Microbiome

RESEARCH

Open Access

Deciphering the coordinated roles of the host genome, duodenal mucosal genes, and microbiota in regulating complex traits in chickens

Fangren Lan[1,2,3†], Xiqiong Wang[1,2,3†], Qianqian Zhou[1,2,3†], Xiaochang Li[1,2,3], Jiaming Jin[1,2,3], Wenxin Zhang[1,2,3], Chaoliang Wen[1,2,3], Guiqin Wu[4], Guangqi Li[4], Yiyuan Yan[4], Ning Yang[1,2,3*] and Congjiao Sun[1,2,3*]

Abstract

Background The complex interactions between host genetics and the gut microbiome are well documented. However, the specific impacts of gene expression patterns and microbial composition on each other remain to be further explored.

Results Here, we investigated this complex interplay in a sizable population of 705 hens, employing integrative analyses to examine the relationships among the host genome, mucosal gene expression, and gut microbiota. Specific microbial taxa, such as the cecal family Christensenellaceae, which showed a heritability of 0.365, were strongly correlated with host genomic variants. We proposed a novel concept of regulatability (r_b^2), which was derived from h^2, to quantify the cumulative effects of gene expression on the given phenotypes. The duodenal mucosal transcriptome emerged as a potent influencer of duodenal microbial taxa, with much higher r_b^2 values (0.17 ± 0.01, mean ± SE) than h^2 values (0.02 ± 0.00). A comparative analysis of chickens and humans revealed similar average microbiability values of genes (0.18 vs. 0.20) and significant differences in average r_b^2 values of microbes (0.17 vs. 0.04). Besides, cis (h_{cis}^2) and trans heritability (h_{trans}^2) were estimated to assess the effects of genetic variations inside and outside the cis window of the gene on its expression. Higher h_{trans}^2 values than h_{cis}^2 values and a greater prevalence of trans-regulated genes than cis-regulated genes underscored the significant role of loci outside the cis window in shaping gene expression levels. Furthermore, our exploration of the regulatory effects of duodenal mucosal genes and the microbiota on 18 complex traits enhanced our understanding of the regulatory mechanisms, in which the *CHST14* gene and its regulatory relationships with *Lactobacillus salivarius* jointly facilitated the deposition of abdominal fat by modulating the concentration of bile salt hydrolase, and further triglycerides, total cholesterol, and free fatty acids absorption and metabolism.

Conclusions Our findings highlighted a novel concept of r_b^2 to quantify the phenotypic variance attributed to gene expression and emphasize the superior role of intestinal mucosal gene expressions over host genomic variations

†Fangren Lan, Xiqiong Wang and Qianqian Zhou contributed equally to the work.

*Correspondence:
Ning Yang
nyang@cau.edu.cn
Congjiao Sun
cjsun@cau.edu.cn
Full list of author information is available at the end of the article

in elucidating host–microbe interactions for complex traits. This understanding could assist in devising strategies to modulate host–microbe interactions, ultimately improving economic traits in chickens.

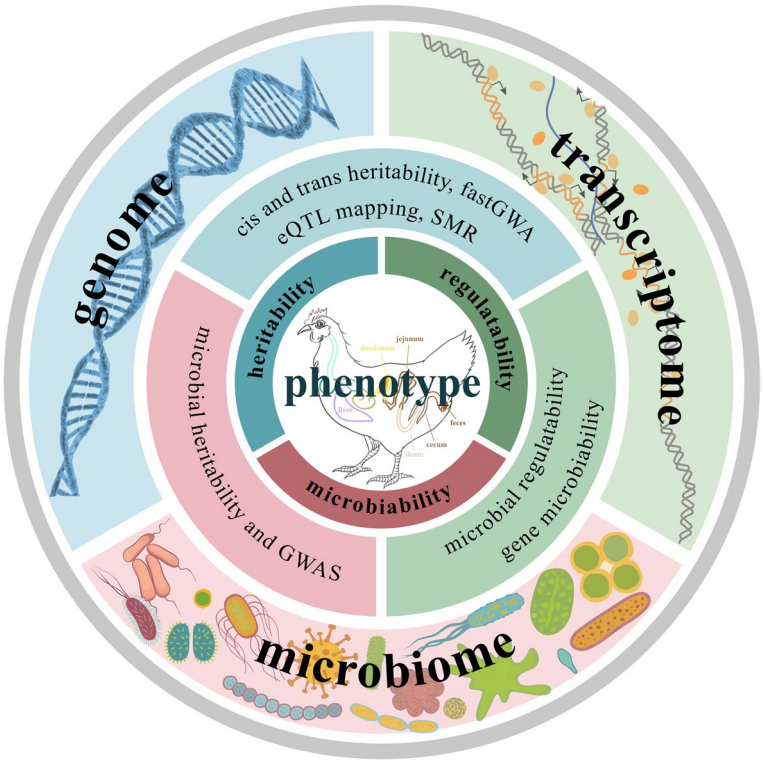
Keywords Chickens, Host genetics, Gut microbiota, Regulatability, Integrative analysis

Graphical Abstract

The host genome, hepatic and duodenal transcriptomes, and gut microbiome were integrated to elucidate the regulatory mechanisms underlying the complex phenotypes of chickens. Heritability estimation and GWASs from the genome, regulatability estimation from the transcriptome, and microbiability estimation from the microbiome were collectively calculated for the host phenotypes.

By integrating genome and transcriptome data, the cis- and trans-heritability of genes were estimated, and methods such as eQTL mapping and SMR were used to elucidate the genetic regulatory mechanisms of phenotypes. By integrating genome and microbiome data, we conducted microbial heritability estimation and GWASs to explore the extent to which the host genome shapes the microbiota. By integrating transcriptome and microbiome data, we estimated the microbiability of genes and the regulatability of the microbiota, exploring the degree of interaction between host gene expression and the gut microbiota.

By combining all the analytical methods mentioned above, a more advanced and comprehensive understanding of the extent and manner of host and microbiota interactions that regulate host phenotypes can be achieved.



Background

The gut microbiota is influenced by numerous factors, including environmental exposure [1, 2], dietary fluctuations [3, 4], age-related dynamics [5, 6], sex-specific influences [7, 8], and underlying genetic distinctions [9, 10]. Host genetics interact with the gut microbiota in a complex and dynamic manner, with host genes impacting the presence/absence and quantity of microbes [11]

and shaping the environment for microbial survival [12]; in turn, microbes can directly or indirectly participate in host energy metabolism and nutrient absorption [13, 14], reciprocally affecting cellular functions in intricate ways through fermentation and metabolite production. However, the comprehensive characterization of interactions between the host and its intestinal microbiome remains a challenge, requiring a range of advanced and nuanced approaches.

Researchers have identified single-nucleotide polymorphisms (SNPs) in the host that are associated with the abundance of specific gut microbes. Variations in the genes *Irak3*, *Lyz1*, and *Lyz2*, which belong to the immunomodulatory *Irak3*–Toll-like receptor 2 (*TLR2*) signaling pathway, were found to shape the relative abundances of *Lactococcus* and Coriobacteriaceae by altering the ability of the host to immunologically respond to peptidoglycan [15]. The genetic variants located at or near the lactase (*LCT*) gene consistently exhibited associations with the Actinobacteria clade, particularly the *Bifidobacterium* genus and its species, which have been independently observed in various populations, including British [16], Dutch [17], Canadian [18], and Finnish [19] populations, as well as in the meta-analysis conducted by the MiBioGen consortium [20]. This locus is the most rigorously validated finding in microbial quantitative trait locus (mbQTL) studies to date. The *LCT* gene encodes the lactase enzyme, which breaks down lactose into glucose and galactose. In addition, a range of studies have been performed in humans to interpret the associations between the microbiome and the *ABO* locus, but the specific and nuanced connections have varied across different studies. The *ABO* locus was associated with the abundance of *Faecalibacterium* and *Bacteroides* in a German cohort [21], *Faecalicatena lactaris* and *Collinsella* in a Finnish cohort [22], and *Bifidobacterium* and *Collinsella* in a Dutch population [17]. In pigs, Yang et al. [23] revealed compelling evidence supporting the effect of the host genotype on the abundance of specific intestinal bacteria and systematically elucidated the underlying mechanism, a 2.3-kb deletion in the *ABO* gene that reduces the concentration of N-acetylgalactosamine in the gut, consequently diminishing the abundance of Erysipelotrichaceae, which is known for importing and catabolizing this compound.

From the perspective of quantitative genetics, the contribution of the microbiome to the phenotypic variance of a given trait can be quantified via estimates of microbiability (m^2) [22], which is analogous to the classical concept of heritability (h^2) [24]. This insight emerged when Difford et al. [25] reported convincing estimates of the h^2 (0.21) and m^2 (0.13) associated with CH_4 emissions in a substantial cohort of 750 cows. Similar investigations have been rigorously conducted in farm animal populations, illuminating the scientific landscape through analogous studies on feed efficiency in pigs [26–28] and abdominal fat deposition [29], and feed efficiency [30] in chickens, collectively contributing to a deeper understanding of the complex dynamics and shared principles governing these diverse animal models.

Gene regulation plays a pivotal role in numerous complex traits, exemplified by the fact that variants associated with these traits are notably enriched in regulatory regions compared with protein-coding regions [31–33]. Numerous studies on expression quantitative trait loci (eQTLs), lncRNA QTLs (lncQTLs), exon QTLs (exQTLs), splicing variation QTLs (sQTLs), and 3′UTR alternative polyadenylation QTLs (3a′QTLs) have significantly increased our understanding of the intricate landscape of genetic variants and their causal roles in regulating gene expression, thereby influencing phenotypic diversity in humans [34], cattle [35], pigs [36], and chickens [37]. These efforts are crucial for elucidating the functions of numerous genetic variants, particularly those with largely unexplored regulatory roles, thereby revealing the intricate coordination of molecular processes governing human traits. Previous studies have focused predominantly on cis effects. Although a substantial impact is consistently observed in eQTLs near the transcription start sites (TSSs) of genes [38], eQTLs can exert their pronounced influence over considerable distances, commonly referred to as trans effects. Trans-eQTLs contribute to the regulation of multiple genes, potentially playing a broader role in coordinating the complex regulatory networks that govern gene expression across the entire genome through long-range regulatory mechanisms [39, 40].

In this study, we comprehensively characterized the complex interactions among host genomic variants, mucosal gene expression, and intestinal microbial composition affecting various traits, such as feed efficiency, egg production, carcass traits, and hepatic and serum biochemical indicators. We analyzed over 20,000 genes in the liver and duodenum, coupled with hundreds of duodenal microbial taxa, and observed their reciprocal influence on each other. Notably, we developed and adopted innovative approaches, including h^2, m^2, and regulatability (r_b^2, the relative proportion of the total variance attributable to the gene expression profile, calculated in analogy to h^2 and m^2), designed to address common challenges encountered in multiomics integrations, such as high dimensionality, sparsity, and multicollinearity. These indices provide a holistic perspective on the combined impact of the genomic variants, genes, and microbiota on the phenotypes. And these methodologies facilitated the identification of robust and biologically meaningful associations between gut microbes and host genes.

Methods

Experimental design

All animal experiments were conducted according to the ethical policies and procedures approved by the Institutional Animal Care and Use Committee of China Agricultural University, China (Issue No. 32303202–1-1). A cohort of 705 female chickens from a pure Rhode Island Red line was housed at Beijing Huadu Yukou Poultry Breeding Co. Ltd. (Pinggu, China). The birds were kept in individual cages and were treated identically, receiving corn–soybean-based diets and water *ad libitum*. No antibiotics or other antimicrobial treatments were administered prior to or during the study.

