

RESEARCH

Open Access



Host genetic architecture and gut microbiota cooperatively regulate early growth in goats

Hui-Long Mou¹, Zi-Xuan Wang¹, Ming-Ding Zhang¹, Yao-Long Liu¹, Tao Ren¹, Peng-Cheng Ruan¹, Yan-Guo Han¹, Yan Zeng¹, Hao-Yuan Zhang¹, Jia-Guo Lu¹, Yang Zhang¹, Cheng-Li Liu², Yong-Fu Huang¹, Yong-Ju Zhao³, Zhong-Quan Zhao³, Xiao Pan⁴, Li-Ping Zhou⁴, Simone Ceccobelli^{5*} and Guang-Xin E.^{1*} 

Abstract

Early postnatal growth is a critical determinant of meat production efficiency and long-term genetic improvement in goats; however, the molecular mechanisms underlying individual variation in growth performance remain poorly understood. In this study, a total of 123 Hechuan white goats were included. First, a genome-wide association study (GWAS) for average daily gain (ADG) was performed using all 123 individuals. Subsequently, based on the coefficient of variation of ADG (CV=65.6%), an extreme phenotype sampling (EPS) strategy was applied to select 39 individuals with extreme growth phenotypes for subsequent metabolomic, microbiome, and integrated mGWAS analyses. The results showed that ADG approximately followed a normal distribution across the 123 goats. GWAS identified 22 loci significantly associated with ADG, mapping to genes including DLK1, NCAPG2, LCORL, CNTNAP2, and SLC8A1, which are involved in pathways related to skeletal muscle development, cell cycle regulation, ion transport, and immune function. Metabolomic profiling detected 1,589 putative metabolites, revealing differential enrichment of lipid, amino acid, and bile acid metabolic pathways between fast- and slow-growing goats. Gut microbiome analysis demonstrated that Christensenellaceae_R-7_group and Monoglobus were significantly enriched in fast-growing individuals, whereas Desulfovibrio was more abundant in slow-growing goats. Integrated mGWAS analysis further revealed extensive effects of host genetic variation on gut microbiota and fecal metabolites. Specifically, 11 bacterial genera were significantly associated with host genomic variants, among which Desulfovibrio exhibited the highest number of associated loci. Integration of multiple variant types consistently linked Desulfovibrio, Eubacterium_hallii_group, and Candidatus_Saccharimonas with genes such as ARHGAP24 and IGF2BP2. In addition, 14 metabolites were significantly associated with host genetic variants, with Lysopc(14:1(9Z)/0:0) and glycocholic acid showing the strongest associations. Notably, the peak signal for Lysopc was located within HMGA2. Collectively, these findings define a coordinated host genome–gut microbiota–metabolite network underlying early growth variation in goats and provide a mechanistic foundation for precision breeding and targeted nutritional strategies in goat production systems.

Keywords Goat, Growth performance, GWAS, Metabolomics, Gut microbiota

*Correspondence:
Simone Ceccobelli
s.ceccobelli@staff.univpm.it
Guang-Xin E.
eguangxin@126.com

¹College of Animal Science and Technology, Southwest University, Chongqing 400715, China

²School of Biological and Food Engineering, Chongqing Three Gorges University, Chongqing 404020, China

³State Key Laboratory of Resource Insects, Southwest University, Chongqing 400715, China

⁴Chongqing Hechuan Animal Husbandry Station, Chongqing 401520, China

⁵Department of Agricultural, Food and Environmental Sciences, Università Politecnica Delle Marche, Via Brecce Bianche 10, 60131 Ancona, Italy



© The Author(s) 2026. **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

The goat industry plays an important role in global animal production systems, supplying meat, milk, and fiber, while supporting income generation in many rural regions [1]. The Hechuan white goat is a representative indigenous breed in southwest China and has become closely linked to local agricultural development and rural revitalization initiatives. However, because breeding programs have historically relied on extensive management with limited structured genetic selection, the population shows considerable variation in productive performance and slow overall genetic progress. Improving the genetic basis of key economic traits is therefore necessary to enhance productivity and sustain competitiveness. Among these traits, average daily gain (ADG) is one of the most informative indicators of growth performance, as it directly influences feeding duration, feed efficiency, and production costs, making it a practical target for genetic improvement [2]. Growth rate during early life is particularly important, as it affects weaning weight, feed conversion efficiency, subsequent growth potential, carcass traits, and ultimately economic returns [3, 4]. For these reasons, understanding the biological mechanisms that regulate early growth has become a priority in ruminant breeding research.

With advances in high-throughput genotyping, genome-wide association studies (GWAS) have emerged as a pivotal approach for exploring the genetic architecture of complex traits. By associating dense single nucleotide polymorphism (SNP) markers with phenotypic variation, GWAS facilitates the identification of genomic regions and candidate genes influencing economically important traits [5, 6]. This approach has been successfully applied to traits such as milk yield in cattle [7], muscle development in pigs [8] and egg weight in chickens [9], illustrating its value for livestock improvement programs [10, 11]. In goats, GWAS has uncovered loci linked to horn type, meat quality and growth-related traits [12, 13]. In addition to GWAS, runs of homozygosity (ROH) analysis provides essential metrics regarding population structure, inbreeding levels, and historical selection. ROH-based approaches have been applied to identify genomic regions linked to growth, reproduction, and coat color in local goat populations, offering practical insights for both conservation and breeding strategies [14–16]. Nevertheless, growth is a multifaceted trait that cannot be fully explained by host genomic variation alone. Increasing evidence suggests that growth outcomes reflect interactions between host genetics and environmental factors, among which the gut microbiota plays a central role. The gut microbiota—comprising bacteria, fungi, viruses, and archaea—is often characterized as the host’s “second genome” due to its profound impact on systemic metabolism and physiological functions [17,

18]. Microbial community composition is influenced by host genotype, diet, and management conditions, and in turn affects nutrient utilization, immune function, and growth performance [19–21]. In ruminants, several studies have reported associations between gut microbial composition and growth-related traits [22–24], including differences in weaning weight observed in Hainan black goats [25].

Metabolomics provides an additional layer of information by capturing small-molecule metabolites that represent the downstream output of host and microbial activity. These metabolic profiles integrate the effects of genetic background, diet, and microbial metabolism, and have been increasingly used in livestock studies focusing on nutrition, health, and meat quality [26–29]. As a result, combining genomic, microbiome, and metabolomic data offers a systems-level perspective for investigating the mechanisms linking genotype to phenotype.

In the present study, an integrative multi-omics approach was applied to the Hechuan white goat population elucidating the biological mechanisms underlying early growth performance. Specifically, GWAS was conducted to identify loci associated with ADG during the early growth period and screened candidate SNPs and genes. In parallel, we characterized the fecal microbiota and metabolomic profiles, employing correlation and integrative association analyses to identify microbial taxa and metabolic pathways associated with average daily gain (ADG). Subsequently, mGWAS was utilized to link these ADG-associated gut microbial and metabolic signatures to host genetic variants. Ultimately, this study aims to provide mechanistic insights into early growth regulation, thereby supporting precision breeding and nutritional management strategies for goats.

Materials and methods

Experimental animals and sample collection

A total of 123 Hechuan white goats were selected as experimental subjects. Blood samples were collected for genomic DNA extraction from a conservation population at Chongqing Yinghe Agricultural Science and Technology Co., Ltd., China (106.171° E, 29.954° N). Before weaning, all kids were reared under natural suckling conditions with their dams. After weaning at 60 days of age, they were transferred to an indoor housing system and provided ad libitum access to a standardized basal diet and high-quality forage.

Body weight (BW) was recorded at 60 and 90 days of age. Average daily gain (ADG, g/day) for the early growth period (60–90 days) was calculated as: $ADG = (BW_{90} - BW_{60}) / 30$. The normality of ADG distribution was assessed using the Shapiro-Wilk test and visualized via histogram. Pearson correlation analysis was performed between ADG (60–90 days) and subsequent BW at 2–3,

6, and 9 months of age. Statistical significance was set at $P < 0.05$. All statistical analyses and visualizations (scatter plots with linear regression) were conducted in R (v4.4.1) using the ggplot2 and ggpubr packages.

To capture biological extremes for integrative multi-omics analysis, a subset of 39 kids was selected from the initial cohort ($n = 123$). Selection was primarily based on sex-specific ADG rankings. Additional criteria included high sample quality, complete metadata, and availability for matched multi-omics profiling. The selected animals were classified into four groups: Fast-growing males (M_F, $n = 10$, ADG > 120 g/day), Fast-growing females (F_F, $n = 10$, ADG > 110 g/day), Slow-growing males (M_S, $n = 10$, ADG < 40 g/day), Slow-growing females (F_S, $n = 9$, ADG < 35 g/day).

After overnight fasting, blood samples (approximately 5 mL) were collected via jugular venipuncture from all 123 animals, immediately aliquoted, and stored at -80°C for genomic DNA extraction. From the selected subset of 39 kids exhibiting extreme growth phenotypes, matched fresh fecal samples were aseptically obtained by rectal collection directly into sterile cryovials, snap-frozen in liquid nitrogen within 5 min, and maintained at -80°C for subsequent microbiome profiling (16 S rRNA gene sequencing) and metabolomic analysis. Prior to sampling, all animals were confirmed to be clinically healthy, with no history of antibiotic administration or clinical signs of infectious disease. All animal procedures were approved by the Laboratory Animal Ethics Committee of Southwest University (Approval No. IACUC-20240710-06) and performed in strict accordance with the committee's guidelines for the care and use of laboratory animals.

Whole-genome sequencing and data processing

Genomic DNA was extracted from whole blood using the TaKaRa Universal DNA Extraction Kit following the manufacture instructions, and DNA concentration was measured using a NanoDrop™ One spectrophotometer (Thermo Fisher Scientific, USA). DNA integrity was assessed by 1% agarose gel electrophoresis. Sequencing libraries were constructed using the Annoroad® Universal DNA Library Prep Kit v2.0 (Illumina®, USA) and sequenced on a DNBSEQ-T7 platform (BGI, China) to generate whole-genome data.

Raw reads were quality filtered using Fastp (v0.20.1) [30] to obtain clean reads with the following parameters: sliding-window trimming (4 bp, mean quality $\geq Q15$), 3 bp trimming from both ends, and minimum read length of 40 bp. Filtered reads were aligned to the goat reference genome (*Capra hircus* ARS1.2) using BWA-MEM2 (v2.2.1) [31] to generate BAM files. Subsequently, the resulting BAM files were sorted, PCR duplicates were removed, and the final alignments were indexed using Samtools (version 1.17) [32].

