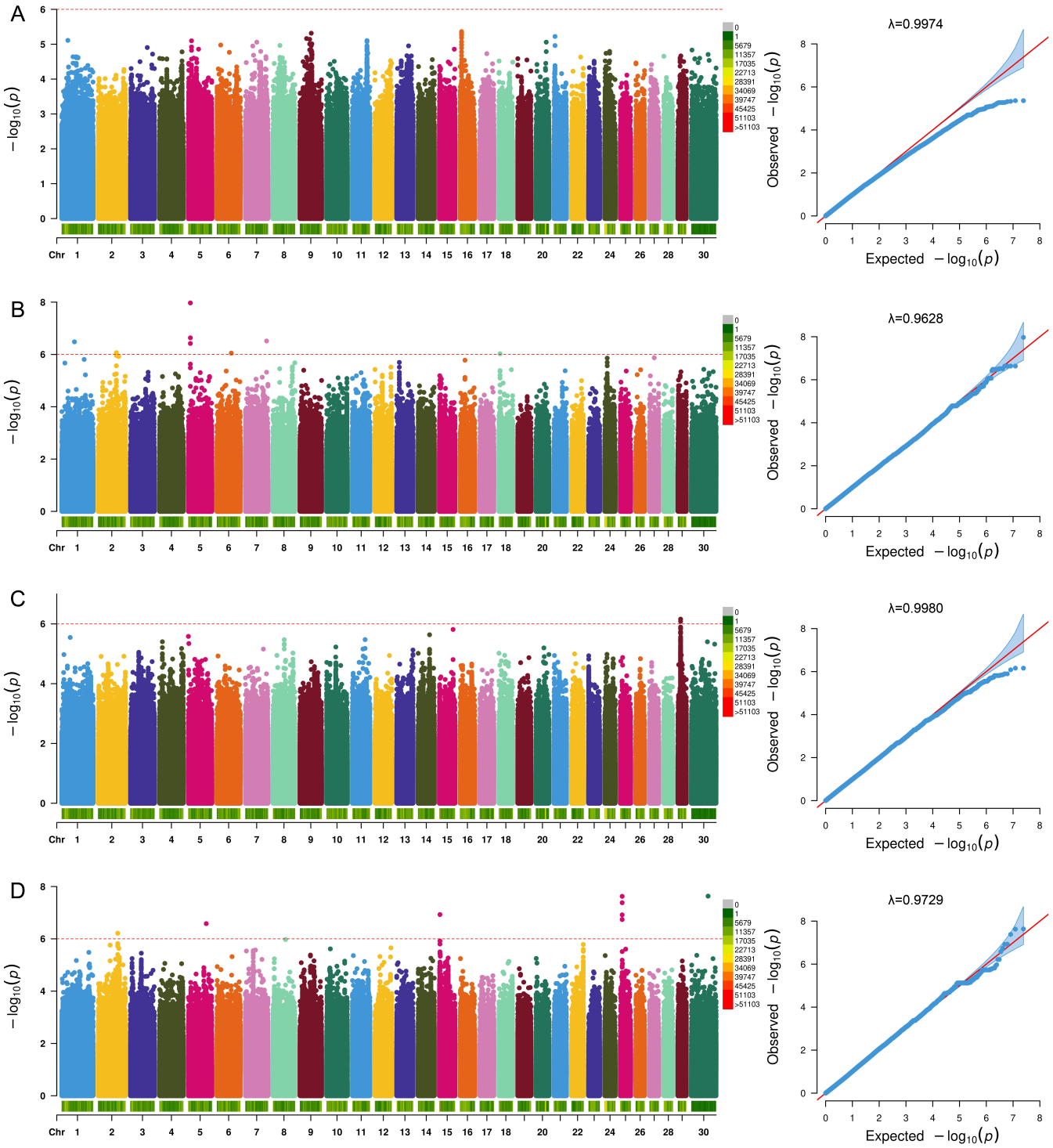


Figure 5

Manhattan plots and QQ plots of growth traits in Tianzhu white yak aged 4 years and older. A: Manhattan and QQ plots of BW traits; B: Manhattan and QQ plots of BH traits; C: Manhattan and QQ plots of BOL traits; D: Manhattan and QQ plots of CC traits. The dashed line indicates the genome-wide significance threshold, which is set at 1.0×10^{-6}.