Feed efficiency traits, including the feed conversion ratio (FCR) and residual feed intake (RFI), were measured from 68 to 72 weeks of age as described previously [41]. The total egg number (EN) from the first egg until the age of 90 weeks was recorded. At age 90 weeks, body weight (BW) and egg weight (EW) were measured using electronic scales with accuracies of 5 g and 0.1 g, respectively. Whole blood samples were collected from the wing vein and stored at – 20 °C. The serum was separated by centrifugation at 3000 × *g* for 15 min and stored at – 20 °C until the measurement of serum biochemical indicators, including total cholesterol (STC), triglycerides (STG), total bile acids (STBA), high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, and very low-density lipoprotein (VLDL) cholesterol. Fecal samples were obtained by introducing a sterile moistened swab (BKMAM®, Hunan, China) into the rectum and rotating it. The birds were subsequently euthanized by cervical dislocation and dissected. After the abdomen was opened, the liver and abdominal fat (surrounding the gizzard, cloaca, and adjacent abdominal muscles) were carefully dissected. The liver weight (LW) and abdominal fat weight (AFW) were weighed promptly using an electronic scale accurate to 0.1 g, after which the liver percentage (LP) and abdominal fat percentage (AFP) were calculated. Liver tissues were collected for host whole-genome resequencing, transcriptomic sequencing, and measurement of hepatic biochemical indicators, including total cholesterol (HTC) and triglycerides (HTG). The duodenal, jejunal, ileal, and cecal chyme, as well as the mucosa from the same segments, were collected immediately after the birds were sacrificed. The entire slaughter and sampling process took 3 days to complete. All the samples were snap-frozen in liquid nitrogen, transferred to the laboratory, and immediately stored at – 80 °C for long-term storage until subsequent sequencing.

Whole-genome resequencing and data processing

Genomic DNA was extracted from liver samples using a Tiangen DNA Extraction Kit (Tiangen Biotech, Beijing, China, DP304-2) according to the manufacturer's instructions. DNA concentrations were measured using a Nanodrop-2000 instrument (Thermo Fisher Scientific, Waltham, MA, USA), and the DNA quality was assessed via 1% agarose gel electrophoresis. Afterward, 686 DNA samples were selected for subsequent whole-genome resequencing. The DNA was amplified by PCR with 500 bp inserts to construct libraries. The TruSeq Nano DNA LT Library Preparation Kit (Illumina, CA, USA) was used for DNA library construction. Whole-genome resequencing was conducted using the Illumina HiSeq 2500 Sequencer (Illumina, Inc., San Diego, CA, USA) to generate 150 bp paired-end reads. To ensure data quality reads containing adaptor contamination, low-quality reads, and reads with more than 5% N bases were removed using FastQC software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The clean reads were then aligned to the chicken reference genome (GRCg6a, https://ftp.ensembl.org/pub/release-106/fasta/gallus_gallus/) using the Burrows–Wheeler aligner (BWA, version 0.7.15) [42] with default parameters.

We subsequently utilized SAMtools (version 1.3.1) [43] to sort the reads and remove low-quality reads with the parameter "-q 4". Duplicate reads in the PCR results were removed using Picard tools (http://broadinstitute.github.io/picard/). For SNP calling, we employed the HaplotypeCaller protocol in the Genome Analysis Toolkit (GATK, version 4.2.0.0) [44]. To obtain high-quality SNPs, the GATK VariantFiltration protocol was used to filter the SNPs with the following parameters: QD < 2.0, ReadPosRankSum < – 8.0, FS > 60.0, QUAL < 30.0, DP < 4.0, MQ < 40.0, and MappingQualityRankSum < – 12.5. Finally, PLINK (version 1.90) [45] was used to filter the SNP data according to the following parameters: sample call rate > 90%, SNP call rate > 90%, and minor allele frequency (MAF) > 1%. The remaining SNPs and individual birds were imputed using BEAGLE (version 5.1) [46], and the PLINK analysis was re-executed using the same criteria as described above. Following these steps, a total of 5,904,820 SNPs distributed across 32 chromosomes and 686 birds were retained for subsequent analysis.

Liver tissue and duodenal mucosa transcriptome sequencing and analysis

To complete the RNA extraction, an Eastep® Super Total RNA Extraction Kit (Promega, Shanghai, China, LS1040) was used for liver tissues, and a TRIzol kit was used for duodenal mucosa tissues according to the manufacturer's instructions. The RNA concentration and purity were determined with a NanoDrop-2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and the integrity was assessed with the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent

Technologies, CA, USA). Following purification, the mRNA was fragmented into short fragments using divalent cations under elevated temperature (Magnesium RNA Fragmentation Module [NEB, cat. e6150, USA] at 94 °C for 5–7 min). The cleaved RNA fragments were subsequently reverse transcribed to create cDNA using SuperScript™ II Reverse Transcriptase (Invitrogen, cat. 1896649, USA). The cDNAs were subsequently used to synthesize U-labeled double-stranded DNAs using *E. coli* DNA polymerase I (NEB, cat. m0209, USA), RNase H (NEB, cat. m0297, USA) and dUTP Solution (Thermo Fisher, cat. R0133, USA). An A-base was then added to the blunt ends of each strand, preparing them for ligation to the indexed adapters. Each adapter contained a T-base overhang for ligation with the A-tailed fragmented DNA. Dual-index adapters were ligated to the fragments, and size selection was performed with AMPureXP beads. After treatment of the U-labeled double-stranded DNA with heat-labile UDG enzyme (NEB, cat. m0280, USA), the ligated products were amplified via PCR under the following conditions: initial denaturation at 95 °C for 3 min; 8 cycles of denaturation at 98 °C for 15 s, annealing at 60 °C for 15 s, and extension at 72 °C for 30 s; and a final extension at 72 °C for 5 min. The average insert size for the final cDNA library was 300 ± 50 bp. Libraries for transcriptome sequencing were constructed according to the standard Illumina RNA-seq instructions and sequenced on an Illumina NovaSeq platform, generating 150 bp paired-end reads. Quality control was performed using Fastp (version 0.20.1) [47] to eliminate reads containing adaptor contamination, low-quality bases, and undetermined bases. The quality-controlled sequencing data were aligned to the chicken reference genome (GRCg6a) using HISAT2 (version 2.0.5) [48] with default parameters. FeatureCounts (version 1.6.3) [49] and StringTie (version 1.3.1c) [50] were used to tally the reads and calculate the transcripts per million (TPM) values. The read counts were subsequently normalized between samples via the trimmed mean of *M* values (TMM) method [51]. Distinct patterns of gene expression in the liver and duodenum were discerned and visualized with a principal component analysis (PCA) plot. Differentially expressed genes (DEGs) were identified using the edgeR package in R (version 4.0.2), and notable enrichment of biological processes was identified.

16S rRNA gene sequencing

The gut microbial DNA was isolated from the gut digesta and fecal samples from the duodenum, jejunum, ileum, cecum, and feces separately using a Tiangen DNA Extraction Kit (Tiangen Biotech, Beijing, China) and a QIAamp DNA Stool Mini Kit (QIAGEN, Hilden, Germany, D4015-01) as appropriate, in accordance with the manufacturer's standard protocol. The hypervariable V4 region of the 16S gene was amplified from the extracted DNA using an Ion Plus Fragment Library Kit 48 rxns (Thermo Scientific), following a previously described method [52]. Sequencing was performed on an Ion S5™ XL platform according to the manufacturer's instructions, generating 400 bp single-end reads.

Further bioinformatic analyses were carried out using Quantitative Insights Into Microbial Ecology (QIIME2, version 2019.10) [53]. After the barcode and primer sequences were trimmed, the original high-throughput sequencing data were subjected to preliminary quality screening using the QIIME2 plugin DADA2 (version 1.16) [54], with the sequences trimmed to a final length of 252 bp. The remaining high-quality sequences were then clustered and classified by amplicon sequence variants (ASVs) with 100% identity [55]. ASVs present in less than 1% (seven individuals) of the samples and with an average relative abundance of less than 10^{-6} were excluded for subsequent analyses. Taxonomic assignments for each ASV were made via similarity searching against the SILVA 16S rRNA gene sequence reference database (Release 132) [56]. The Shannon index was used as an α-diversity metric and calculated with the vegan package (version 2.6–4) [57] in R (version 4.0.2). The Bray–Curtis dissimilarity was used as a β-diversity measure and subjected to principal coordinate analysis (PCoA) with vegdist in the vegan package (version 2.6–4).

Heritability, microbiability, and regulatability estimation

To assess the impact of host genetics on various phenotypes and the gut microbiota, we employed a restricted maximum likelihood analysis within the framework of the following model using GCTA (version 1.93.2) [58]:

$$Y = Kc + g + e$$

Here, Y represents a vector of corrected phenotypes encompassing recorded traits, gene expression, and the relative abundance or presence/absence of microbial taxa. K denotes a design matrix that includes batch effects and the top 10 host genetic principal components; c is a vector of effects for the covariates. The vector g represents the total effects of all SNPs following a normal distribution $\sim N\ (0,\ G\sigma^2_g)$, where G and $G\sigma^2_g$ represent the genetic relatedness matrix (GRM) and genetic variance, respectively. Finally, e accounts for residual errors. The 5,904,820 meticulously filtered SNPs were used to construct the GRM with an estimation model expressed as the following equation:

$$g_{ij} = \frac{1}{N}\sum_{V=1}^{N}\frac{(x_{iv} - 2\overline{p_v})(x_{jv} - 2\overline{p_v})}{2\overline{p_v}(1 - \overline{p_v})}$$

In this equation, g_{ij} denotes the genetic relationship between individuals i and j; x_{iv} and x_{jv} represent the number of reference alleles in hens i and j, respectively; p_v denotes the reference allele frequency; and N is the total number of SNPs.

Gene expression is notably shaped by both local (cis) and non-local (trans) eQTLs. The heritability of genes encompasses two fundamental components: cis-heritability and trans-heritability. cis-heritability describes the genetic influence of a QTL on the expression of genes within a close genetic distance, typically near a gene or its neighboring region; this index accounts for the impact of genomic variations and regulatory elements in the nearby genetic sequence and their contribution to gene expression levels. The cis-GRM was established to capture the variance explained by all measured SNPs in the cis window surrounding the gene. The cis window extended from 1 Mbps upstream to 1 Mbps downstream of the TSS. The variance explained by the cis-GRM is referred to as cis-h^2 (h^2_{cis}). In contrast, trans-h^2 (h^2_{trans}) characterizes the heritability of genes originating from more distant genomic locations or regulatory elements. This index represents the overall genetic relatedness between individuals and captures the genetic effects that contribute to gene expression variability but are not directly linked to gene expression. A second GRM, termed the residual GRM, was created for the autosomes that included closely related individuals (genetic correlation ≥ 0.05). This matrix accounts for the genetic correlation across the autosomes, and the variance explained by this matrix is referred to as h^2_{trans}. This index eliminates distant relatedness from the matrix by nullifying off-diagonal elements less than 0.05. The cis-GRM focuses on a limited number of SNPs that have considerable power in detecting cis-genetic effects [59], and the second GRM absorbs all genetic variance not explained by the SNPs within the cis window or those in high linkage disequilibrium (LD) with the cis window SNPs. The sum of these two effects approximates the estimated total heritability (h^2_{total}).

Similar to heritability, the phenotypic variances attributed to the gut microbial community and gene expression profile were defined as microbiability (m^2) and regulatability (r^2_b), respectively, and estimated with GCTA software using the microbial relationship matrix (MRM) and regulative relationship matrix (RRM). All filtered ASVs in the duodenum were normalized by zero-centering and scaling to unit variance to construct the MRM as previously described [26, 29, 60] with an R script based on the following equation:

$$m_{ij} = \frac{1}{N}\sum_{a=1}^{N} \frac{(x_{ia} - \overline{x_a})(x_{ja} - \overline{x_a})}{\sigma_a^2}$$

where m_{ij} represents the estimated duodenal microbial relationship between birds i and j; x_{ia} and x_{ja} denote the relative abundances of the ASV a in birds i and j, respectively; x_a represents the average relative abundance of the ASV a in the population; σ_a^2 represents the variance in the abundance of the ASV a; and N represents the total number of ASVs used for the relatedness calculation.