Single nucleotide polymorphisms (SNPs) and small insertions/deletions (InDels) were identified using the GATK (v4.5.0.0) [33] 'Best Practices' pipeline with base quality recalibration (BQSR). Raw variants were subjected to hard filtering using the following thresholds: for SNPs, QD < 2.0, FS > 60.0, MQ < 40.0, and SOR > 3.0; for indels, QD < 2.0, FS > 200.0, and SOR > 10.0. Variants tagged as low-quality (LowQual) were excluded. For downstream analyses, the variant set was further filtered to retain only those with a genotype call rate > 90% and a minor allele frequency (MAF) > 0.05.

Copy number variations (CNVs) were detected using CNVcaller [34] with a sliding window size of 1,000 bp and a step size of 500 bp. CNV regions (CNVRs) were determined based on copy number status, window distribution, and correlations between adjacent windows. Finally, CNVRs not anchored to specific chromosomes were excluded from downstream analyses.

Structural variants (SVs) were detected using a consensus approach across three independent callers: Manta (v1.6.0) [35] Lumpy (v0.2.13) [36], and Delly (v1.1.7) [37]. All methods were performed using default parameters. High-confidence SVs were defined as those identified by at least two callers and merged using SURVIVOR (v1.0.3) [38] with a maximum breakpoint distance of 1000 bp, consistent SV type and strand orientation, and a minimum size of 50 bp.

The final, high-confidence variant sets (SNPs, InDels, CNVs, and SVs) were used as input for subsequent genome-wide association studies (GWAS).

To integrate CNVs and SVs into the GWAS framework, CNVs were encoded as ordinal variables based on copy number status: 0 (homozygous deletion), 1 (heterozygous deletion or duplication), and 2 (homozygous duplication). For complex multi-allelic CNV states, a letter-based scheme (A–J) was applied. SVs were encoded as binary variables: 0 (absence) and 1 (presence), with specific letter assignments for each variant type (inversion, deletion, duplication, insertion, translocation). The complete replacement rules are provided in Supplementary Table S1.

Genome-wide association analysis and candidate gene annotation

A genome-wide association study (GWAS) for average daily gain was conducted using the multi-locus mixed model (MLMM) implemented in GAPIT3 [39]. The model included sex as a fixed effect and incorporated a genomic relationship matrix (derived from autosomal genetic variants) as a random effect to account for population stratification. The MLMM procedure iteratively identified significant Genetic variants as covariates until no new Genetic variants met the genome-wide significance threshold (Bonferroni-corrected $P < 0.05/N$, where

N is the number of tested Genetic variants). Protein-coding genes within ± 0.5 Mb of significant Genetic variants were defined as positional candidates. Candidate genes were annotated using ANNOVAR [40]. Gene Ontology (GO) biological process terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were carried out using KOBAS (<http://kobas.cbi.pku.edu.cn/>). Enrichment significance was tested using the hypergeometric test, and *P*-values were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) method. Terms and pathways with an FDR-adjusted *P*-value < 0.05 were considered statistically significant.

Metabolomic analysis and data processing

Fecal metabolomic profiling was performed using an ultra-high-performance liquid chromatography-Q Exactive HF-X system (Thermo Fisher Scientific). Metabolites were extracted from 50 mg of feces using 80% methanol containing 2-chloro-L-phenylalanine as an internal standard, with quality control samples prepared from pooled aliquots of all samples. Chromatographic separation was carried out on an ACQUITY UPLC HSS T3 column (100×2.1 mm, $1.8 \mu\text{m}$) with a mobile phase consisting of (A) 5% acetonitrile containing 0.1% formic acid and (B) acetonitrile/isopropanol (1:1, v/v) containing 0.1% formic acid at a flow rate of 0.40 mL/min. Mass spectrometry data were acquired in both positive and negative electrospray ionization modes (*m/z* 70–1050) with data-dependent MS/MS acquisition.

After peak detection and alignment, features with $> 20\%$ missing values within any experimental group were removed. Missing values were imputed using the half-minimum method. Data were normalized by total ion current, and features showing relative standard deviation $> 30\%$ in quality control samples were excluded before log₁₀ transformation. Metabolite annotation was performed by matching accurate mass (tolerance ± 5 ppm) and MS/MS spectra against the Human Metabolome Database (HMDB, <http://www.hmdb.ca>) and Metlin database (<https://metlin.scripps.edu>). Orthogonal partial least squares discriminant analysis (OPLS-DA) was conducted using the ropls R package, with model validity assessed via cross-validation. Differential metabolites were identified based on variable importance in projection (VIP) > 1 , fold change (FC) > 2 , and false discovery rate-adjusted *P* < 0.05 . Pathway enrichment analysis was performed using over-representation analysis against the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, with significance defined as false discovery rate < 0.05 .

Microbiome analysis and data processing

Fecal microbial DNA was extracted with the E.Z.N.A.[®] Soil DNA Kit (Omega Bio-tek, USA) following the manufacturer's protocol. DNA integrity was evaluated by 1% agarose gel electrophoresis and concentrations were measured using a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA). The V3–V4 hypervariable region of the bacterial 16S rRNA gene was amplified using primers 338F_806R (F: 5'-ACTCCTACGGGAGG CAGCAG-3'; R: 5'-GGACTACHVGGGTWTCTAAT-3') [41] in a 20 μL PCR system containing 10 ng of template DNA, 4 μL of 5 $\times$ TransStart FastPfu buffer, 2 μL of 2.5 mM dNTPs, 0.8 μL of each primer (5 μM), and 0.4 μL of TransStart FastPfu DNA polymerase. PCR cycling conditions were: 95 $^{\circ}\text{C}$ for 3 min; 27 cycles of 95 $^{\circ}\text{C}$ for 30 s, 55 $^{\circ}\text{C}$ for 30 s, 72 $^{\circ}\text{C}$ for 30 s; final extension at 72 $^{\circ}\text{C}$ for 10 min. Amplicons were purified with a gel extraction kit (YuHua, China) and quantified with a Qubit 4.0 Fluorometer (Thermo Fisher Scientific). Libraries were prepared with the NEXTFLEX[®] Rapid DNA-Seq Kit (Bioo Scientific, USA) and sequenced on an Illumina NextSeq 2000 platform (2×150 bp).

Raw reads were quality-filtered with Fastp (v0.19.6) to trim low-quality bases (quality score < 20) in a sliding 50-bp window and discard reads shorter than 50 bp or containing ambiguous nucleotides. Paired-end reads were merged using FLASH (v1.2.11; minimum overlap 10 bp, maximum mismatch ratio 0.2). Operational taxonomic units (OTUs) were clustered at 97% similarity using the UPARSE pipeline, and chimeric sequences were removed with UCHIME. Alpha-diversity indices (Chao1, Shannon) were calculated in Mothur [42] and compared between groups using the Wilcoxon rank-sum test. Beta-diversity was assessed with Bray–Curtis distances, visualized by principal coordinates analysis (PCoA), and tested by permutational multivariate analysis of variance (PERMANOVA; 999 permutations) at *P* < 0.05 . Differentially abundant taxa were identified with linear discriminant analysis effect size (LEfSe) using a logarithmic LDA score threshold > 2.0 .

Integrative host–microbiome–metabolome analysis

Genus-level microbial taxa exhibiting significant differential relative abundance and metabolites with VIP > 1 , *P* < 0.05 , and FC > 1.5 were treated as quantitative traits for microbial/metabolite GWAS (mGWAS). mGWAS was performed using the MLM implemented in GAPIT3 with Bonferroni correction (*P* $< 0.05/N$, where *N* represents the number of effective Genetic variants). Genes located within ± 0.5 Mb of the significant loci were annotated to identify host candidate genes associated with specific microbial taxa or metabolites.

Results

Early growth traits and body weight correlations in hechuan white goats

Analysis of average daily gain (ADG) in 123 Hechuan white goat kids revealed an approximately normal distribution across the population (Fig. 1A), indicating that this trait was well suited for subsequent GWAS. Correlation analyses demonstrated that the average daily gain for samples aged 2–3 months showed a significant positive correlation with their body weights at 6 and 9 months of age. (Fig. 1B–D). These results indicate that growth performance during the early post-weaning period is likely linked to later body weight, suggesting that ADG may serve as a reliable indicator for early selection in breeding programs.

Association of SNPs, InDels, CNVs, and SVs with early post-weaning ADG in Hechuan white goats

GWAS identified genetic variants significantly associated with ADG during the early growth phase in Hechuan white goats. To establish the genetic basis for this analysis, we performed whole-genome resequencing. Following stringent quality control and variant calling, a total of 3,637,201 high-confidence variants were detected across the genome. This comprehensive variant set comprised 2,358,515 single nucleotide polymorphisms (SNPs),

1,206,297 insertions and deletions (InDels), 59,164 copy number variation (CNV) regions, and 13,225 structural variants (SVs), with an average sequencing data yield of approximately 18.29 Gb per individual (Supplementary Table S2). Subsequent GWAS results based on these respective variant types are detailed below.

- **SNP Analysis:** Four SNPs were significantly associated with ADG and located on chromosomes 1, 2, 13, and 25 (Fig. 2A). Among them, chr2:59,803,820 was located within an intronic region of *CNTNAP5* ($P = 6.85 \times 10^{-20}$, MAF = 0.446), and chr25:7,550,971 was located within an intron of *PMM2* ($P = 2.80 \times 10^{-18}$, MAF = 0.314). The remaining two significant SNPs were mapped to chromosomes 1 and 13.
- **CNV Analysis:** Eight CNVRs were significantly associated with ADG and distributed across chromosomes 4, 7, 8, 11, 17, 21, and 22 (Fig. 2B). Notably, chr4:522,001–532,000 encompassed *NCAPG2* ($P = 1.32 \times 10^{-11}$, MAF = 0.430), while chr21:64,976,001–64,978,000 overlapped *DLK1* ($P = 7.53 \times 10^{-16}$, MAF = 0.145). Additional CNVRs contained genes such as *TLR4* and *SLC8A1*.
- **InDel Analysis:** Five InDels were significantly associated with ADG and located on chromosomes