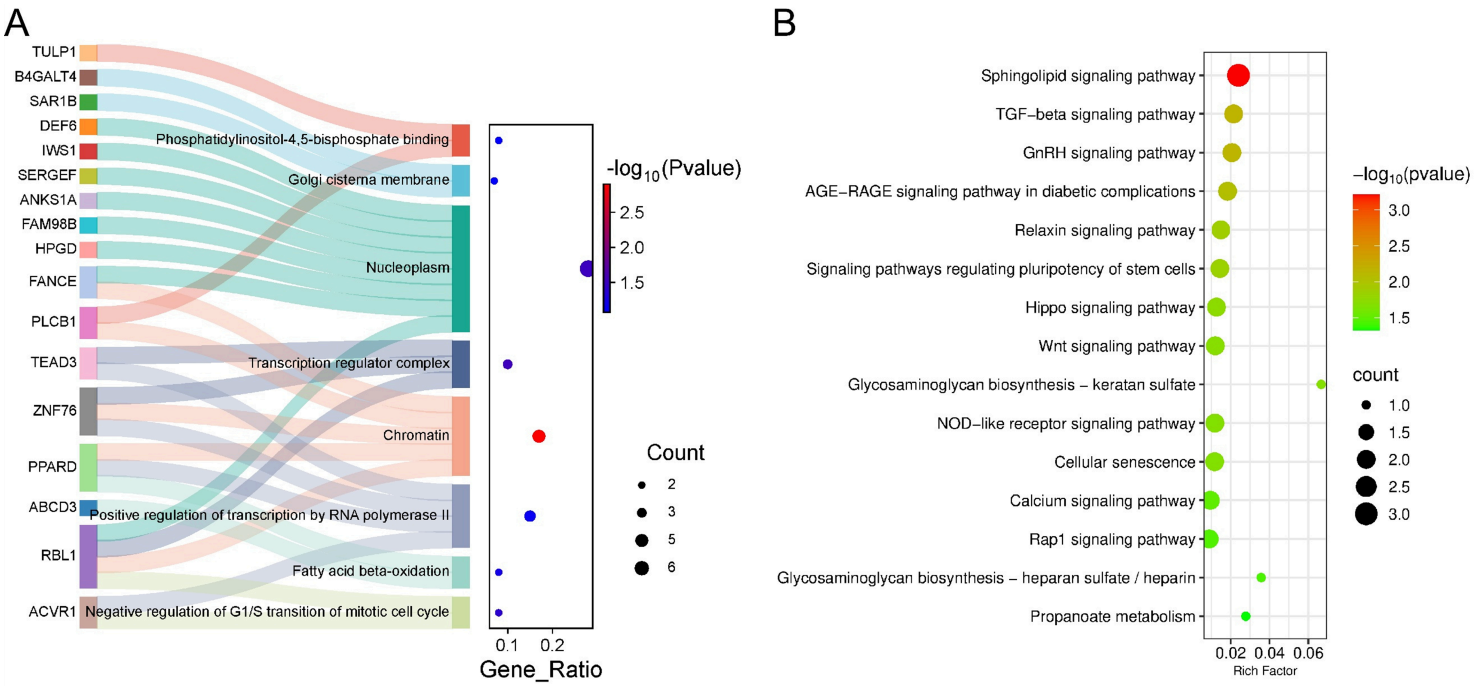


Figure 6

Functional enrichment results for candidate genes identified by GWAS in Tianzhu white yak. A: GO enrichment analysis; B: KEGG enrichment analysis.