Similarly, genes were filtered according to the criteria of ≥ 0.1 TPM and ≥ 6 unnormalized reads in $\geq 20\%$ of the samples. The TPM values were used to create the RRM as follows:

$$r_{ij} = \frac{1}{N_d}\sum_{t=1}^{N_d} \frac{(x_{it} - \overline{x_t})(x_{jt} - \overline{x_t})}{\sigma_t^2}$$

where r_{ij} represents the estimated relationship between birds i and j; x_{it} and x_{jt} are the expression levels of gene t in birds i and j, respectively; x_t signifies the average expression of gene t in the population; σ_t^2 is the variance of the expression of gene t; and N_d is the total number of genes used for the relatedness calculation.

The h^2 estimates of various recorded phenotypes, intestinal microbial taxa, and duodenal mucosal genes were estimated, along with the m^2 estimates of the phenotypes and genes and the r^2_b estimates of the phenotypes and microbial taxa. Notably, prior to that, the abundance of taxa in a specific segment that was present in at least 60% of individuals was encoded using normalized abundance via rank-based inverse normal transformations in R (version 4.0.2). Relative abundance values of 0 were considered not detected and were converted to NA. Otherwise, taxa present in $< 60\%$ but $\geq 30\%$ of the samples were dichotomized into presence/absence patterns, and the phenotype was encoded as a binary vector to prevent zero inflation, which led to a bimodal distribution [61]. Microorganisms that were detected in $< 30\%$ of the samples were excluded from this analysis as previously reported [29].

Genome-wide association study (GWAS)

We employed a univariate linear mixed model (LMM) in GEMMA (version 0.98.4) [62] to conduct GWASs aiming to identify significant host genetic markers influencing recorded phenotypes and the abundance or presence/absence of heritable genera. The statistical model applied was as follows:

$$y = Wa + x\beta + u + \varepsilon$$

where y represents the corrected phenotypic values (recorded traits, abundance or presence/absence of heritable microbial taxa); W denotes a matrix of covariates (including the top 5 principal components and batch effects) controlling for population structure; α is a vector

of corresponding effects that compose the intercept; x represents the marker genotypes; β is the effect of the markers; u is a vector of random polygenic effects with a covariance structure; and ε is a vector of residual errors. The significance of the associations between SNPs and phenotypes was assessed through the likelihood ratio test. The genome-wide significance threshold was determined using a modified Bonferroni correction with the R (version 4.0.2) package simpleM (https://github.com/LTibbs/SimpleM) as previously described [63]. Using this approach, 150,802 valid inspections were obtained, leading to the determination of genome-wide significance and suggestive significance thresholds of 3.32×10^{-7} (0.05/150802) and 6.63×10^{-6} (1/150802), respectively.

In addition to the traits and heritable microbial taxa, GWASs of filtered genes and all microbial taxa with detection rates exceeding 30% were conducted using GCTA software with the support of fastGWA [64]. Similarly, LMM was employed throughout the analysis. For fastGWA, a fully dense GRM was generated, on which a sparse GRM was built at a cutoff value of 0.05. FastGWA was then run using this sparse GRM with expression level and the abundance or presence/absence of the microbial taxa as the dependent variable and SNP genotype value as the independent variable. The significant and suggestively significant p value thresholds were 3.32×10^{-7} and 6.63×10^{-6}, respectively, as described above.

eQTL mapping and Mendelian randomization analysis

The eQTLs are powerful genetic indices used in genomics to investigate the genetic basis of gene expression variations. By studying the genetic regulation of gene expression, we can gain insights into the molecular mechanisms underlying complex traits and diseases. To identify genetic variants or loci that are associated with the expression levels of specific genes, we incorporated several covariates, including the top 5 genetic principal components, batch effects, and the top 3 probabilistic estimation of expression residuals (PEER) factors. The choice of 60 PEER factors for use in calculations was based on the sample size corresponding to previous research: 15 for $n<150$, 30 for $150\leq n<250$, 45 for $250\leq n<350$, and 60 for $n\geq350$ [65].

We subsequently performed a comprehensive analysis encompassing cis-eQTL and trans-eQTL mappings. For cis-eQTLs, rigorous permutations were executed to generate gene expression-level summary statistics, yielding empirical p values. To account for multiple tests, we applied false discovery rate (FDR) correction specifically for cis-eQTL analysis, setting a stringent p value threshold of 8.05×10^{-6}. In trans-QTL mapping, a set of 5,315,471 common genetic variants passed stringent quality control criteria (MAF > 5% and location ± 5 Mbps outside the TSS), and permutations were performed to compute nominal associations between all genes and genotypes. Considering the similarity of the SNPs used in trans-QTL mapping to those used in the GWASs, we established a p value threshold of 3.32×10^{-7}.

Subsequently, top-eQTL-based summary data-based Mendelian randomization (SMR) [66] analysis was performed to prioritize genes underlying the GWAS findings. This methodology comprises two crucial steps: (i) identifying variants independently associated with the exposure factor and (ii) calculating causal estimates. Prior to this analysis, we meticulously prepared a BESD file, updating the SNP and gene coordinates and the frequency of the effect allele. For each GWAS summary statistic, we picked SNPs that were significantly and suggestively associated with the traits as SMR input files to discern connections with significant cis-eQTLs and trans-eQTLs. Ultimately, we retained only those variants displaying associations at an FDR-corrected p value < 0.05. Moreover, the bidirectional generalized summary data-based Mendelian randomization (GSMR) method was employed to investigate the putative causal associations between specific phenotypes and microbial taxa.

Analysis of the associations among the host genome, gene expression profile, and gut microbial community

To investigate the effects of host genetics on hepatic and duodenal mucosal gene expression and the gut microbiota, we evaluated the correlation between host genetic kinship (based on the GRM) and duodenal microbial similarity (based on MRM) for each gut section through Spearman's correlation-based Mantel tests with 10,000 permutations.

To elucidate the relationships and distinctions between duodenal mucosal gene expression and the duodenal microbiota, Procrustes analysis was employed. Procrustes analysis is a powerful tool for investigating and quantifying the concordances and divergences between shapes or configurations and determining the optimal alignment by correlating ordination spaces rather than individual features [67]. This method involved computing the Euclidean distance for the gene expression matrix and the Bray–Curtis distance matrices for the ASV-level microbial matrix using the R function "vegdist" in the vegan (version 2.6–4) package [57]. The "protest" function in the vegan (version 2.6–4) package was subsequently used to derive the symmetric Procrustes correlation coefficients (r), the sum of squared distance (M^2), the p value, and the Procrustes residuals via 1000 permutations. The Procrustes correlations were visually depicted using ggplot2 (version 3.5.1.9000).

In addition, we downloaded previously published human colonic mucosal gene expression data (RNA-seq, NCBI Sequence Read Archive [SRA] under BioProject accession number PRJNA816986.) and gut microbiome data (16S rRNA, NCBI SRA under BioProject accession number PRJNA284355) from 88 individuals in an irritable bowel disease (IBD) cohort reported by Blekhman [68, 69] to compare host genetics and intestinal microbial interaction patterns between chickens and humans. Specifically, the m^2 values of human colonic genes and the r_b^2 values of the microbiota were separately estimated. This cross-species exploration allowed for a unique perspective on the role of host genetics in shaping the gut microbiome and offered valuable insights and a detailed examination of the similarities and differences between the species, by which we can uncover universal principles as well as species-specific factors that contribute to the complex interplay between host genetics and the intestinal microbiota.

Reverse transcription-quantitative real-time polymerase chain reaction (RT–qPCR) of CHST14

To validate the regulatory effects of the SNPs 3:10,534,146 and 3:11,263,397 on the expression of *CHST14*, we conducted RT–qPCR on 10 individuals randomly selected from 3 genotypes. Total RNA was extracted from the duodenal mucosa using an RNAsimple Total RNA Kit (Tiangen Biotech, Beijing, China, DP419). After removing the genomic DNA, 1 μg of total RNA was reverse transcribed using a PrimeScript™ RT reagent Kit with gDNA Eraser (Perfect Real Time) (Takara) to obtain cDNA. Then, 2 μL of diluted cDNA was used to perform a real-time PCR assay using TB Green® *Premix Ex Taq*™ (Tli RNaseH Plus) (Takara). PCR amplification was monitored with an Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher Scientific). The cycling conditions were as follows: 95 °C for 30 s, 40 amplification cycles (95 °C for 5 s, 60 °C for 30 s), and 1 melting curve cycle (95 °C for 15 s, 60 °C for 1 min, 95 °C for 15 s), with fluorescence detection at the end of each cycle. For quantification, the data were analyzed via the $2^{-\triangle\triangle CT}$ method and normalized to the housekeeping gene *β-actin*. The differences among the three genotypes were determined using the Wilcoxon test. The sequences of the primers used in this study were as follows:

CHST14: F: 5′-GGAAATAGTCAGGCGGTAT-3′;
R: 5′-GCACAGGTTGTAGATGGG-3′;
β-actin: F: 5′-TTGTTGACAATGGCTCCGGT-3′;
R: 5′-TCTGGGCTTCATCACCAACG-3′

Duodenal chyme and liver assays

To explore the regulatory relationships among *CHST14*, *Lactobacillus salivarius*, and AFW, 10 individuals from each extreme of *CHST14* expression levels were selected for the measurement of biochemical indicators. The levels of bile salt hydrolase (BSH), triglycerides (TG), total cholesterol (TC), free fatty acids (FFA), and total bile acids (TBA) in the duodenal chyme were quantified using commercial kits (Enzyme-linked Biotechnology Co., Ltd., Shanghai, China) following the manufacturer's protocols. Liver tissues from the same 20 individuals were weighed (ranging from 0.01 to 0.05 g) on a precision electronic scale. The total protein content was quantified using a protein assay kit (Nanjing Jiancheng Institute of Biological Engineering, China, A045-2) according to the manufacturer's guidelines. HTG and HTC levels were subsequently measured using specific assay kits (Nanjing Jiancheng Institute of Biological Engineering, China A110-2–1) according to the manufacturer's instructions.

Results

Overall characterization of multiomics data

A total of 3519 microbial DNA samples were used for 16S rRNA gene sequencing, including 705 duodenal, 705 jejunal, 704 ileal, 705 cecal, and 700 fecal samples, generating 174.2 million quality-filtered sequences with an average of 49,497 reads (Supplementary Table S1). Then, 6087 (duodenum), 5987 (jejunum), 3751 (ileum), 3215 (cecum), and 7428 (feces) ASVs were clustered with 100% sequence identity in each gut segment and subsequently classified into 3930 species, 2329 genera, 1003 families, 467 orders, 161 classes, and 52 phyla (Supplementary Fig. S1A). The digestive tract comprises several distinct habitats that influence the heterogeneous spatial arrangement of the resident microbiota, leading to spatial changes in the diversity and composition of intestinal microbes. The alpha diversities differed significantly among the five sampling sites ($p < 0.05$, ANOVA, Supplementary Fig. S1B), with the cecum having the highest diversity and the ileum having the lowest diversity (4.66 ± 0.02 vs. 2.39 ± 0.04, Supplementary Fig. S1B). PCoA was performed, revealing distinct differences in the gut microbial community among the hindgut (duodenum and jejunum), ileum, cecum, and feces, as evidenced by the separate clustering of these groups (Supplementary Fig. S1C). At the phylum level, the three segments of the small intestine (SI) harbored similar dominant microorganism communities, with Firmicutes being the predominant phylum, followed by Proteobacteria, Bacteroidetes, and Actinobacteria (Supplementary Fig. S1D). However, evident differences were observed in the cecum, with Bacteroidetes (53.91%) and Firmicutes (36.83%) being the most abundant phyla. At the genus level, *Lactobacillus*

represented the majority of the genera in the duodenum (46.06%), jejunum (54.25%), and feces (22.02%), whereas *Romboutsia* was the most abundant genus (30.32%) in the ileum, and *Bacteroides* constituted a notable fraction (30.04%) in the cecum (Supplementary Fig. S1E).