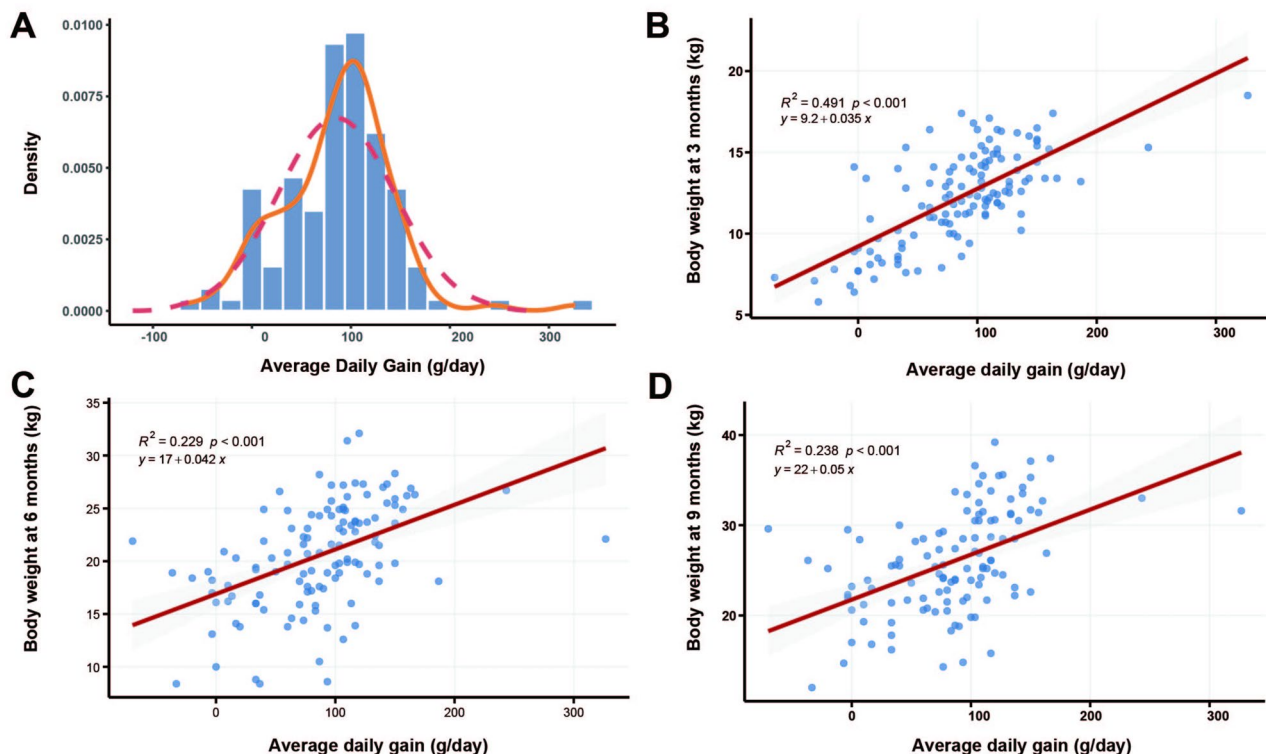


Fig. 1 Distribution of average daily gain and its correlations with body weight traits in Hechuan white goat kids during 60–90 days of age. **(A)** Density curve of average daily gain distribution. **(B–D)** Scatter plots of correlations between average daily gain and body weight at **(B)** 3, **(C)** 6, and **(D)** 9 months of age

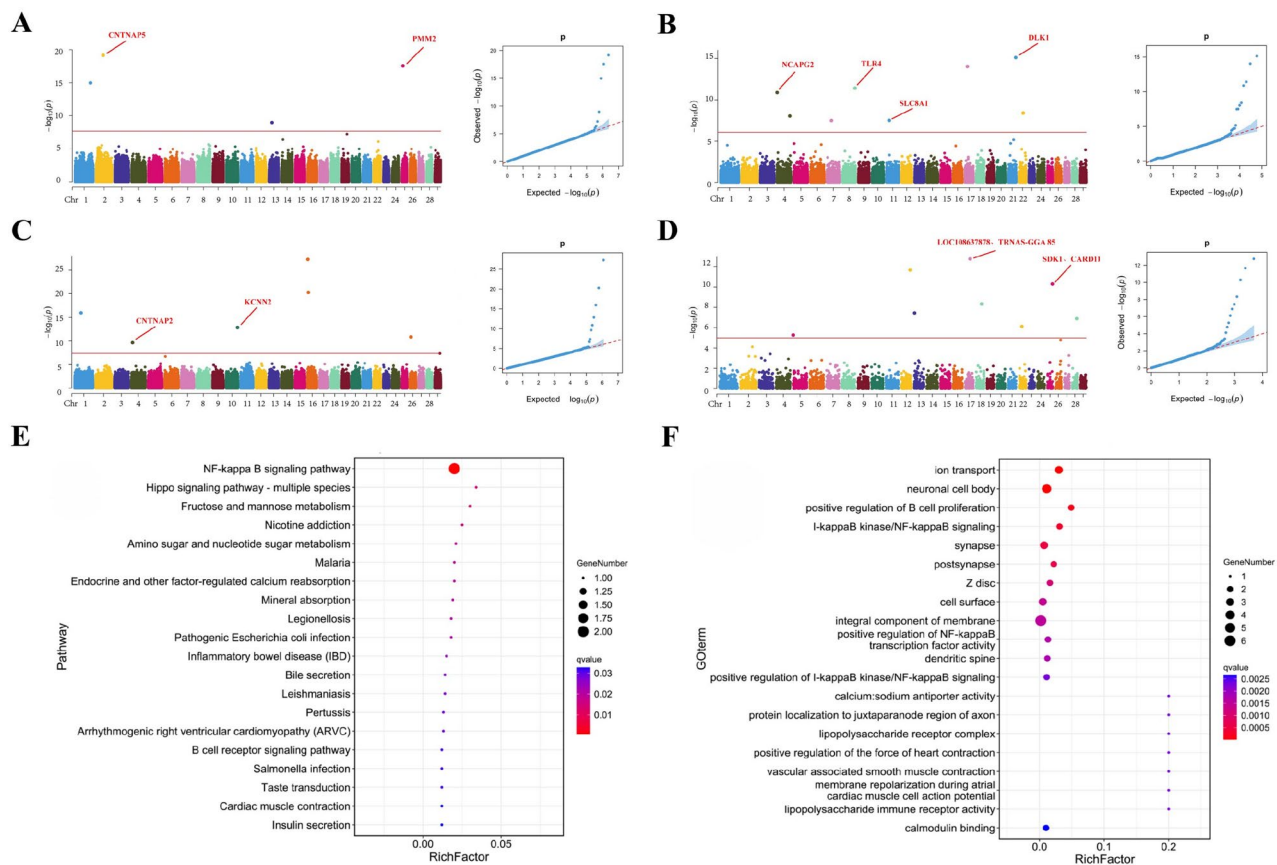


Fig. 2 Genome-wide association and functional enrichment analysis for early growth traits in Hechuan white goats. Manhattan plots showing genomic variants significantly associated with ADG, including SNPs(A), CNVRs(B), InDels(C), and SVs(D). (E) Top 20 significantly enriched KEGG pathways. (F) Top 20 significantly enriched GO terms

1, 4, 10, 16, and 26 (Fig. 2C). Notably, chr4:10,470,783 was located within an intron of *CNTNAP2* ($P = 2.21 \times 10^{-10}$, MAF = 0.236), and chr10:97,991,557 mapped to an intronic region of *KCNN2* ($P = 1.38 \times 10^{-13}$, MAF = 0.186). The remaining InDels were distributed across chromosomes 1, 16, and 26.

- SV Analysis: Five candidate SV regions associated with ADG were identified on chromosomes 12, 13, 17, 18, and 25 (Fig. 2D). For example, chr25:41,188,671 overlapped *SDK1* and *CARD11* ($P = 1.91 \times 10^{-6}$, MAF = 0.241), whereas chr17:42,399,750 was located between *LOC108637878* and *TRNAS-GGA_85* ($P = 2.42 \times 10^{-9}$, MAF = 0.417).

Taken together, these results (Supplementary Table S3) indicate that multiple forms of genomic variation contribute to early growth performance, highlighting a complex genetic architecture underlying ADG.

Furthermore, functional enrichment analysis was performed to explore the biological roles of the candidate genes. KEGG pathway enrichment analysis revealed that

a total of 8 differentially candidate genes were mapped to 55 KEGG pathways, among which 15 pathways were significantly enriched. These mainly included the NF- κ B signaling pathway, Hippo signaling pathway-multiple species, as well as pathways related to physiological regulation, such as insulin secretion and cardiac muscle contraction. In addition, the Hippo signaling pathway (involving genes such as *WTIP*) and the amino sugar and nucleotide sugar metabolism pathway (involving genes such as *PMM2*) were also significantly enriched (Fig. 2E, Supplementary Table S4).

GO enrichment analysis indicated that 15 candidate genes were associated with 289 GO terms, among which 214 terms were significantly enriched. The significantly enriched GO terms primarily involved biological processes, molecular functions, and cellular components related to ion channel regulation, nervous system function, and immune response. Notable specific terms included ion transport, neuronal cell body, and regulation of cardiac muscle contraction by calcium ion signaling (Fig. 2F, Supplementary Table S5).

Differential fecal metabolites reflect early growth rate in Hechuan white goats

Metabolite annotation was performed based on two confidence levels: B(i) (highest accuracy, exact MS/MS match using an in-house library) and B(ii) (exact MS/MS match using in silico spectra from public/theoretical databases). Results were filtered to retain only those with a B(i) score ≥ 35 or a B(ii) score ≥ 40 . Based on the aforementioned criteria, untargeted metabolomic profiling identified a total of 1,589 tentative metabolites. OPLS-DA analysis revealed clear separation between fast- and slow-growing males (M_F vs. M_S) and females (F_F vs. F_S) (Fig. 3A, B), indicating that early growth rate substantially influences the metabolic profile. In males, 252 metabolites exhibited significant differential abundance (111 upregulated, 141 downregulated), whereas 208 differential metabolites were detected in females (119 upregulated, 88 downregulated) (Fig. 3C, D, Supplementary Table S6). HMDB classification indicated that these metabolites mainly belonged to alkaloids, lipids and lipid-like molecules, organic acids and derivatives, amino acids, and phenylpropanoids. A total of 82 metabolites were shared between sexes, whereas the remaining metabolites exhibited sex-specific patterns (Fig. 3E).

KEGG pathway enrichment revealed that differential metabolites in males (49 metabolites) mapped to 100 pathways, predominantly metabolic pathways (62 pathways). Secondary pathway analysis indicated enrichment in lipid metabolism, amino acid metabolism, secondary metabolite biosynthesis, and digestive system pathways (Fig. 3F). After FDR correction, 12 pathways were remained significant ($P_{\text{adjust}} < 0.05$), including choline metabolism, cholesterol metabolism, and fat digestion and absorption (Fig. 3H). In females, 48 differential metabolites were enriched in 109 pathways, mainly involving secondary metabolite biosynthesis, amino acid metabolism, and membrane transport (Fig. 3G). FDR correction identified 12 significantly enriched pathways ($P_{\text{adjust}} < 0.05$), including flavonoid degradation, ABC transporters, and pyrimidine metabolism (Fig. 3I, Supplementary Table S7).

Gut microbiome profiling in relation to early growth of goat

Alpha- and beta-diversity analyses were conducted to evaluate gut microbial differences between growth groups (Table 1). At the genus level, fast-growing males (M_F) exhibited higher species richness (Sobs index) and a trend toward greater community diversity (Shannon index) compared to slow-growing males (M_S), although differences were not statistically significant. In females, fast-growing individuals (F_F) showed a significantly higher Shannon indices and a lower Simpson indices

than slow-growing females (F_S) ($P < 0.05$), indicating elevated diversity and more even microbial communities.