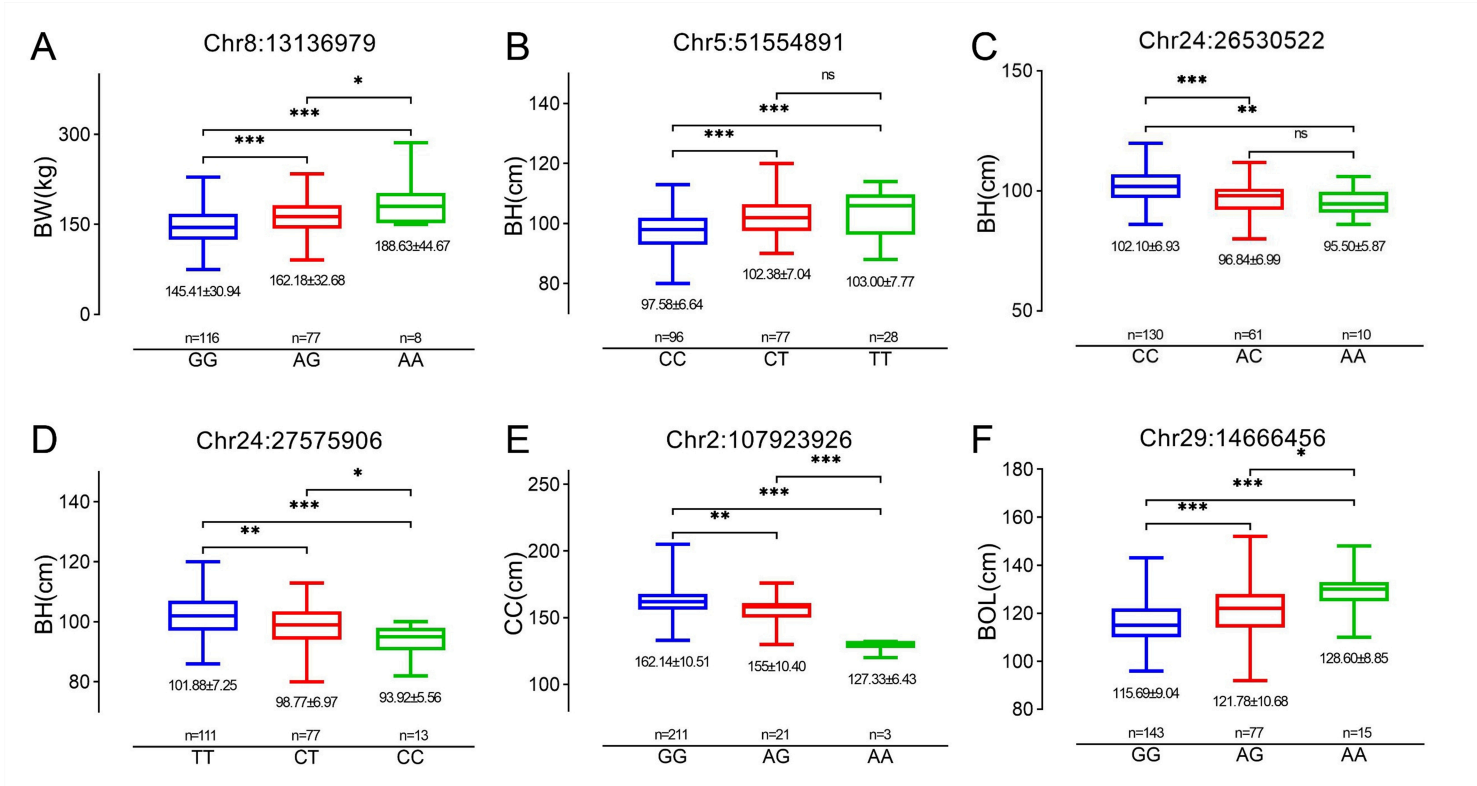


Figure 7

Genotyping results for key SNP loci in Tianzhu white yak. Figures A, B, C, D, E, and F show chr8:13136979, chr5:51554891, chr24:26530522, and chr24:27575906, respectively, showing the phenotypic values and their significant relationships among the three genotypes. ns indicates that the difference is not significant ($P \geq 0.05$), * indicates that $0.01 \leq P < 0.05$, ** indicates that $0.001 \leq P < 0.01$, and *** indicates $P < 0.001$.

Supplementary Files

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

- SupplementaryMaterial.zip