Hepatic DNA from 686 individuals was extracted for whole-genome resequencing. Up to 7.74 Tb of clean reads were generated, and each individual reached an 8.13-fold depth and 95.06% genome coverage, allowing us to call variants with high confidence. After quality control was performed, a final set of 5,904,820 high-quality SNPs (6.17 SNPs per kb) remained for subsequent analysis (Supplementary Table S2). With respect to the transcriptomic data, the clean data of 680 hepatic transcriptomic samples ranged from 5.24 to 10.51 G for each individual, with an average of 6.37 G. Similarly, the transcriptomic clean data of 655 duodenal mucosal samples varied between 5.40 and 8.87 G for each individual, with an average of 6.33 G (Supplementary Table S3). PCA revealed distinct clustering of samples from these two tissues (Supplementary Fig. S2A). A total of 4123 DEGs were identified (Supplementary Table S4); hepatic genes were enriched in cell communication, whereas duodenal genes were enriched in processes related to digestion and absorption (Supplementary Fig. S2B).

With respect to the previously published human colonic mucosal gene expression of 88 individuals in an IBD cohort, the average alignment rate of RNA-seq data was 87.11% for 88 samples, with the number of aligned reads per sample ranging from 14,365,657 to 31,530,487, with an average value of 22,586,647. For the 16S rRNA data, an average of 62,100 high-quality reads per sample was generated, ranging from 6559 to 173,471.

Genetic determinants of the gut microbiome

The microbiome is highly dynamic, influenced by various factors, and capable of adapting and evolving on its own, making it difficult to determine the extent to which the gut microbiota is shaped by the host genome. Hence, correlation analysis between host genetic kinship and intestinal microbial similarity was first performed on individuals for whom both genetic and microbial data were available. The calculation process was iterated 10,000 times but yielded an average correlation coefficient of nearly zero for each gut segment (Fig. 1A), indicating the weak correlation between host genetics and the gut microbiota from an overall perspective. However, this analysis was unable to identify the specific microorganisms regulated by host genetics; thus, we estimated the h^2 for each taxon separately, focusing on those with a detection rate greater than 30% (Supplementary Fig. S3). The results indicated that a modest proportion of taxa (8.80% for the duodenum, 4.00% for the jejunum, 14.08% for the ileum, 23.08% for the cecum, and 9.58% for feces at the genus level) presented significant SNP-based heritability in terms of both quantity and cumulative relative abundance (3.33% for the duodenum, 2.95% for the jejunum, 4.49% for the ileum, 9.97% for the cecum, and 10.17% for feces at the genus level) (Fig. 1B and Supplementary Table S5). This result suggested that host genetics played a moderate role in shaping the gut microbial composition in various intestinal segments, but the extent of this influence varied between segments. Several distinct taxa, for example, Christensenellaceae, Fusobacteriaceae, Chitinophagaceae, Solirubrobacteraceae, and Corynebacteriaceae at the family level, were regulated by host genetics in multiple intestinal segments, with the most significant heritability ranging from 0.135 to 0.365, and exhibited variations across different intestinal segments (Fig. 1C). The family Christensenellaceae in the cecum presented the highest heritability, reaching 0.365, with a detection rate of 100% and a relative abundance of 1.98%. Two distinct peaks on chromosome 4 were associated with microbial taxa in the family Christensenellaceae (Supplementary Fig. S4). One peak (4:73258088–4:7360984, chromosome:position) was associated with the regulation of the family Christensenellaceae and the genus *Christensenellaceae R-7 group*, and the other peak (4:8928910–4:8939639) was linked to the regulation of the family Christensenellaceae, the genus *f_Christensenellaceae; g_uncultured*, and the species *f_Christensenellaceae; g_uncultured; s_uncultured_bacterium*. The presence and abundance of these specific microorganisms exhibited a high dependency on host genomic variations. The SMR method was subsequently employed on the GWAS summary datasets for the family Christensenellaceae. Two genes (*COPS7B* and *SCG2*) were significantly associated with the first genomic peak (4:73258088–4:7360984), and one gene (*SEMA6D*) was associated with the second peak (4:8928910–4:8939639) (Supplementary Table S6). In total, genome-wide significant loci were identified for 11 microbial families, 24 genera, and 15 species, with the majority (3/11 at the family level, 11/24 at the genus level and 11/15 at the species level) originating from the cecum (Fig. 2). No pleiotropic SNPs were found to affect multiple microbial taxa at a single taxonomic level, whereas SNPs exhibiting pleiotropy were discovered to influence specific microbial taxa across various taxonomic levels, including family, genus, and species.

The intestinal transcriptome is superior in revealing host–microbial interactions

A combined analysis of the transcriptome and microbiome was conducted to provide a more holistic and precise understanding of the effects of host genetics on microorganisms, thus presenting a more rigorous

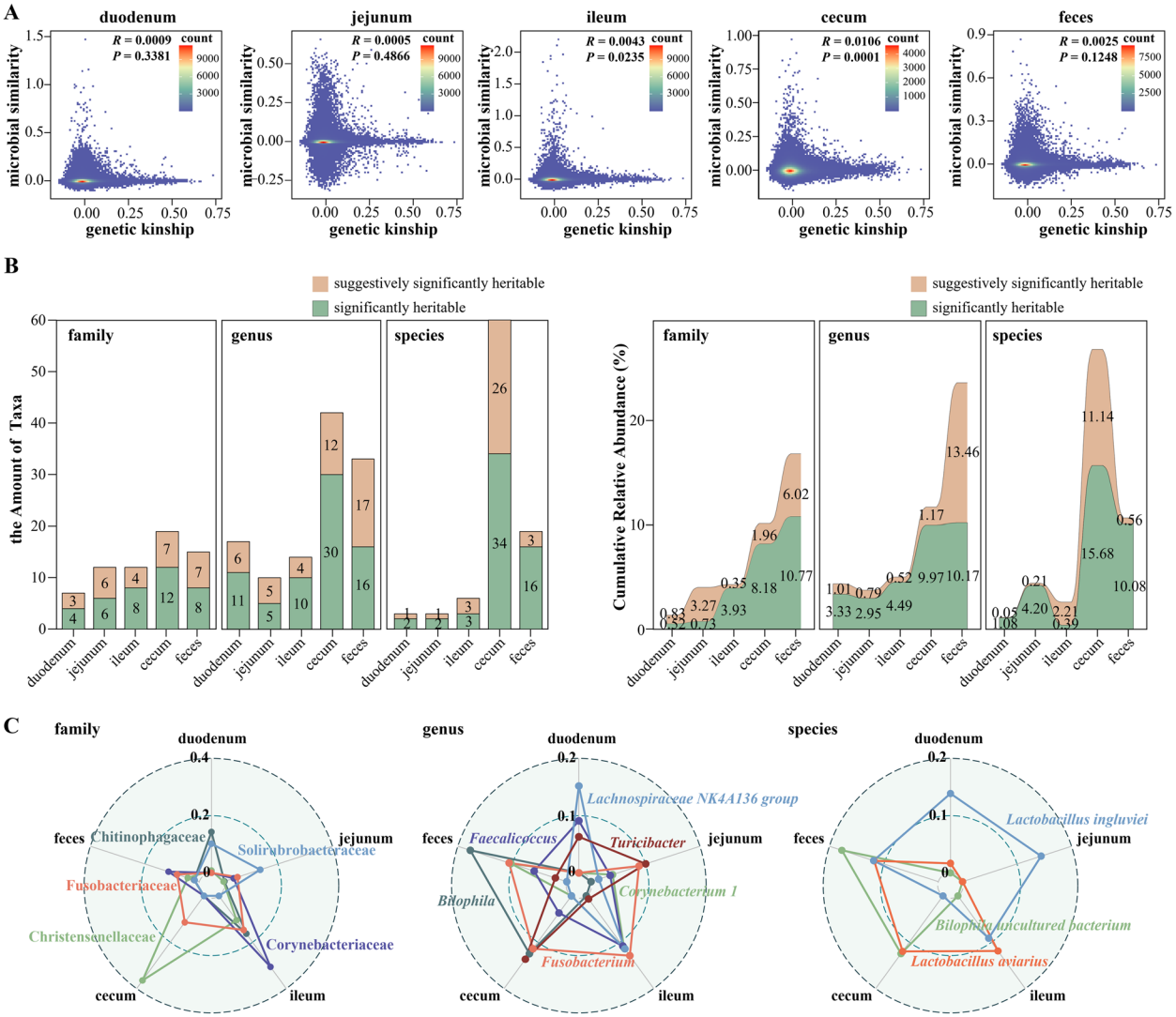


Fig. 1 The role of the host genome in shaping the gut microbiota. **A** Relationships between the host genetic kinship matrix and gut microbial similarity matrices. Scatter plots of the host genetic kinship of pairs of individuals (x-axis) and their microbial similarities according to the five sample sites (*y*-axis). **B** The amount (left) and cumulative relative abundance (right) of heritable microbial taxa grouped by sampling sites from the family to the species level. Green represents the significantly heritable (*p* value less than 0.05) microbes, and yellow represents suggestively significantly heritable (*p* value between 0.05 and 0.1) microbes. The numbers represent the corresponding amount and cumulative relative abundance (percentage). **C** SNP-based heritability estimates of microbial taxa at multiple gut sites from the family to the species level

(See figure on next page.)

Fig. 2 Microbial GWASs at the family (**A**), genus (**B**), and species (**C**) levels. The horizontal red and blue lines indicate the genome-wide significance ($p = 3.32 \times 10^{-7}$) and suggestive significance ($p = 6.63 \times 10^{-6}$) thresholds, respectively, in the Manhattan plots. Each point represents a SNP. Each taxon has been renamed systematically, reflecting its taxonomic level, gut segment, and sequential number of significant loci along the chromosomes. The first character indicates the taxonomic level, where F stands for family, G stands for genus, and S stands for species. The second number represents the order of significant loci on the chromosomes. The third character denotes the gut segment, with D, J, I, C, and F referring to the duodenum, jejunum, ileum, cecum, and feces, respectively. For example, F1: I_Christensenellaceae in **A** represents the family Christensenellaceae in the ileum, whereas G6: C_Ruminiclostridium in **B** denotes the genus *Ruminiclostridium* in the cecum