Beta-diversity analysis based on Bray-Curtis distances revealed limited separation between M_F and M_S groups (Fig. 4A), whereas F_F and F_S displayed clear segregation along principal coordinates, with F_F samples forming a tighter cluster (Fig. 4B). Microbial composition analysis at the genus level showed that all groups were dominated by *Christensenellaceae_R-7_group*, *unclassified_f_Lachnospiraceae*, *norank_o_Clostridia_UCG-014*, and *UCG-005*, with relative abundances ranging from 7.79 to 12.16%, 8.39–11.99%, 5.14–8.40%, and 8.43–10.22%, respectively (Fig. 4C, D).

Differential abundance analysis using the Wilcoxon rank-sum test identified 10 genera that differed significantly between the M_F and M_S groups. Genera such as *Marvinbryantia*, *norank_f_Flavobacteriaceae*, *norank_f_Eubacterium_coprostanoligenes_group*, *Acetitomaculum*, *dga-11_gut_group*, *Dorea*, and *Lachnospiraceae_NK3A20_group* were enriched in M_F ($P < 0.05$), whereas *Subdoligranulum*, *Eubacterium_hallii_group*, and *Candidatus_Saccharimonas* were more abundant in M_S (Fig. 4E, Supplementary Table S8). In females, 10 genera exhibited significant abundance changes between F_F and F_S. Genera including *norank_f_Ruminococcaceae*, *Family_XIII_AD3011_group*, *norank_f_UCG-010*, *Eubacterium_coprostanoligenes_group*, *Rikenellaceae_RC9_gut_group*, *Monoglobus*, and *UCG-009* were enriched in F_F ($P < 0.05$), whereas *norank_f_Muribaculaceae*, *Blautia*, and *Desulfovibrio* were more abundant in F_S ($P < 0.05$) (Fig. 4F, Supplementary Table S9).

Host genome–microbiome and metabolome associations via mGWAS

Based on the 16S rRNA gene sequencing of the gut microbiota from 39 extreme individuals to elucidate the regulatory basis of early growth traits.

Microbiome–host genome associations

To investigate the interplay between gut microbiota and host genetic variation, metagenome-wide association studies (mGWAS) were performed using early growth-differentiated individuals (Fig. 5A, Supplementary Table S10).

- SNP Analysis: identified 11 bacterial genera with significant host-genome associations. *Desulfovibrio* showed the highest number of associated loci (8 SNPs), distributed across seven chromosomes (5, 6, 10, 11, 12, 21, and 29). These SNPs implicated genes involved in inflammatory and metabolic regulation, including *ARHGAP24*, *DPH6*, *FANCL*, and *ATP12A*. *Eubacterium_hallii_group* was associated with 4 SNPs located on chromosomes 1, 6, 11, and

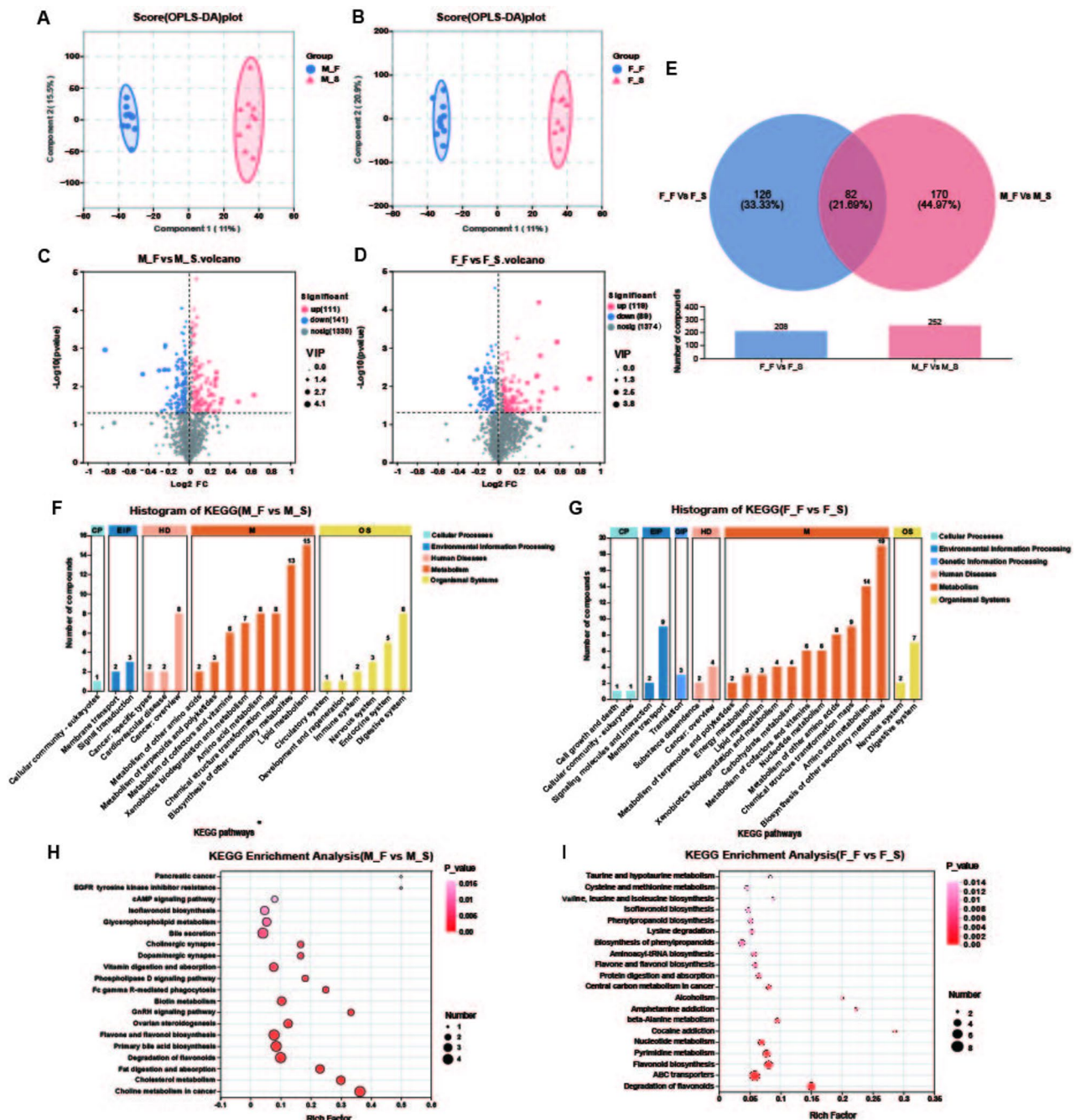


Fig. 3 Metabolic profiling and pathway enrichment analysis associated with growth in Hechuan white goats. **(A)** OPLS-DA score plot of the male group (M_F vs. M_S). **(B)** OPLS-DA score plot of the female group (F_F vs. F_S). **(C)** Volcano plot of differential metabolites in the male group. **(D)** Volcano plot of differential metabolites in the female group. **(E)** Venn diagram of shared and sex-specific metabolites. **(F)** KEGG level 2 pathway classification in the male group. **(G)** KEGG level 2 pathway classification in the female group. **(H)** Significantly enriched KEGG pathways in the male group. **(I)** Significantly enriched KEGG pathways in the female group. M_F: fast-growing males, M_S: slow-growing males, F_F: fast-growing females, F_S: slow-growing females

23, mapping to *LEKR1* and *NRXN1*. *Subdoligranulum* was linked to 3 SNPs spanning chromosomes 1, 7, and 14, which included *HUNK*.

- InDel Analysis: identified significant associations for 9 bacterial genera. *Eubacterium hallii* group exhibited the highest number of InDels (6),

distributed across chromosomes 2, 4, 9, 10, 14, and 24, involving genes *DPP6* and *IQGAP2*. *Monoglobus* and *norank_f_Ruminococcaceae* each carried 4 InDels, with *Monoglobus* variants mapping to *PLXNA4* and *UNC79*. *Desulfovibrio* and

Table 1 Microbial community α diversity Index in different groups of Hechuan White Goats

Items	Males group		P-values	Females group		P-values
	M_F	M_S		F_F	F_S	
Shannon	3.221	3.074	0.21	3.400	3.232	0.04
Simpson	0.069	0.098	0.13	0.055	0.075	0.006
Sobs	125.8	119.3	0.25	131.3	126.6	0.48
Shannoneven	0.021	0.643	0.26	0.020	0.667	0.01
Simpson even	0.018	0.102	0.23	0.143	0.111	0.01

Note: M_F: fast-growing males, M_S: slow-growing males, F_F: fast-growing females, F_S: slow-growing females

Marvinbryantia were associated with 3 InDels each, the latter including gene *MALTI*.

- CNV Analysis: revealed significant copy-number variations associated with 11 bacterial genera. *Candidatus_Saccharimonas* exhibited the most extensive CNV associations (8 regions), spanning chromosomes 1, 4, 5, 11, 14, 16, and 22, and implicating key genes including *IGF2BP2*, *ATXN10* and *DHX30*. *norank_f__Ruminococcaceae* was linked to 7 CNV regions across chromosomes 1, 4, 9, 11, 13, and 23, involving genes such as *STARD3NL* and *CADM2*. *norank_f__UCG-010* showed 5 CNV hits on chromosomes 7, 14, 22, and 29, with associations near *ATP6V0D2* and *ZCWPW2*.
- SV Analysis: identified structural variants associated with 13 bacterial genera. *norank_f__Flavobacteriaceae* exhibited the most extensive SV associations (8 regions), spanning chromosomes 1, 5, 6, 11, 15, and 16, and implicating genes such as *NAALADL2* and *DDX10*. *Desulfovibrio* was linked to 7 SV regions across chromosomes 1, 7, 17, 18, 21, and 24, including *TOMM70* and *ARHGAP24*. *norank_f__UCG-010* showed 6 SV hits on chromosomes 1, 2, 6, and 18, involving genes *TECRL* and *LNK1*.

Integration across the four types of genomic variation (SNP, InDel, CNV, and SV) indicated that *Desulfovibrio*, *Eubacterium_hallii_group*, and *Candidatus_Saccharimonas* consistently exhibit strong host genomic associations, implicating genes such as *ARHGAP24*, *DPH6*, *FANCL*, *ATP12A*, *LEKR1*, *NRXN1*, *HUNK*, and *IGF2BP2*, providing comprehensive evidence of host–microbe interactions.

Host genome–fecal metabolite associations

Metagenome-wide association analysis revealed extensive correlations between host genetic variants and fecal metabolites. Fourteen metabolites exhibited significant associations across different types of genomic variation. SNPs displayed the most widespread associations, with 8 metabolites linked to 29 significant loci; InDels involved 2 metabolites and 6 loci; CNVs encompassed 7 metabolites across 21 regions; and SVs were associated with 9

metabolites across 22 regions (Fig. 5B, Supplementary Table S11).

- SNP Analysis: Lysopc(14:1(9Z)/0:0) exhibited the strongest and highest number of SNP associations (7 loci), with the peak signal located within *HMGA2*. The flavonoid Luteolin also showed an extremely strong association, mapping to *PECAM1*. Choline was associated with one intergenic SNP on chr18. Taurocholic Acid to the *SCARA5* and *MGAT4D*, while Glycocholic Acid was associated with *CTNNA2*. Additional significant SNP associations were identified for Sebacoyl-L-Carnitine (linked to *SLC2A13*), Eriodictyol (linked to *DEPTOR*, *SCN3B*, *TMEM117*), (+)-Taxifolin (linked to *GRM7*, *PECAM1*, and *SNCAIP*).
- InDel Analysis: phosphatidic acid Pa(A-13:0/I-15:0) was associated with three InDel loci, with the strongest signal located within the glutamate receptor gene *GRIK2*, *TTC7B*, and *KCTD12*. The cholesteryl ester Ce (18:3(9Z,12Z,15Z)) was linked to three InDel loci, implicating or located near genes *MIR137*, *RASSF2*, *CYB5A*, and *TIMM21*.
- CNV Analysis: The bile acid taurocholic acid was associated with 4 CNV regions, with the strongest signal overlapping the *MIPOL1* gene. Another bile acid, glycocholic acid, displayed the highest number of CNV associations (7 regions), spanning loci near genes including *ADAMTS20*, *FANCL*, *IRX1*, and *NUB1*. The flavonoid apigenin was linked to 3 CNV regions, with one located within *PFKFB3*. Additional significant CNV associations were found for Dg (8:0/10:0/0:0), chrysoeriol, phloretin, and choline.
- SV Analysis: The strongest SV signal for Lysopc(14:1(9Z)/0:0) was located near the lipase gene *LIP1*, with additional associations implicating *ROBO2*, *TRPC3*. For glycocholic acid, significant SVs were found near *ASTN2*, *TAC1/SDHAF3*, and *HECW1*. Another bile acid, taurocholic acid, was linked to 1 SV region. The flavonoid metabolites eriodictyol and (+)-taxifolin shared a common SV association (chr12:15,509,156). Additional significant SV associations were identified for choline, sebacoyl-

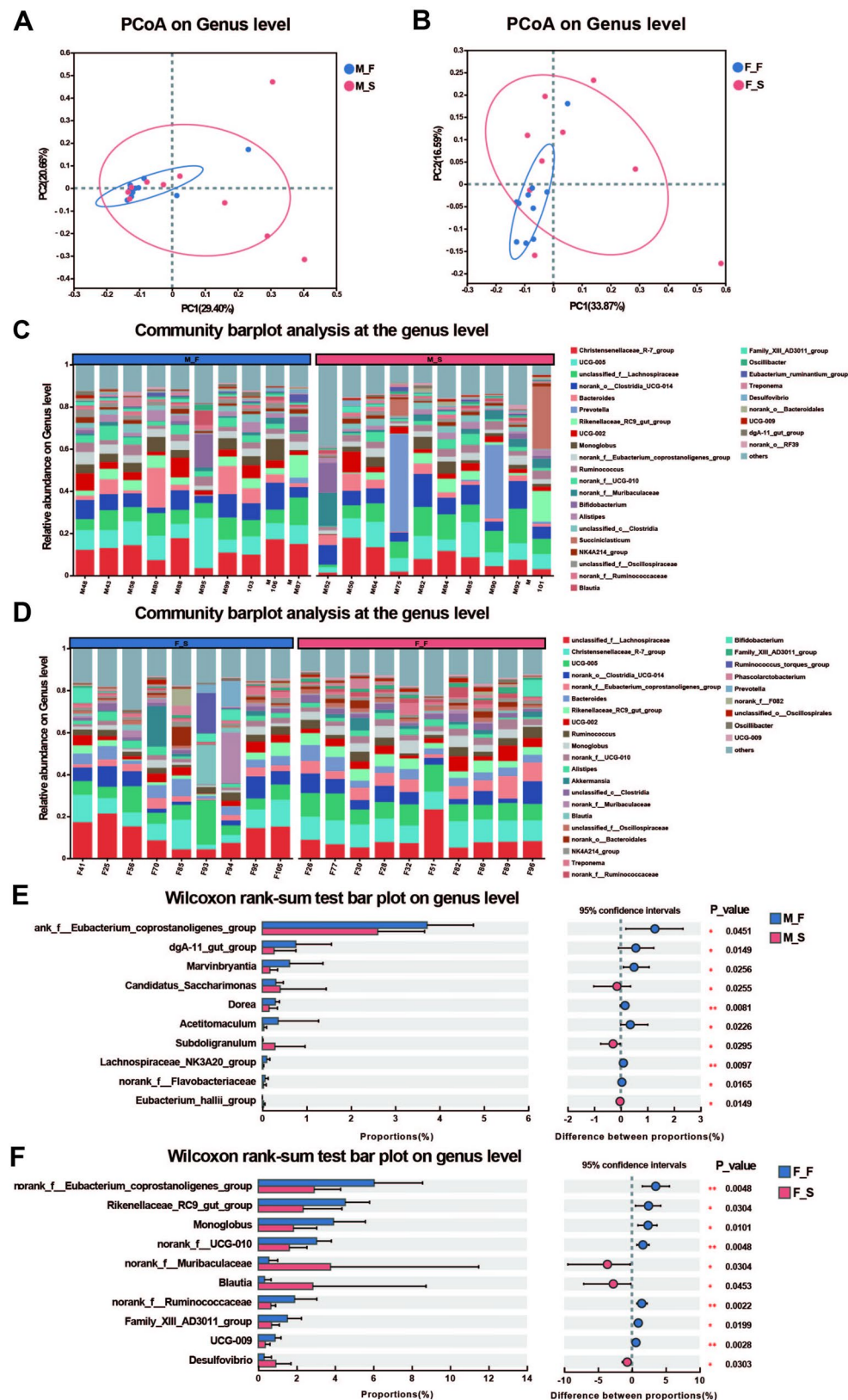


Fig. 4 Analysis of fecal microbial composition and differential genera in Hechuan white goats with divergent growth rates. **(A)** Beta diversity analysis of the male group (M_F vs. M_S). **(B)** Beta diversity analysis of the female group (F_F vs. F_S). **(C)** Composition of dominant bacterial genera in the male group. **(D)** Composition of dominant bacterial genera in the female group. **(E)** Significantly differential genera in the male group. **(F)** Significantly differential genera in the female group. M_F: fast-growing males, M_S: slow-growing males, F_F: fast-growing females, F_S: slow-growing females

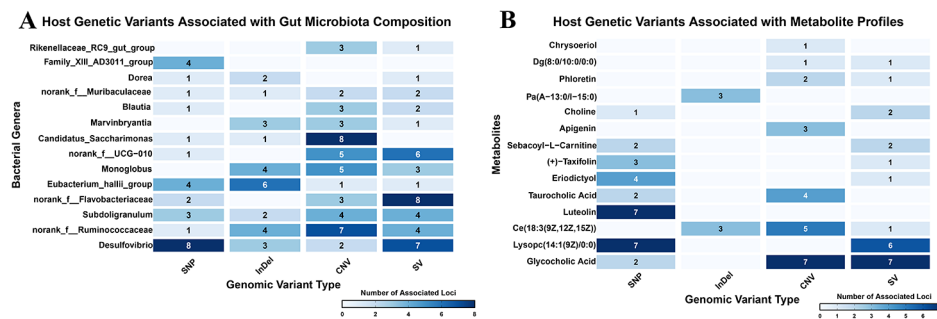


Fig. 5 Host genetic associations with gut microbiota and metabolites in Hechuan white goats. **(A)** Host genetic associations with gut microbiota composition in Hechuan white goats. **(B)** Host genetic associations with Metabolites in Hechuan white goats

L-carnitine, Dg (8:0/10:0/0:0), phloretin, and Ce(18:3(9Z,12Z,15Z)).

Integration across all variant types demonstrated that Lysopc (14:1(9Z)/0:0) and glycocholic acid were linked to the greatest number of host genomic loci, predominantly through SNPs and SVs. These associations spanned both intragenic and intergenic regions, revealing a multi-layered host genomic regulation of gut metabolites.

Discussion

In this study, we systematically elucidated the genetic basis and host–microbiome interactions underlying early growth in Hechuan white goats by integrating multi-omics data, including genomics, metabolomics, and microbiome profiling. Early average daily gain (ADG) exhibited an approximately normal distribution across the population and was significantly positively correlated with later body weight traits, indicating that ADG serves as a reliable predictor of individual post-weaning growth potential and provides a theoretical foundation for early selection and genetic improvement strategies.

GWAS identified multiple candidate genes across various chromosomes significantly associated with ADG, encompassing key biological processes such as muscle development, skeletal growth, energy metabolism, and immune regulation. Among them, *DLK1*, a critical imprinted gene, plays a central role in muscle and adipose tissue development. Previous studies have shown that *DLK1* promotes myogenic differentiation while inhibiting adipogenesis via the Notch signaling pathway [43, 44] and enhances muscle growth through the *ACVR2B*/Myostatin pathway [45]. High *DLK1* expression has been consistently linked to muscle fiber hypertrophy and accelerated growth in cattle, pigs, and sheep [46]. The significant association observed in this study further supports its pivotal role in early muscle development and growth in goats.

The *NCAPG2–LCORL* locus has been repeatedly reported as a major determinant of body size and growth traits in livestock. *NCAPG2*, a regulator of the

chromosome condensation complex, promotes cell proliferation and tissue growth through the *STAT3/c-MYC* signaling pathway [47, 48], while *LCORL* participates in lipid metabolism and skeletal muscle differentiation [49]. Variants in this locus are strongly associated with body length, height, and carcass weight [50, 51]. Collectively, our results suggest that *DLK1* and *NCAPG2–LCORL* may form a coordinated “muscle–skeletal developmental axis” that cooperatively regulates early growth in goats.

Notably, several genes involved in neural and ion channel functions, including *CNTNAP2*, *CNTNAP5*, *SLC8A1*, and *KCNN2*, were also significantly associated with ADG. *CNTNAP* family members contribute to synapse formation and neural signaling, and their variants may influence energy metabolism and growth rate via the hypothalamic–pituitary–growth axis [52]. *SLC8A1* encodes the sodium–calcium exchanger, which is essential for calcium homeostasis and muscle contraction [53], while *KCNN2* regulates calcium-activated potassium channels, integrating neuromuscular signaling. Furthermore, enrichment of immune-related genes such as *TLR4* and *CARD11* indicates that inflammation and immune homeostasis contribute to early growth [54]. These findings collectively suggest that early growth is governed by interconnected regulatory networks spanning musculo-skeletal, neural, and immune systems.