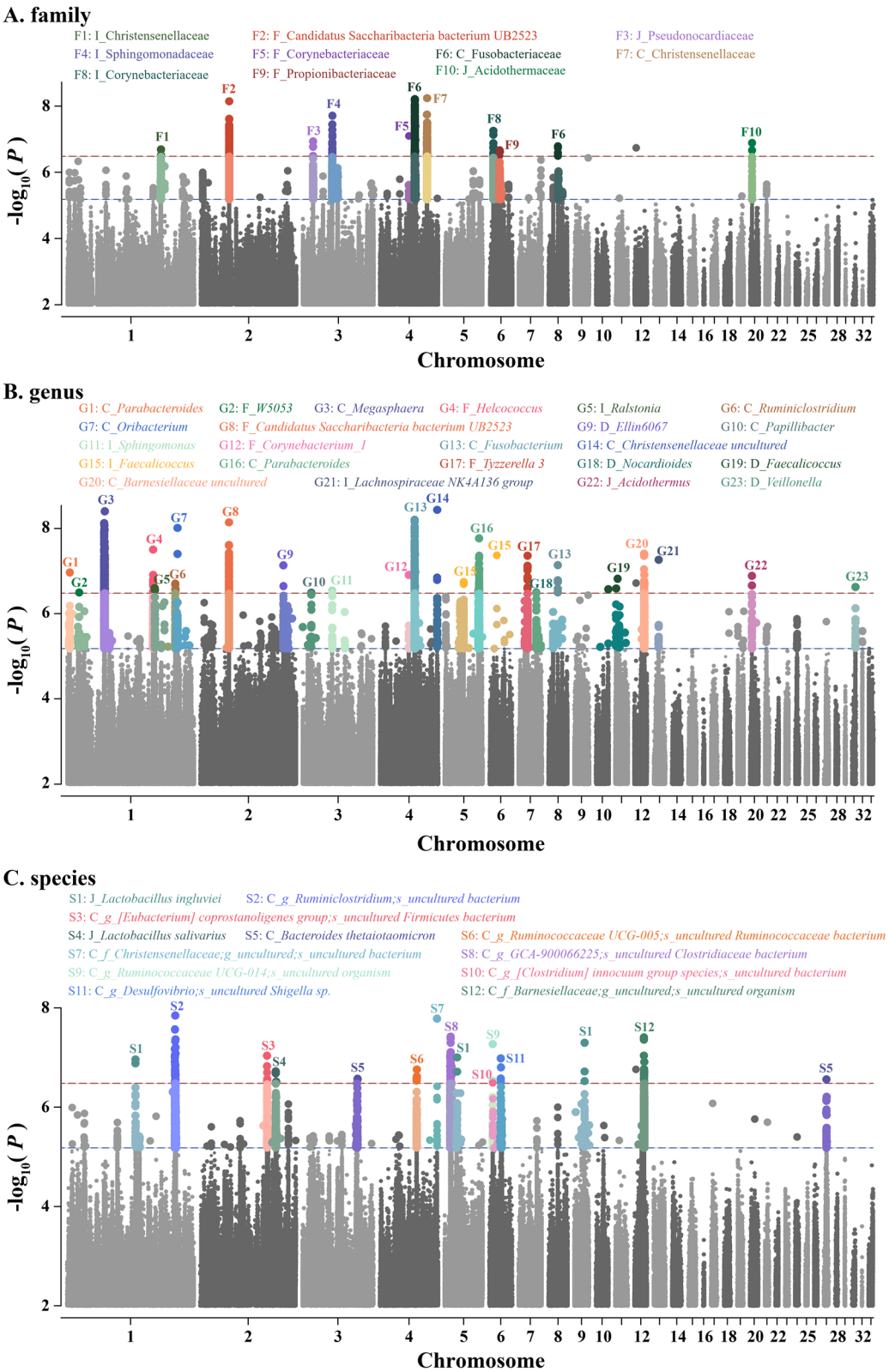


Fig. 2 (See legend on previous page.)

scientific approach for gaining profound insights into host–microbial interactions.

Duodenal gene expression significantly contributes to the variation in the duodenal microbial community

Procrustes analysis revealed a notable correlation between the gene expression profile in the duodenal mucosa and the structure of the duodenal microbial community ($M^2 = 0.9844$, $r = 0.1249$, $p = 0.001$) (Fig. 3A). Moreover, no significant correlation was detected between the hepatic gene expression profile and the duodenal microbial community structure ($M^2 = 0.9967$, $r = 0.0574$, $p = 0.172$) (Fig. 3A). Therefore, a more in-depth scientific approach is needed to study the interaction between host genetics and intestinal microbes from the perspective of gene expression levels.

Analogous to h^2, m^2, and r_b^2 were defined as the relative proportion of the total variance attributable to the gut microbial community and the gene expression profile, respectively. Higher values of m^2 and r_b^2 indicated more substantial roles of the microbial community and gene expression profile in shaping the phenotypes of target traits. Given that the mucosal gene expression profile of a specific digestive tract exhibited a close connection with the corresponding microbial community, we proceeded to estimate the m^2 of each duodenal mucosal gene and the r_b^2 of each duodenal taxon, resulting in an average m^2 of 0.18 and an average r_b^2 of 0.17. For the majority (86.27%) of genes, a greater m^2 (average, 0.21; range, 0–0.55) than h^2_{total} (average, 0.02; range, 0–0.34) was observed (Fig. 3B and Supplementary Table S7), suggesting that the microbiome contributed more to the expression of most mucosal genes than the genome. In addition, the combined sum of the m^2 and h^2_{total} values of some genes were greater than 1, indicating partial overlap and suggesting that gene expression in the duodenal mucosa was jointly regulated by host genome and gut microbiome interactions. Considering the microbial taxa as binary or quantitative phenotypes, the r_b^2 of the duodenal microbial families, genera and species were estimated, and the taxa most highly regulated by gene expression were identified from the top 3 phyla, namely, Firmicutes, Proteobacteria, and Bacteroidetes (Fig. 3C). The impact of gene expression on shaping microbes was significantly more pronounced than that of host genomic variation, as evidenced by the notably larger r_b^2 compared with h^2 (0.17 ± 0.01 vs. 0.02 ± 0.00, Fig. 3C and Supplementary Table S5). At various taxonomic levels, particularly among those microbes with a detection rate exceeding 60%, the family Bifidobacteriaceae (0.65), the genus *Bifidobacterium* (0.63), and the species *Lactobacillus vaginalis* (0.54) presented the most pronounced r_b^2 values. Notably, the taxon chain in the phylum Firmicutes, including the family Lactobacillaceae, genus *Lactobacillus*, and species *L. vaginalis* and *L. murinus*, was largely regulated by gene expression, with r_b^2 values of 0.46, 0.46, 0.54, and 0.27, respectively.

Comparison of host genetics and intestinal microbial interaction patterns between chickens and humans

Based on orthologous genes between chickens and humans, a comparative analysis was undertaken to elucidate the patterns of host genetics and intestinal microbial interactions in both chickens and humans. Therefore, in addition to the multiomics data generated in this study, we utilized previously generated and published human colonic mucosal gene expression and gut microbiome data from 88 individuals in an IBD cohort.

The gene expression profile of the human colonic mucosa was not significantly correlated with the colonic microbial community (Fig. 4A; $M^2 = 0.9691$, $r = 0.1558$, $p = 0.210$), and the pattern was distinct from that observed in the chicken duodenum (Fig. 3A). Furthermore, the m^2 of each mucosal gene and the r_b^2 of each microbial taxon were simultaneously estimated (Supplementary Tables S8 and S9). The taxa were categorized into five groups according to their r_b^2 values, with the group exhibiting minor values representing the largest proportion (Fig. 4B, C). Specifically, the microbial taxa with r_b^2 values less than 0.1 accounted for 44.90% and 81.19% of the total microbial taxa in chickens and humans, respectively. These results strongly suggested the limited influence of host genetics on the overall microbial community in both humans and chickens. However, the regulatory effects of the gene expression profile on the gut microbiota in the chicken duodenum were relatively more substantial. Specifically, the average r_b^2 value of chicken microbial taxa (0.17) differed significantly from that of human microbial taxa (0.04) (Supplementary Tables S5 and S8). Moreover, 4.08% of the chicken duodenal taxa and 0.99% of the human colonic taxa had high r_b^2 values (>0.4), and 24.49% and 2.98% had moderate r_b^2 values (0.2–0.4). *P_Saccharibacteria;s_*, formerly known as TM7, in the human colon, and *L. vaginalis* in the chicken duodenum were the species most significantly regulated by mucosal genes, with remarkably high r_b^2 values of 0.41 and 0.54, respectively. In general, the proportion of shared species, including *Bacteroides uniformis*, *B. fragilis*, and *Parabacteroides distasonis*, was relatively low, and the r_b^2 differed significantly between chickens and humans (*B. uniformis*, 0.121 vs. 0.079; *B. fragilis*, 0.193 vs. 0.091; *P. distasonis*, 0.004 vs. 0.000) (Fig. 4B). Similarly, the gene set was also classified into five groups according to their m^2 values. The average m^2 value of chickens (0.18) was analogous to that of humans (0.20) (Supplementary Tables S7.2 and S9). However, genes in chickens were uniformly distributed

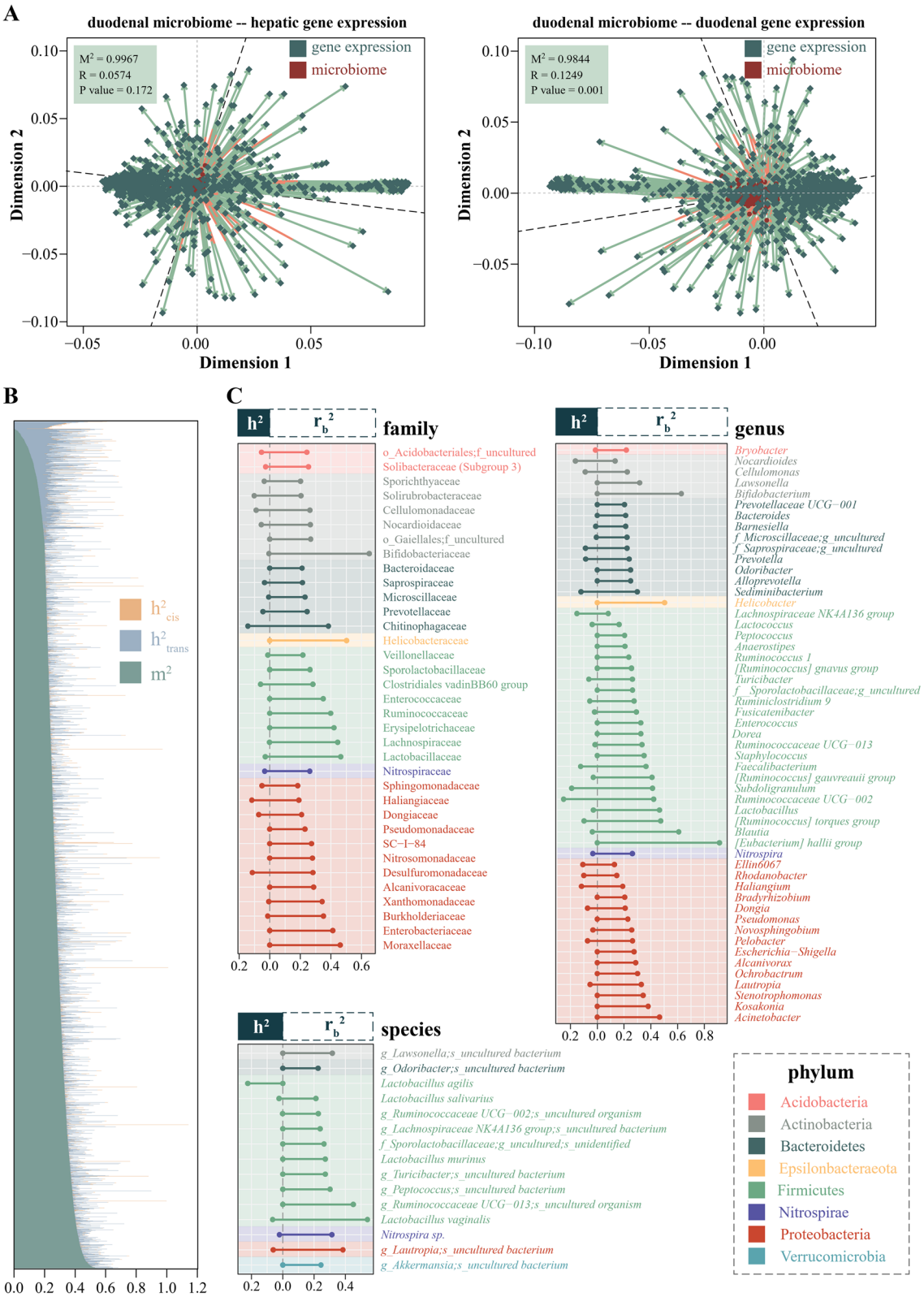


Fig. 3 Mutual influences between the duodenal microbiota and mucosal gene expression. **A** Procrustes analysis of the gene expression profile and duodenal microbial community. **B** Comparison of the heritability (both cis- and trans-) and microbiability of duodenal mucosal genes. Each line represents a gene. **C** Heritability (left) and regulatability (right) estimates of the duodenal microbiota from the family level to the species level