Metabolomic profiling revealed pronounced differences in metabolic signatures between fast- and slow-growing individuals. A total of 1,589 metabolites were detected by LC-MS, and multivariate analyses showed a clear separation between the fast-growth (F_F, M_F) and slow-growth (F_S, M_S) groups. Differential metabolites predominantly included amino acids, lipids, organic acids, and bile acids, reflecting systemic differences in energy metabolism, nutrient absorption, and oxidative stress responses. KEGG pathway enrichment demonstrated that metabolites related to energy metabolism and nutrient absorption were strongly enriched in fast-growing individuals, including choline metabolism, fatty acid β-oxidation, and bile acid biosynthesis and transport. Taurine-conjugated bile acids, such as

taurocholic acid and glycocholic acid, not only facilitate lipid emulsification and absorption but also act as signaling molecules regulating glucose, lipid, and energy metabolism [55, 56]. Lipid metabolites, including diglycerides (e.g., DG(8:0/10:0/0:0)) and cholesteryl esters (e.g., CE(18:3(9Z,12Z,15Z))), modulate lipid metabolism and maintain intestinal and hepatic metabolic homeostasis [57, 58]. Lysophosphatidylcholine (LysoPC(14:1(9Z)/0:0)) contributes to improved growth performance and gut health [59], while choline enhances fat digestibility [60], and lysophosphatidic acid (LPA), the hydrolysis product of phosphatidic acid (PA(A-13:0/I-15:0)), alleviates intestinal inflammation [61]. Flavonoids, including taxifolin, luteolin, and eriodictyol, were enriched in fast-growing individuals, contributing to anti-inflammatory, antioxidant, and gut barrier-protective effects [62–65]. These results indicate that fast-growing individuals exhibit enhanced metabolic capacity, lipid utilization, and anti-inflammatory regulation.

Microbiome profiling further revealed correlations between gut microbial composition and host metabolism. Alpha diversity was significantly higher in fast-growing individuals, suggesting a more stable and functionally diverse microbiome. Dominant genera included *Christensenellaceae* R-7 group, unclassified *Ruminococcaceae* UCG-014 and UCG-005, and uncultured Clostridia. Growth performance was closely associated with the abundance of specific microbial taxa, which could be categorized as growth-promoting or growth-inhibiting. In fast-growing individuals, beneficial genera such as *Monoglobus*, *Acetitomaculum*, and *Marvinbryantia* were significantly enriched. *Monoglobus*, a core functional genus, efficiently ferments dietary fiber and has been associated with enhanced nutrient metabolism in ruminants [66, 67]. Its growth-promoting effects are likely mediated by improvements in intestinal barrier function, including positive correlations with mucin gene *Muc2* expression and mucosal repair in both ulcerative colitis models and engineered bacterial interventions [68, 69]. *Acetitomaculum* contributes to cellulose degradation and intramuscular fat deposition [70], whereas *Marvinbryantia* produces short-chain fatty acids (SCFAs) and suppresses pro-inflammatory cytokines, alleviating gut inflammation [71]. In contrast, growth-retarded individuals showed enrichment of potentially harmful taxa, including *Desulfovibrio*, *Candidatus Saccharimonas*, and *Subdoligranulum*, which may compromise intestinal barrier integrity or alter lipid metabolism, thereby restricting growth [72, 73]. These findings provide potential targets for microbiome-based interventions to enhance growth performance.

The metabolome-wide GWAS (mGWAS) GWAS identified key gut metabolites associated with early growth and their host genetic regulators. Notably, *Lysopc*,

taurocholic acid, and glycocholic acid were linked to multiple genes. A variant within the intronic region of *HMGA2* (chr5:47188877) was strongly associated with *Lysopc*. *HMGA2*, a chromatin regulator, plays conserved roles in body weight, skeletal development, and fat deposition, with gene-editing studies confirming that its loss results in reduced body size and organ development defects [74]. Other genes within this locus, including *E2F6* and *KHDRBS3*, contribute to cell proliferation and tissue development. *SCARA5* and *MGAT4D* were associated with taurocholic acid, implicating roles in iron metabolism, PI3K/Akt signaling, and protein glycosylation [75, 76]. Associations between *SLC2A13* and *FHIT* with sebacoyl-L-carnitine suggest their involvement in energy metabolism and oxidative stress adaptation. Collectively, these results indicate that *HMGA2*, *SCARA5*, *SLC2A13*, and other genes modulate tissue development, lipid metabolism, and energy transport, forming genetic couplings with gut metabolites that influence early growth.

Microbiome–host gene interactions revealed by mGWAS demonstrated that *Desulfovibrio*-associated SNPs were distributed across chromosomes 5, 6, 10, 11, 12, 21, and 29, involving *ARHGAP24*, *DPH6*, *FANCL*, and *ATP12A*. *ARHGAP24* regulates inflammation via Rac1/NF- κ B signaling [77], *DPH6* participates in translational modification and metabolic adaptation [78], *FANCL* regulates DNA repair and oxidative stress response, and *ATP12A* modulates gut pH and epithelial barrier function [79]. *Eubacterium hallii* group was associated with *LEKR1* and *NRXN1*, implicating insulin secretion regulation and neural energy signaling, respectively [80]. *Candidatus Saccharimonas* was linked to *IGF2BP2*, which promotes glycolysis and muscle growth via m6A modification [81]. These results suggest that host genetic variants not only shape microbial colonization patterns but also indirectly influence growth via systemic metabolic modulation.

Integrated analysis of SNPs, INDELs, CNVs, and SVs revealed that *Desulfovibrio*, *Eubacterium hallii* group, *Subdoligranulum*, and *Candidatus Saccharimonas* were consistently associated with host genetic variants across multiple genomic layers, involving genes such as *ARHGAP24*, *DPH6* and *IGF2BP2*. These genes participate in inflammation regulation, mitochondrial homeostasis, metabolic signaling, and cell growth, highlighting the decisive role of host genetics in shaping gut microbiota and energy metabolism. Restricted by the population size of the Hechuan white goat, this study employed an extreme phenotype sampling strategy, selecting samples from individuals with significant differences in growth performance to conduct a metabolome-wide association study (mGWAS). Although the sample size utilized for the mGWAS analysis in this study was relatively limited

($n = 39$), the extreme phenotype design enhanced the detectability of genetic effects to a certain extent. Furthermore, previous studies have demonstrated that this strategy can increase the detection power of association signals by maximizing phenotypic variance, making it particularly suitable for the preliminary screening of key candidate loci and metabolic markers associated with complex quantitative traits [82, 83]. However, compared to large-scale population studies, a limited sample size may still reduce the statistical power to detect loci with small-to-medium effects or low-frequency variants, and could compromise the accuracy of effect size estimation [84]. Given the complex polygenic architecture underlying growth traits, the candidate metabolites and associated genetic loci identified in this study provide important clues for elucidating the regulatory mechanisms of early growth in Hechuan white goats. Nevertheless, their biological significance and genetic effects require further validation in larger, independent cohorts.

Conclusion

This study provides the first comprehensive, integrative multi-omics framework elucidating the biological regulation of early growth in Hechuan white goats. By jointly analyzing host genomic variation, gut microbiome composition, and metabolic profiles, we demonstrate that average daily gain (ADG) during the early growth phase is governed by coordinated interactions among musculoskeletal development, energy metabolism, immune homeostasis, and microbial activity in the gastrointestinal tract.

At the host genetic level, the *DLK1–NCAPG2–LCORL* gene axis emerged as a central regulator of musculoskeletal growth, supporting its pivotal role in body size and growth performance. In parallel, genes such as *CNTNAP2* and *NRXN1* suggest an integration between neural development and metabolic signaling, highlighting the contribution of neuro–endocrine pathways to growth regulation. From a microbial perspective, the enrichment of *Monoglobus* and *Acetitomaculum* indicates enhanced energy harvest through short-chain fatty acid (SCFA)–mediated pathways, whereas the increased abundance of *Desulfovibrio* may reflect inflammatory or metabolic constraints that negatively affect growth efficiency.

The integration of genome-wide association studies (GWAS) with microbiome and metabolomic profiling allowed the construction of a host genome–gut microbiome–metabolite interaction network, providing mechanistic insights into how genetic predisposition and microbial metabolic functions jointly shape early growth trajectories. This integrative approach offers a robust theoretical foundation for early-life selection strategies, precision nutritional interventions, and microbiome-informed management aimed at optimizing growth

performance in ruminants. Notably, this study represents the first systematic integration of GWAS, metabolomics, and microbiome data to dissect early growth regulation in goats, underscoring its novelty and translational relevance to livestock production.

Despite these advances, several limitations should be acknowledged. The relatively limited sample size may reduce the power to detect rare genetic variants, and the absence of longitudinal sampling across multiple growth stages constrains causal inference. Moreover, functional characterization of the gut microbiome relied primarily on 16 S rRNA sequencing, without metagenomic or transcriptomic validation. Environmental and dietary variability, together with the restricted genetic diversity of the conservation population studied, may also influence the observed associations. Future studies with expanded cohorts, longitudinal metagenomic tracking, and functional in vitro validation are warranted to fully untangle the underlying molecular mechanisms.

Future research should therefore expand population size, incorporate longitudinal and multi-stage sampling designs, and integrate metagenomic, metatranscriptomic, and metabolomic datasets to capture the dynamic nature of host–microbiome–metabolite interactions. Such efforts will further refine the mechanistic understanding of growth regulation and accelerate the development of data-driven strategies for genetic improvement and precision nutrition in goats and other ruminant species.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42523-026-00597-y>.

Supplementary Material 1

Author contributions

Hui-Long Mou: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology. Ming-Ding Zhang: Supervision, Methodology, Investigation. Zi-Xuan Wang and Zhong-Quan Zhao: Visualization, Validation. Yong-Ju Zhao, Yao-Long Liu and Cheng-Li Liu: Validation, Software, Formal analysis. Tao Ren and Peng-Cheng Ruan: Software, Investigation. Simone Ceccobelli, Yang Zhang and Hao-Yuan Zhang: Visualization, Validation. Jia-Guo Lu and Yong-Fu Huang: Software. Xiao Pan, Li-Ping Zhou and Yan Zeng: Formal analysis, Data curation. Yan-Guo Han: Software, Methodology. Guang-Xin E: Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Funding

This work was supported by the National Natural Science Foundation of China (32272834), Fundamental Research Funds for the Central Universities (SWU-XDJH202301), Chongqing Hechuan District Scientific Research Project (HCKJ-2024-038), and Chongqing Higher Education Teaching Reform Research Project (243382).

Data availability

The whole-genome resequencing data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1394697. This BioProject encompasses two datasets: Host genome resequencing data: Available under SRA accessions

SRR36611752–SRR36611874. 16S rRNA gene amplicon sequencing data: Available under SRA accessions SRR36670798–SRR36670836.

Declarations

Competing interests

The authors declare no competing interests.

Received: 9 January 2026 / Accepted: 29 June 2026

Published online: 11 July 2026

References

1. Lu CD. The role of goats in the world: society, science, and sustainability. *Small Ruminant Res.* 2023;227.
2. Jiang Y, An Z, Li W, Xia S, Ding Q, Zhong J, Wang H, Xu Y, Chen K, Shen Y. Integrated multi-omics reveals the impact of ruminal keystone bacteria and microbial metabolites on average daily gain in Xuzhou cattle. *Microbiol Spectr.* 2025;13.
3. Vickery HMM, Neal RAA, Stergiadis S, Meagher RKK. Gradually weaning goat kids may improve weight gains while reducing weaning stress and increasing creep feed intakes. *Front Vet Sci.* 2023;10.
4. Abdelsattar MMM, Zhao W, Saleem AMM, Kholif AEE, Vargas-Bello-Perez E, Zhang N. Physical, metabolic, and microbial rumen development in goat kids: a review on the challenges and strategies of early weaning. *Animals-Basel.* 2023;13.
5. Moaen-ud-Din M, Muner RD, Khan MS. Genome wide association study identifies novel candidate genes for growth and body conformation traits in goats. *Sci Rep-UK.* 2022;12.
6. 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.
7. Turner LB, Harrison BE, Bunch RJ, Neto LRP, Li Y, Barendse W. A genome-wide association study of tick burden and milk composition in cattle. *ANIM PROD SCI.* 2010;50:235–45.
8. Fan B, Onteru SK, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF. Genome-wide association study identifies loci for body composition and structural soundness traits in pigs. *PLoS ONE.* 2011;6.
9. Wolc A, Arango J, Settar P, Fulton JE, O'Sullivan NP, Preisinger R, Habier D, Fernando R, Garrick DJ, Hill WG, Dekkers JCM. Genome-wide association analysis and genetic architecture of egg weight and egg uniformity in layer chickens. *ANIM GENET.* 2012;43:87–96.
10. Silva EFP, Gaia RC, Mulim HA, Pinto LFB, lung LHS, Brito LF, Pedrosa VB. Genome-Wide association study of conformation traits in Brazilian holstein cattle. *Animals-Basel.* 2024;14.
11. Kaminski S, Tarczynski K, Olenski K, Zybert A, Sieczkowska H, Krzeczio-Nieczy-poruk E, Antosik K, Szwaczkowski T. Genome-wide association study for meat quality traits in a multi-breed pig population. *ANN ANIM SCI.* 2024;24:767–77.
12. Guo J, Jiang R, Mao A, Liu GE, Zhan S, Li L, Zhong T, Wang L, Cao J, Chen Y, et al. Genome-wide association study reveals 14 new SNPs and confirms two structural variants highly associated with the horned / polled phenotype in goats. *BMC Genomics.* 2021;22.
13. Selionova MI, Aibazov AMM, Sermyagin AA, Belous AA, Zinovieva NA. Genome-wide analysis in the search for candidate genes associated with meat productivity traits in meat-and-dairy goats. *Teoriâ i Praktika Pererabotki Mâsa.* 2022;7:288–95.
14. Ceballos FC, Joshi PK, Clark DW, Ramsay M, Wilson JF. Runs of homozygosity: windows into population history and trait architecture. *NAT REV GENET.* 2018;19:220.
15. Huang H, Wang Q, Yang Y, Zhong W, He F, Li J. The mycobiome as integral part of the gut microbiome: crucial role of symbiotic fungi in health and disease. *Gut Microbes.* 2024;16.
16. Pegolo S, Bisutti V, Mota LFM, Cecchinato A, Amalfitano N, Dettori ML, Pazzola M, Vacca GM, Bittante G. Genome-wide landscape of genetic diversity, runs of homozygosity, and runs of heterozygosity in five Alpine and Mediterranean goat breeds. *J Anim Sci Biotechnol.* 2025;16.
17. Teranishi Y, Inoue M, Yamamoto NG, Kihara T, Wiehager B, Ishikawa T, Winblad B, Schedin-Weiss S, Frykman S, Tjernberg LO. Proton myo-inositol cotransporter is a novel γ -secretase associated protein that regulates A β production without affecting Notch cleavage. *FEBS J.* 2015;282:3438–51.
18. Zhernakova DV, Wang D, Liu L, Andreu-Sanchez S, Zhang Y, Ruiz-Moreno AJ, Peng H, Plomp N, Castillo-Izquierdo AD, Gacesa R, et al. Host genetic regulation of human gut microbial structural variation. *Nature.* 2024;625.
19. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature.* 2010;464:59–70.
20. Yang W, Cong Y. Gut microbiota-derived metabolites in the regulation of host immune responses and immune-related inflammatory diseases. *CELL MOL IMMUNOL.* 2021;18:866–77.
21. Simon MREC, Sina C, Ferrario PG, Daniel H. Gut microbiome analysis for personalized nutrition: the state of science. *Mol Nutr Food Res.* 2023;67.
22. Fang S, Chen X, Pan J, Chen Q, Zhou L, Wang C, Xiao T, Gan QF. Dynamic distribution of gut microbiota in meat rabbits at different growth stages and relationship with average daily gain (ADG). *BMC Microbiol.* 2020;20.
23. Zhang A, Yang Y, Li Y, Zheng Y, Wang H, Cui H, Yin W, Lv M, Liang Y, Chen W. Effects of wheat-based fermented liquid feed on growth performance, nutrient digestibility, gut microbiota, intestinal morphology, and barrier function in grower-finisher pigs. *J Anim Sci.* 2024;102.
24. Zhou P, Yan H, Zhang Y, Qi R, Zhang H, Liu J. Growth performance, bile acid profile, fecal microbiome and serum metabolomics of growing-finishing pigs fed diets with bile acids supplementation. *J Anim Sci.* 2023;101.
25. Li L, Li K, Bian Z, Chen Z, Li B, Cui K, Wang F. Association between body weight and distal gut microbes in Hainan black goats at weaning age. *Front Microbiol.* 2022;13.
26. Clarke CJ, Haselden JN. Metabolic profiling as a tool for understanding mechanisms of toxicity. *TOXICOL PATHOL.* 2008;36:140–7.
27. Liu C, Yang Y, Wang M, Jiang W, Du Y, Hao Z, Chen L, Zhu K, Liu B, Niu L, et al. Effects of L-arginine on gut microbiota and muscle metabolism in fattening pigs based on omics analysis. *Front Microbiol.* 2024;15.
28. Li Y, Jia M, Chen J, Liu F, Ren Q, Yan X, Xing B, Pan C, Wang J. A comparative metabolomics study of the potential marker compounds in feces from different hybrid offspring of huainan pigs. *Animals-Basel.* 2024;14.
29. Wang Z, Li S, Ding X, Du X, Zhao A. Study on the differences in fecal metabolites and microbial diversity of jiangshan black-bone chickens with different earlobe colors. *Animals-Basel.* 2024;14.
30. Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *IMETA.* 2023;2.
31. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinf (Oxford England).* 2009;25:1754–60.
32. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. Twelve years of SAMtools and BCFtools. *GIGASCIENCE.* 2021;10:giab008.
33. Ran S, Mu C, Zhu W. Diversity and community pattern of sulfate-reducing bacteria in piglet gut. *J ANIM SCI BIOTECHNO.* 2019;10:40.
34. Wang X, Zheng Z, Cai Y, Chen T, Li C, Fu W, Jiang Y. CNVcaller: highly efficient and widely applicable software for detecting copy number variations in large populations. *GIGASCIENCE.* 2017;6:1–12.
35. Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, Cox AJ, Kruglyak S, Saunders CT. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinf (Oxford England).* 2016;32:1220–2.
36. Layer RM, Chiang C, Quinlan AR, Hall IM. LUMPY: a probabilistic framework for structural variant discovery. *GENOME BIOL.* 2014;15:R84.
37. Ryckman KK, Feenstra B, Shaffer JR, Bream ENA, Geller F, Feingold E, Weeks DE, Gadow E, Cosentino V, Saleme C, et al. Replication of a genome-wide association study of birth weight in preterm neonates. *J Pediatr.* 2012;160:19–24.
38. Jeffares DC, Jolly C, Hoti M, Speed D, Shaw L, Rallis C, Balloux F, Dessimoz C, Bähler J, Sedlazeck FJ. Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *NAT COMMUN.* 2017;8:14061.
39. Wang J, Zhang Z. GAPIT Version 3: Boosting Power and Accuracy for Genomic Association and Prediction. *Genom Proteom Bioinform.* 2021;19:629–40.
40. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *NUCLEIC ACIDS RES.* 2010;38:e164.
41. Liu C, Zhao D, Ma W, Guo Y, Wang A, Wang Q, Lee DJ. Denitrifying sulfide removal process on high-salinity wastewaters in the presence of Halomonas sp. *APPL MICROBIOL BIOT.* 2016;100:1421–6.
42. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, et al. Introducing mothur: open-source, platform-independent, community-supported software for