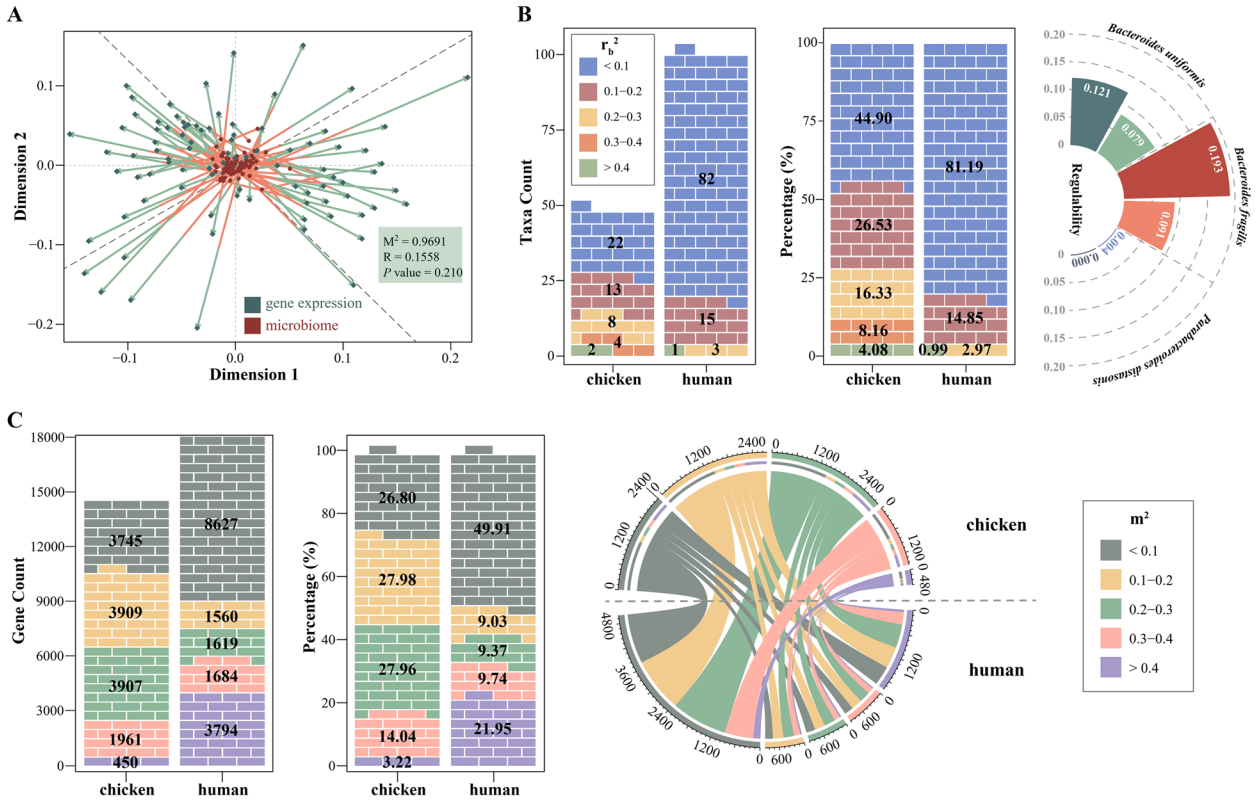


Fig. 4 Comparative analysis of host genetics and intestinal microbial interaction patterns between chickens and humans. **A** Procrustes analysis between the colonic gene expression profile and the microbial community in humans. **B** Comparison of the regulatability estimates of microbial species between chickens and humans. Shared taxa were specifically visualized. **C** Comparison of the microbiability estimates of mucosal genes between chickens and humans

across these groups, with only 3.22% exhibiting high m^2 values (> 0.4), whereas the majority of human genes presented m^2 values below 0.1 (49.91%) or above 0.4 (21.95%) (Fig. 4C). For the shared genes, a weak correlation with coefficients of 0.02 (both Pearson and Spearman) was detected between humans and chickens. However, the extent of overlap between chickens and humans within a specific group had a clear positive correlation with the size of the group, and the average correlation coefficient was 0.983 for chickens and 0.996 for humans.

Trans-eQTLs are considerable contributors to gene expression

The above results underscored the predominant role of gene expression over genetic variation in the regulation of the gut microbiota. This finding prompted an inquiry: How do genetic variations govern gene expression? To explore the effects of host genetics on hepatic and duodenal gene expression, we examined the correlation between host genetic kinship and gene expression similarities. The Spearman correlation coefficients of the GRM with either of the two RRMs were close to zero (Fig. 5A). We distinguished between the genetic components close to the gene (h^2_{cis}) and away from the gene (h^2_{trans}). Together, h^2_{cis} and h^2_{trans} constituted the total heritability (h^2_{total}). The heritability values of 12,191 genes in the liver and 13,654 genes in the duodenum were estimated, the majority of which presented relatively low h^2_{cis} and h^2_{trans} values and more than 60% were identified as non-heritable ($h^2_{\text{total}} < 0.01$) (Fig. 5B and Supplementary Table S7). The average contributions of h^2_{cis} to h^2_{total} (mean h^2_{cis}/(mean h^2_{cis} + mean h^2_{trans})) were 17.62% (liver) and 14.96% (duodenum), ranging from 0 to 100%. Conversely, loci outside the cis window of a gene explained a significant proportion of h^2_{total}. The correlations between h^2_{cis} and h^2_{trans}, h^2_{cis} and h^2_{total}, and h^2_{trans} and h^2_{total} were 0.06, 0.40, and 0.94, respectively, in the liver and 0.05, 0.42, and 0.93, respectively, in the duodenum ($p < 0.01$).

To gain an extensive understanding of genetic regulation within the chicken transcriptome, we conducted a thorough mapping of cis- and trans-eQTLs in both the liver and duodenal mucosa. Our findings revealed that cis-regulated genes accounted for only a minor fraction of the total genetic regulation, whereas trans-regulated

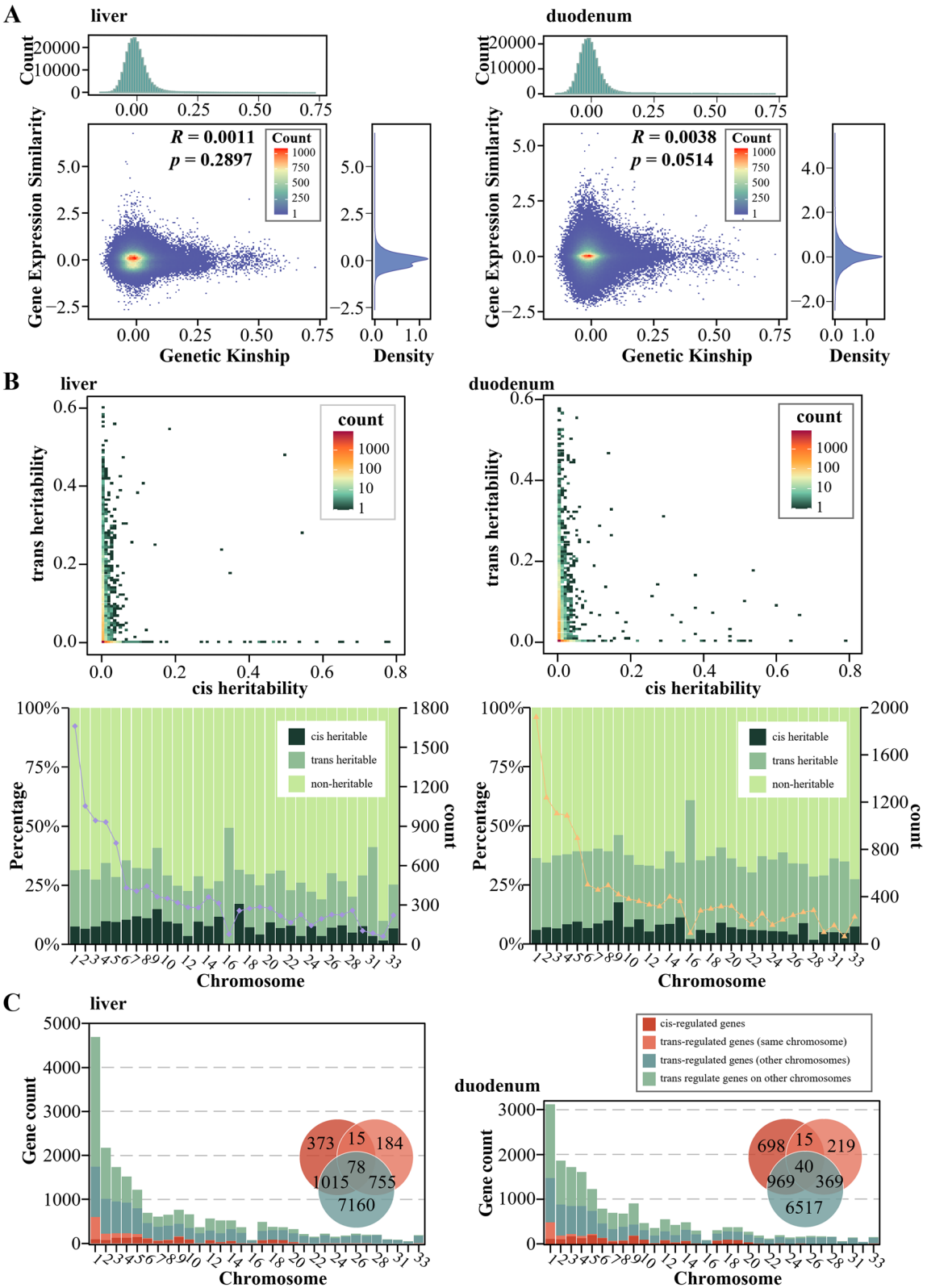


Fig. 5 Relationships between genetics and gene expression. **A** Correlation of the host genetic kinship and gene expression similarity in the liver and duodenum. **B** Comparison of h^2_{cis} and h^2_{trans} and distribution of genes with different heritable patterns in the liver and duodenum. **C** Statistics of hepatic and duodenal mucosal genes with different regulatory patterns categorized by chromosome

genes constituted the majority (Fig. 5C), emphasizing the limited influence of cis-eQTLs on gene expression and the more substantial effects of trans-eQTLs. Additionally, the number of genes regulated by each chromosome was positively correlated with chromosome size. Notably, cis-regulated genes were predominantly located on the first 20 chromosomes, with the duodenal mucosa showing a slightly greater proportion than the liver (Fig. 5C).

Deciphering the regulatory mechanisms of complex traits: Joint contributions of duodenal mucosal genes and the microbiota

The estimations of SNP-based h^2, m^2, and r_b^2 were successively performed for the most complex traits of chickens, including AFP, AFW, BW, EN, EW, FCR, RFI, LW, LP, HTC, HTG, HTBA, STC, STG, STBA, HDL, LDL and VLDL, followed by GWAS and SMR analysis (Supplementary Table S10). In parallel, GWAS and SMR analyses were conducted on the microbial taxa (14 phyla, 18 classes, 44 orders, 59 families, 98 genera, and 95 species) (Supplementary Table S11). This extensive analysis revealed several intriguing findings regarding the regulatory effects of duodenal mucosal genes on host phenotypes and microbial taxa (Fig. 6A and Supplementary Table S12). Complex traits involving complicated biological processes and microbial taxa at various taxonomic levels were characterized by polygenic inheritance—a consequence of the collective influence of numerous genes. *SLC1A4*, *COMMD1*, *FAM179A*, and *VRK2* were identified as potential candidate genes involved in the

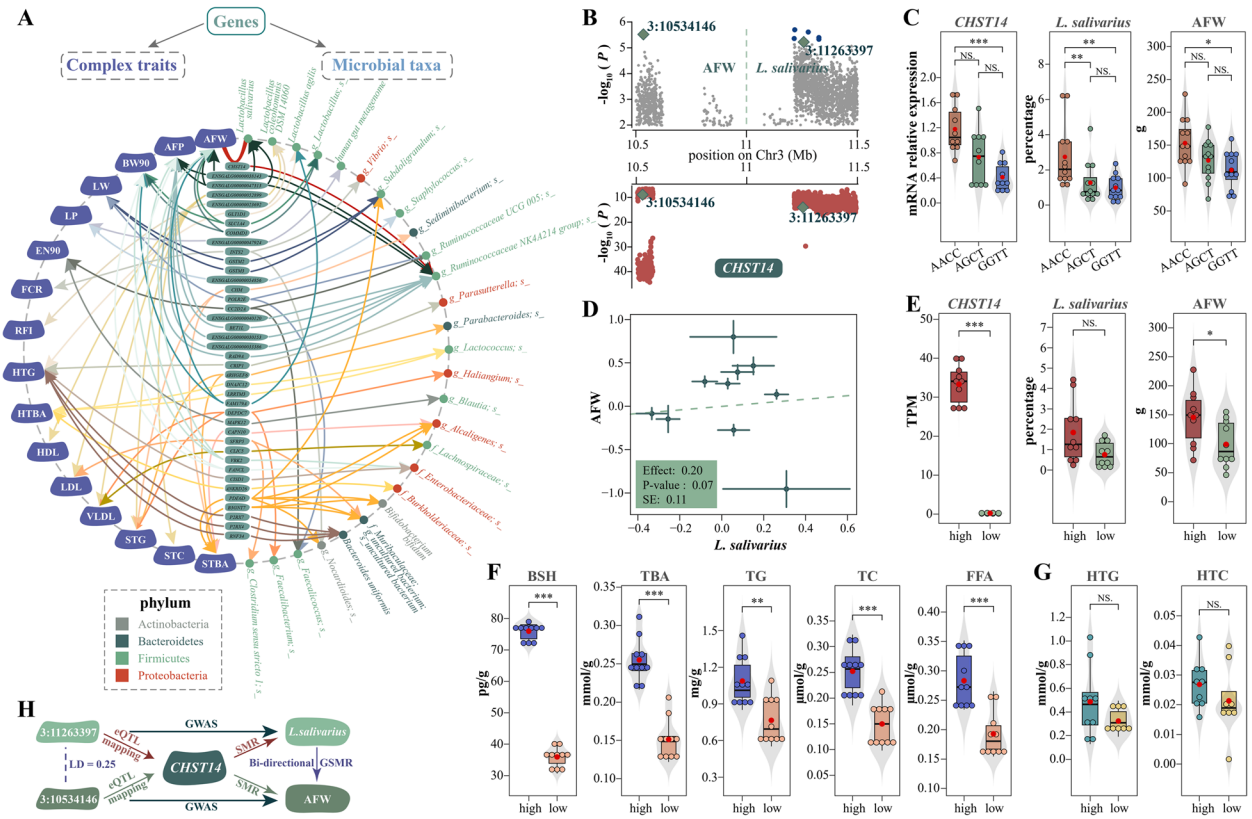


Fig. 6 Joint contributions of duodenal mucosal genes and the microbiota to host phenotypes. **A** Regulatory effects of duodenal mucosal genes on host phenotypes and microbial taxa at the species level. **B** Manhattan plots of GWASs for AFW and *Lactobacillus salivarius*. The gray, dark blue, and reddish-brown dots indicate non-significant, suggestively significant and significant SNPs, respectively. Colocalization of trans-eQTLs of the duodenal mucosal gene *CHST14* identified two colocalized SNPs, highlighted by green diamonds. **C** Boxplots showing *CHST14* expression levels, AFW, and *L. salivarius* abundance across the three genotypes of the two eVariants (3:10534146 and 3:11263397). **D** Putative causal association between AFW and *L. salivarius* generated by GSMR. **E** Individuals with high *CHST14* expression exhibited an increased relative abundance of *L. salivarius* and higher AFW. **F** Levels of BSH, TBA, TG, TC, and FFA in the duodenal chyme were elevated in individuals with high *CHST14* expression. **G** HTG and HTC were upregulated in individuals with high *CHST14* expression. **H** Summarized results of the relationships among *CHST14*, *L. salivarius*, and AFW. AFW: abdominal fat weight; AFP: abdominal fat percentage; BW90: body weight at age 90 weeks; LW: liver weight; LP: liver percentage; EN90: total egg number until the age of 90 weeks; FCR: feed conversion ratio; RFI: residual feed intake; HTG: hepatic triglyceride; HTBA: hepatic total bile acid; HDL: serum high-density lipoprotein; LDL: serum low-density lipoprotein; VLDL: serum very low-density lipoprotein; STG: serum triglyceride; STC: serum total cholesterol; STBA: serum total bile acid; BSH: bile salt hydrolase; TBA: duodenal total bile acid; TG: duodenal triglyceride; TC: duodenal total cholesterol; FFA: duodenal free fatty acid; HTG: hepatic triglyceride; HTC: hepatic total cholesterol

regulation of BW and AFW. *ENSGALG00000052899* was associated with several serum biochemical indicators, including HDL, LDL, VLDL, STG, and STC. In addition, mucosal genes played pivotal roles in regulating the gut microbiota. The microbial taxa within Firmicutes, especially *L. salivarius* and the *Ruminococcaceae NK4A214 group*, were strongly influenced by the genes *CHST14*, *ENSGALG00000047313*, and *ENSGALG00000038143*. Notably, the simultaneous examination of genetic variants, gene expression, and the microbiota revealed complex interactions underlying trait and microbial regulation. *PDE6D* and *B3GNT7* concurrently regulated both STBA and *Bifidobacterium bifidum*. Additionally, the *P2RX4*, *P2RX7*, and *RNF34* genes were identified as factors influencing HTG and *B. uniformis*.

Below, AFW was used as an example for further elaboration, focusing on the interplay of the *CHST14* gene and *L. salivarius*. The SNP 3:10534146 was found to be suggestively significantly associated with AFW (Fig. 6B). eQTL mapping and SMR analysis were subsequently performed, revealing a significant association between the *CHST14* gene and SNPs 3:10534146 and 3:11263397 (Fig. 6B). Notably, the SNP 3:11263397 was strongly correlated with *L. salivarius* according to the GWAS results (Fig. 6B). To validate the regulatory roles of the SNPs 3:10534146 and 3:11263397 in regulating *CHST14* expression, we performed RT−qPCR on 10 individuals selected from three genotypes. Additionally, we assessed the relative abundance of *L. salivarius* and AFW across groups with different genotypes. Our analysis revealed significant differences in *CHST14* expression levels, *L. salivarius* abundance, and AFW among individuals with different genotypes for these two variants (Fig. 6C, $p < 0.05$, Wilcoxon rank-sum test). Notably, the linkage disequilibrium (LD) between these two SNPs was 0.25. Moreover, the bidirectional GSMR method was used to investigate a putative causal association between AFW and *L. salivarius*. The findings revealed a nonsignificant p value ($p > 0.1$) in the forward analysis but a suggestively significant p value ($p = 0.07$) in the reverse analysis, indicating an estimated effect of *L. salivarius* on AFW of 0.20 (Fig. 6D).

Validation of the coordinated roles of CHST14 and L. salivarius in regulating abdominal fat deposition

Ten individuals from each of the high- and low-expression tails of CHST14 were selected to investigate the regulatory mechanism involving *CHST14*, *L. salivarius*, and AFW. *L. salivarius* exhibited an enhanced abundance in the high-*CHST14* group (Fig. 6E and Supplementary Table S13, $p < 0.1$, Wilcoxon rank-sum test). Given that *L. salivarius* is a primary producer of bile salt hydrolase (BSH), we quantified BSH levels and found its significantly higher levels in the high-CHST14 group (Fig. 6F and Supplementary Table S13, $p < 0.05$, Wilcoxon rank-sum test), reinforcing the role of *L. salivarius* in promoting BSH production. Due to the vital functions of BSH in deconjugating bile acids and thereby altering the bile acid pool, we subsequently measured the total bile acid (TBA) in the duodenal chyme. The results showed that individuals with higher BSH levels indeed exhibited significantly elevated TBA levels (Fig. 6F and Supplementary Table S13, $p < 0.05$, Wilcoxon rank-sum test). The increased TBA can further facilitate lipid emulsification, leading to elevated levels of triglycerides (TG), cholesterol, and fatty acids (FA) (Fig. 6F and Supplementary Table S13, $p < 0.05$, Wilcoxon rank-sum test). The newly produced cholesterol and TG can be packaged into chylomicrons within intestinal epithelial cells and transported to the circulatory system, explaining the observed increase in hepatic TG (HTG) and total cholesterol (HTC) in the high-CHST14 group (Fig. 6G and Supplementary Table S13, $p > 0.05$, Wilcoxon rank-sum test). Excess cholesterol can be esterified by cholesterol esterases, leading to storage in lipid droplets, which contributes to fat deposition and increased AFW (Fig. 6E and Supplementary Table S13, $p < 0.05$, Wilcoxon rank-sum test). This finding highlighted the potential influence of *CHST14* not only on AFW but also on broader metabolic processes through its regulatory effect on *L. salivarius* abundance and activity. Although these results warrant cautious interpretation, they support the hypothesis that *CHST14* functions as a critical gene modulating abdominal fat deposition by influencing the abundance and activity of *L. salivarius* and its downstream effects on bile acid metabolism and lipid homeostasis (Fig. 6H).

Discussion

An increasing number of studies in host genetics and gut microbiota support and explore the multifaceted interplay among host genetics, the gut microbiome, and their combined influence on health [19, 69–71], metabolism [30, 72], and behaviors [73–75]. We employed high-throughput sequencing, multiomics analyses, and advanced statistical methods and technologies to elucidate these intricate relationships, shedding light on the mechanisms underlying health, disease, and potential therapeutic interventions related to the host–gut microbiota axis.

The gut microbiome is regarded as the second genome of the host [76], indicating the extensive genetic and functional diversity of microorganisms in the gastrointestinal tract and their significant influence on biological processes, including metabolism [77–81], immune function [82], and disease susceptibility [83, 84]. Heritability estimation and GWASs indicated that the quantity and

cumulative relative abundance of microbial taxa shaped by the host genome accounted for only a small fraction (8.80% for the duodenum, 4.00% for the jejunum, 14.08% for the ileum, 23.08% for the cecum and 9.58% for feces at the genus level) of the total taxa, which were consistent with the findings of a previous study in broilers [29]. The extent to which host genetic variations determined microbial abundance was dynamic across different gut sites. Metabolic and nutritional factors [85], selective pressure [86], coevolution [87–89], and microbial genetic adaptation [90, 91] jointly contributed to the distinct dependence of specific microbes on the host genome. The family Christensenellaceae had the highest heritability estimate (0.365), which was consistent with the original microbial heritability findings from a twin-based landmark study by Goodrich et al. [92] (0.39), the extended TwinsUK cohort [16] (0.42), and a study involving 3666 twins from the American Gut project, Flemish Gut Flora Project and the extended TwinsUK cohort [93] (0.31). The family Christensenellaceae played crucial ecological roles in the host's gut, significantly influencing metabolism and overall health [94, 95]. This taxon remained relatively stable within the gut microenvironment, exhibiting high abundance and a 100% detection rate, which minimized the influence of environmental factors and naturally increases its heritability. In addition to its presence in chickens and pigs [23], the family Christensenellaceae was widespread across diverse human populations, with reports in individuals from North America [18], South America [96], Europe [97], Asia [98], Africa [99], and Australia [100]. Throughout evolution, selective pressure exerted by the host may have enhanced the adaptability of Christensenellaceae [101], leading to the retention and reinforcement of its genetic characteristics. The Christensenellaceae family has been reported to play a protective role in maintaining a healthy body mass index (BMI), and individuals with a higher abundance of Christensenellaceae tended to have a lower BMI and less body fat [92, 93, 102, 103]. Additionally, studies have suggested that Christensenellaceae may support a leaner body composition by influencing metabolic pathways and interacting with other beneficial microbes within a broader microbial network that impacts metabolism, energy extraction from food and fat storage [92, 104, 105]. However, the regulatory mechanisms of Christensenellaceae remain unclear. We integrated multiomics data, revealing significant regulatory loci on chromosome 4 and identifying the candidate genes *COPS7B*, *SCG2*, and *SEMA6D* associated with these significant loci, among which *COPS7B* was positively associated with several lipoproteins [106]. Furthermore, individuals with a relatively high abundance of Christensenellaceae presented significantly high levels of HDL [95]. We hypothesized that the *COPS7B* gene regulates the relative abundance of Christensenellaceae by altering the concentration of intestinal lipoproteins.