- describing and comparing microbial communities. *APPL ENVIRON MICROB*. 2009;75:7537–41.
43. Wang M, Jiang P, Yu X, Mi J, Bai Z, Zhang X, Liu Y, Fang X, Yang R, Zhao Z. Analysis of the Bovine DLK1 Gene Polymorphism and Its Relation to Lipid Metabolism in Chinese Simmental. *ANIMALS-BASEL*. 2020;10:923.
44. Zhang L, Kubota M, Nakamura A, Kaji T, Seno S, Uezumi A, Andersen DC, Jensen CH, Si Fukada: Dlk1 regulates quiescence in calcitonin receptor-mutant muscle stem cells. *Stem cells* (Dayton, Ohio). 2021;39:306–317.
45. Fu Y, Hao X, Shang P, Chamba Y, Zhang B, Zhang H. Functional Identification of Porcine DLK1 during Muscle Development. *ANIMALS-BASEL*. 2022;12:1523.
46. You YY, Zhang N, Wang Z, Yin ZH, Bao QY, Lei SX, Xie XJ. DLK1 promoted ischemic angiogenesis through notch1 signaling in endothelial progenitor cells. *ACTA PHARMACOL SIN*. 2024;45:2553–66.
47. Meng F, Zhang S, Song R, Liu Y, Wang J, Liang Y, Wang J, Han J, Song X, Lu Z, et al. NCAPG2 overexpression promotes hepatocellular carcinoma proliferation and metastasis through activating the STAT3 and NF- κ B/miR-188-3p pathways. *EBIOMEDICINE*. 2019;44:237–49.
48. Zhang E, Chen Z, Liu W, Lin L, Wu L, Guan J, Wang J, Kong C, Bi J, Zhang M. NCAPG2 overexpression promotes cancer malignancy and stemness via STAT3/c-MYC signaling. *J TRANSL MED*. 2024;22:12.
49. Rodrigues FM, Majeres LE, Dilger AC, McCann JC, Cassidy CJ, Shike DW, Beever JE. Characterizing differences in the muscle transcriptome between cattle with alternative LCORL-NCAPG haplotypes. *BMC Genomics*. 2025;26:479.
50. Metzger J, Schrimpf R, Philipp U, Distl O. Expression levels of LCORL are associated with body size in horses. *PLoS ONE*. 2013;8:e56497.
51. Lindholm-Perry AK, Kuehn LA, Oliver WT, Sexten AK, Miles JR, Rempel LA, Cushman RA, Freely HC. Adipose and muscle tissue gene expression of two genes (NCAPG and LCORL) located in a chromosomal region associated with cattle feed intake and gain. *PLoS ONE*. 2013;8.
52. Wittekind DA, Scholz M, Kratzsch J, Loeffler M, Horn K, Kirsten H, Witte V, Villringer A, Kluge M. Genome-wide association and transcriptome analysis suggests total serum ghrelin to be linked with GFRAL. *EUR J ENDOCRINOL*. 2021;184:847–56.
53. Huang T, Xu T, Wang Y, Zhou Y, Yu D, Wang Z, He L, Chen Z, Zhang Y, Davidson D, et al. Cannabidiol inhibits human glioma by induction of lethal mitophagy through activating TRPV4. *AUTOPHAGY*. 2021;17:3592–606.
54. Shen X, Lan C, Lin Y, Zhang F, Zhang Y, Chen M, Yan S. Suppression of TLR4 prevents diabetic bone loss by regulating FTO-mediated m6A modification. *INT IMMUNOPHARMACOL*. 2023;122:110510.
55. Wahlström A, Sayin SI, Marschall HU, Bäckhed F. Intestinal Crosstalk between Bile Acids and Microbiota and Its Impact on Host Metabolism. *CELL METAB*. 2016;24:41–50.
56. Sánchez B. Bile acid-microbiota crosstalk in gastrointestinal inflammation and carcinogenesis: a role for bifidobacteria and lactobacilli? *Nat Rev Gastroenterol Hepatol*. 2018;15:205.
57. Mitsuhashi Y, Nagaoka D, Ishioka K, Bigley KE, Okawa M, Otsuji K, Bauer JE. Postprandial lipid-related metabolites are altered in dogs fed dietary diacylglycerol and low glycemic index starch during weight loss. *J Nutr*. 2010;140:1815–23.
58. Li J, Pijut SS, Wang Y, Ji A, Kaur R, Temel RE, van der Westhuyzen DR, Graf GA. Simultaneous Determination of Biliary and Intestinal Cholesterol Secretion Reveals That CETP (Cholesteryl Ester Transfer Protein) Alters Elimination Route in Mice. *Arterioscler Thromb Vasc Biol*. 2019;39:1986–95.
59. Chen C, Jung B, Kim WK. Effects of lysophospholipid on growth performance, carcass yield, intestinal development, and bone quality in broilers. *Poult Sci*. 2019;98:3902–13.
60. Arshad U, Zimpel R, Husnain A, Poindexter MB, Santos JEP. Effect of rumen-protected choline on fat digestibility and lymph metabolome in dairy cows. *J ANIM PHYSIOL N*. 2024;108:950–64.
61. Jenkin KA, He P, Yun CC. Expression of lysophosphatidic acid receptor 5 is necessary for the regulation of intestinal Na⁺/H⁺ exchanger 3 by lysophosphatidic acid in vivo. *Am J Physiol Gastrointest Liver Physiol*. 2018;315:G433–42.
62. Hou J, Hu M, Zhang L, Gao Y, Ma L, Xu Q. Dietary Taxifolin Protects Against Dextran Sulfate Sodium-Induced Colitis via NF- κ B Signaling, Enhancing Intestinal Barrier and Modulating Gut Microbiota. *FRONT IMMUNOL*. 2021;11:631809.
63. Villalva M, Jaime L, de Las Nieves M, Siles-Sanchez S, Santoyo. Bioavailability assessment of yarrow phenolic compounds using an in vitro digestion / Caco-2 cell model: anti-inflammatory activity of basolateral fraction. *Molecules*. 2022;27.
64. Ramos FMM, Ribeiro CB, Cesar TB, Milenkovic D, Cabral L, Noronha MF, Sivieri K. Lemon flavonoids nutraceutical (Eriomin®) attenuates prediabetes intestinal dysbiosis: A double-blind randomized controlled trial. *FOOD SCI NUTR*. 2023;11:7283–95.
65. Ly HT, Pham KD, Le PH, Do THT, Nguyen TTH, Le VM. Pharmacological properties of Ensete glaucum seed extract: Novel insights for antidiabetic effects via modulation of oxidative stress, inflammation, apoptosis and MAPK signaling pathways. *J ETHNOPHARMACOL*. 2024;320:117427.
66. Sallam AM, Reyer H, Wimmers K, Bertolini F, Aboul-Naga A, Braz CU, Rabee AE. Genome-wide landscape of runs of homozygosity and differentiation across Egyptian goat breeds. *BMC Genomics*. 2023;24:573.
67. Wang S, Tang W, Jiang T, Wang R, Zhang R, Ou J, Wang Q, Cheng X, Ren C, Chen J, et al. Effect of Dietary Concentrate-to-Forage Ratios During the Cold Season on Slaughter Performance, Meat Quality, Rumen Fermentation and Gut Microbiota of Tibetan Sheep. *ANIMALS-BASEL*. 2024;14:3305.
68. Liu M, Fu J, Liu Y, Gou W, Yuan W, Shang H. Pectin from comfrey roots alleviate DSS-induced ulcerative colitis in mice through modulating the intestinal barrier. *INT J BIOL MACROMOL*. 2024;282:137016.
69. Yoda M, Takase S, Suzuki K, Murakami A, Namai F, Sato T, Fujii T, Tochio T, Shimamoto T. Development of engineered IL-36 γ -hypersecreting *Lactococcus lactis* to improve the intestinal environment. *WORLD J MICROB BIOT*. 2024;40:363.
70. Jin L, Li K, Li Z, Huang X, Wang L, Wang X, Di S, Cui S, Xu Y. Investigation into Critical Gut Microbes Influencing Intramuscular Fat Deposition in Min Pigs. *ANIMALS-BASEL*. 2024;14:3123.
71. Hu Y, Leu RKL, Christophersen CT, Somashekar R, Conlon MA, Meng XQ, Winter JM, Woodman RJ, McKinnon R, Young GP. Manipulation of the gut microbiota using resistant starch is associated with protection against colitis-associated colorectal cancer in rats. *Carcinogenesis*. 2016;37:366–75.
72. Gebeyew K, Yang C, He Z, Tan Z. Low-protein diets supplemented with methionine and lysine alter the gut microbiota composition and improve the immune status of growing lambs. *APPL MICROBIOL BIOT*. 2021;105:8393–410.
73. Nie Y, Xie XQ, Zhou L, Guan Q, Ren Y, Mao Y, Shi JS, Xu ZH, Geng Y. Desulfovibrio fairfieldensis-Derived Outer Membrane Vesicles Damage Epithelial Barrier and Induce Inflammation and Pyroptosis in Macrophages. *CELLS-BASEL*. 2022;12:89.
74. Chung J, Zhang X, Collins B, Sper RB, Gleason K, Simpson S, Koh S, Sommer J, Flowers WL, Petters RM, Piedrahita JA. High mobility group A2 (HMG A2) deficiency in pigs leads to dwarfism, abnormal fetal resource allocation, and cryptorchidism. *P NATL ACAD SCI USA*. 2018;115:5420–5.
75. Akintayo A, Mayoral J, Asada M, Tang J, Sundaram S, Stanley P. Point mutations that inactivate MGAT4D-L, an inhibitor of MGAT1 and complex N-glycan synthesis. *J Biol Chem*. 2020;295:14053–64.
76. Fang Y, Wu F, Shang G, Yin C. SCARA5 in bone marrow stromal cell-derived exosomes inhibits colorectal cancer progression by inactivating the PI3K/Akt pathway. *Genomics*. 2023;115:110636.
77. Liu H, Wang W, Shen W, Wang L, Zuo Y. ARHGAP24 ameliorates inflammatory response through inactivating Rac1/Akt/NF- κ B pathway in acute pneumonia model of rat. *ANN TRANSL MED*. 2020;8:1289.
78. Uthman S, Bär C, Scheidt V, Liu S, Ten Have S, Giorgini F, Stark MJR, Schaffrath R. The amidation step of diphthamide biosynthesis in yeast requires DPH6, a gene identified through mining the DPH1-DPH5 interaction network. *PLOS GENET*. 2013;9:e1003334.
79. Abdelgied M, Uhl K, Chen OG, Schultz C, Tripp K, Peraino AM, Paithankar S, Chen B, Tamae Kakazu M, Castillo Bahena A, et al. Targeting ATP12A, a Non-gastric Proton Pump α Subunit, for Idiopathic Pulmonary Fibrosis Treatment. *AM J RESP CELL MOL*. 2023;68:638–50.
80. Andersson EA, Harder MN, Pilgaard K, Pisinger C, Stančáková A, Kuusisto J, Grarup N, Færch K, Poulsen P, Witte DR, et al. The birth weight lowering C-allele of rs900400 near LEKR1 and CCNL1 associates with elevated insulin release following an oral glucose challenge. *PLoS ONE*. 2011;6:e27096.
81. Zhou Y, Yan J, Huang H, Liu L, Ren L, Hu J, Jiang X, Zheng Y, Xu L, Zhong F, Li X. The m⁶A reader IGF2BP2 regulates glycolytic metabolism and mediates histone lactylation to enhance hepatic stellate cell activation and liver fibrosis. *Cell Death Dis*. 2024;15.
82. Nicuira SCM, Benavides MV, Okino CH, Ibelli AMG, Minho AP, Esteves SN, Chagas ACDS. Genome-wide association study for haemonchus contortus resistance in morada nova sheep. *Pathogens*. 2022;11:939.
83. Gong C, Zhang Y, Yang C, Ju Z, Liu W, Jiang Q, Wang X, Xiao Y, Wei X, Gao Y, et al. Genome-wide identifications of regulatory variants influencing milk fat percentage in Chinese Holstein cow. *BMC Genomics*. 2025;27:103.

84. Yoon D, Gim J. Genetic architecture of human aging and longevity: Insights from genome-wide association studies. *BMB REP.* 2026;59:2–12.

Publisher's note

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