To gain a better understanding of the intricate interaction between the host and the gut microbiota, it is crucial to determine whether the genome or the intestinal mucosal transcriptome is more influential in host–microbiome interactions. Host–microbial interactions are inherently dynamic and subject to changes over time [107, 108]. The intestinal mucosal transcriptome provides a snapshot of the active genetic processes influenced by the presence of the microbiome, offering a more detailed and dynamic perspective [109–112] on host–microbial interactions than the static genomic sequence does. Moreover, microbial metabolites are capable of engaging host cellular receptors, serving as messengers of information between the intestinal microbiota and host cells [113–116]; correspondingly, the growth of the intestinal microbiota could be impacted by the secretion of metabolites by host mucosal cells [14, 117]. In our study, the duodenal microbial community was more significantly associated with the duodenal mucosal gene expression profile than the host genome, which was indicated by a much greater correlation coefficient. We posited that closer proximity engenders a more robust connection.

The concept of "regulatability", which we introduced in this study, acted as a valuable metric for quantifying the impact of gene expression profiles on host phenotypes. Higher r_b^2 values indicated a more substantial role of gene expression regulation in shaping the given phenotypes. This concept was derived from h^2 and m^2. On the basis of h^2 and m^2 estimates of duodenal mucosal genes, we found that the duodenal microbial community emerged as a more influential predictor of gene expression than genomic variation. The comparison of the h^2 and r_b^2 estimates of the duodenal microbiota revealed the conspicuous role of duodenal genes as greater contributors than genomic variations. Despite the convenience and relative non-invasiveness of sampling for DNA extraction, the connections derived from the host genome and the gut microbiota did not comprehensively capture the diverse and complicated host–microbe interactions that have been extensively discussed but poorly understood, emphasizing the superior role of the intestinal mucosal transcriptome in the exploration of host–microbe interactions. Specifically, taxa in Bifidobacteriaceae and Lactobacillaceae were predominantly regulated by gene expression, exhibiting exceptionally high regulatability. Among these taxa, *Bifidobacterium* was strongly associated with the *LCT* gene in both the Human Microbiome Project (HMP) and the TwinsUK datasets [16, 60, 118]. However, understanding the regulatory mechanisms of

Lactobacillus remains a complex and ongoing endeavor that requires further exploration and investigation.

Modest correlations were detected between the host intestinal microbial community and mucosal gene expression in the chicken duodenum (0.1249) and human colon (0.1558). In addition, the average m^2 value of the human genes (0.20) was close to that of the chicken genes (0.18), and the microbial taxa with relatively low r_b^2 values accounted for the largest proportion of total taxa (44.90% in chickens and 81.19% in humans), collectively validating the broad impact of the gut microbiota on host physiology [119, 120] and the limited influence of host genetics on the colonization of gut microbes [16, 17, 60]. The correlation coefficient for genes shared between humans and chickens was notably low at 0.02, suggesting the limited similarity in specific gene–microbiota interactions between these two species. Notably, the average r_b^2 value of microbes was significantly greater in chickens (0.17) than in humans (0.04). Additionally, genes with m^2 values greater than 0.4 accounted for a substantially larger proportion in chickens (21.95%) than in humans (3.22%). These differences in gene–microbiota interactions can likely be attributed to multiple factors, including species-specific characteristics, variations in diet, anatomical distinctions in intestinal regions, and differences in sample size. Cross-species investigations not only illuminated these disparities but also provided a framework for identifying both universal principles and unique species-specific factors [121]. These findings were crucial for understanding the complex interplay between host genetics and the intestinal microbiota and highlighted the ways in which evolutionary and environmental factors shaped these interactions in different species [122, 123].

Cis- and trans-eQTLs refer to two distinct types of genetic variants that influence gene expression levels that are differentiated based on their physical proximity to the regulated gene [39]. Determination of the explicit effects of trans-eQTLs can be challenging owing to the complexity of the networks they influence [124, 125]. However, eQTL analysis could reveal trans hotspots, which are loci with widespread transcriptional effects, contributing to our understanding of the intricate regulatory mechanisms governing gene expression [126]. Unlike cis-eQTLs, which primarily affect genes in close proximity, trans-eQTLs can influence the expression of genes located on entirely different chromosomes or distant genomic regions. Trans-eQTLs were considered significant contributors to gene expression due to their superior ability to affect multiple genes throughout the entire genome [127], and the combination of many weak trans-eQTL effects was able to explain the major proportion of gene expression heritability [128]. Higher values of h^2_{trans} compared to h^2_{cis}, along with a larger proportion of trans-regulated genes in contrast to cis-regulated genes, indicated that genomic loci outside the cis window played a more significant role in determining gene expression levels, thus confirming and supporting the earlier findings by Wright et al. [129] (relative contribution of h^2_{cis} to h^2_{total}, 23%), Grundberg et al. [130] (0.36) and Klaasjan et al. [131] (20%) in humans and Frank et al. [132] (28%) in yeast. These results implied that genomic loci outside of the cis window had a more prominent role in determining gene expression. Given that SNPs located outside the TSS ± 5 Mbps window vastly outnumber those within the TSS ± 1 Mbps window, it followed that cis-eQTLs exerted a more localized influence, whereas trans-eQTLs contributed broadly to gene expression, highlighting their importance in complex trait regulation. Additionally, many genes do not operate in isolation but are part of intricate regulatory networks. Trans-eQTLs could modulate the expression of genes involved in these networks, leading to coordinated changes in multiple genes that work together [133–135] and providing functional diversity in a range of biological processes [136]. Complex traits and diseases are frequently influenced by the collaborative action of multiple genes. Overall, while cis-eQTLs had a more localized impact, trans-eQTLs exerted a broader influence on gene expression patterns, establishing them as crucial contributors to the regulation of complex biological processes.

Previous studies have revealed that *L. salivarius* is ubiquitous in saliva and the digestive tract and serves as a potent utilizer of acetate, butyrate, propionate and lactate from carbohydrate metabolism [137, 138]. Additionally, *L. salivarius* is a notable producer of BSH [139], a critical enzyme that catalyzes the deconjugation of bile acids by removing glycine or taurine moieties, thereby converting conjugated bile acids into free bile acids [140]. Consistent with these findings, our results revealed elevated BSH levels in the high-*CHST14* group. The increased enzymatic activity of BSH necessitates greater substrate availability and plays a pivotal role in modifying the bile acid pool composition [141], which is consistent with the elevated TBA levels observed. TBA plays a crucial role in lipid emulsification, facilitating the digestion of lipids into TG, cholesterol, and fatty acids [142]. In individuals with higher TBA levels, we observed enrichments of TG, TC, and FFA in the duodenal chyme. Additionally, the packaging of TC and TG into chylomicrons for transport into the circulatory system further reinforces the connection between bile acid dynamics and lipid metabolism [143], which explains why individuals with greater *CHST14* expression presented increased HTG and HTC levels. Excess cholesterol might also undergo esterification, leading to the formation of cholesterol esters stored in lipid

droplets and subsequently fat deposits [144]. Collectively, the increase in TG [145] and cholesterol esters [146] likely drives the accumulation of fat, further increasing abdominal fat deposition. *L. salivarius* has been confirmed to potentially alleviate hepatic fat deposition in chickens [147] and exerts anti-obesity effects [148, 149]. *CHST14*, a promising gene encoding carbohydrate sulfotransferase, plays a pivotal role in catalyzing the sulfonic modification of glycolipids, glycoproteins, and various biomolecules [150, 151]. This modification, characterized by its complex chemistry, functions as a regulatory force in the intricate domain of cell communication, orchestrating essential molecular interactions for cellular function and coordination. Our hypothesis suggested that the *CHST14* gene regulated abdominal fat deposition by influencing the relative abundance of *L. salivarius*. This hypothesis aimed to reveal the complex causal relationship among *CHST14*, *L. salivarius*, and AFW, elucidating the molecular basis of their interaction in the context of abdominal fat deposition.

Conclusions

Our research represents significant progress in elucidating the complex interplay between host genetics and the gut microbiome. Although our analysis revealed a limited influence of host genetics on the overall microbial community, duodenal mucosal gene expression played a substantial regulatory role in microbial taxa at various taxonomic levels. We utilized innovative concepts such as m^2 and r_b^2 to precisely assess the impact of gene expression and microbiome on each other, thereby clarifying the intricate nature of host–microbe interactions. High r_b^2 in specific microbial families and genera emphasized the nuanced and complex nature of host–microbe interactions and provided a holistic view of the combined impact of the microbiota and gene expression on phenotypes. Moreover, our investigation of complex traits, exemplified by the clarification of *CHST14* and its regulatory relationship with duodenal *L. salivarius,* provided deeper insights into the mechanisms controlling AFW. These findings have implications for the mechanisms regulating phenotypic outcomes and may aid in the development of strategies to modulate host–microbe interactions for improved health.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40168-025-02054-5.

Supplementary Material 1: Figure S1. Comparative analysis of gene expression patterns between the liver and duodenum. (A) Visualization of the variance in gene expression between the liver and duodenum according to PCA. (B) Volcano plot showing DEGs between the liver and duodenum. (C) The top 20 significantly enriched biological processes in the liver and duodenum.

Supplementary Material 2: Figure S2. Spatial dynamics of microbial diversity and community composition across diverse gut segments in chickens. (A) Number of taxa with high-quality ASVs classified from phylum to species. (B) Alpha diversity was represented by the Shannon diversity index, and ANOVA was used to determine significant differences (***$P < 0.001$). (C) Principal coordinate analysis (PCoA) plot generated using OTU metrics based on Bray–Curtis dissimilarities. Each point represents a sample. (D) Relative abundance of the dominant microbial phyla in different gut segments. (E) Relative abundance of dominant genera in different gut segments.

Supplementary Material 3: Figure S3. The cumulative relative abundances of microbial taxa with different detection rates in specific segments at the family, genus and species levels.

Supplementary Material 4: Figure S4. Circular Manhattan plots of GWAS results for microbial taxa in the family Christensenellaceae, with concentric layers representing, from inside to outside, the family Christensenellaceae, genus Christensenellaceae R-7 group, genus f_Christensenellaceae; g_uncultured; s_, and species f_Christensenellaceae; g_uncultured; s_uncultured_bacterium. The blue dots represent suggestively significant SNPs, whereas the red dots indicate significant SNPs. Chromosomes 3 and 4 feature peaks that simultaneously regulate microbial taxa across multiple taxonomic levels.

Supplementary Material 5: Table S1. Summary statistics of 16S rRNA gene sequencing. Table S2. Summary statistics of whole-genome resequencing. Table S3.1 Summary statistics of hepatic transcriptome sequencing. Table S4. Differentially expressed genes between liver and duodenal mucosa. Table S5. Estimation of heritability (h2) and regulatability (rb2) of duodenal microbiota. Table S6. SMR test based on the GWAS result of family Christensenellaceae. Table S7.1 Estimation of heritability (h2) and microbiability (m2) of hepatic genes. Table S8. Estimation of and regulatability (rb2) of human colonic microbial species. Table S9. Estimation of microbiability (m2) of human colonic mucosal genes. Table S10. SMR test based on the GWAS results of all recorded phenotypes. Table S11. SMR test based on the GWAS results of duodenal microbiota from phylum to species. Table S12. Regulatory effect of duodenal mucosal genes on the host phenotypes and microbiota from phylum to species.

Authors' contributions

NY and CS conceived the study, participated in the experimental design and critical discussion, and jointly supervised this work. FL and QZ performed the experiments and wrote the manuscript. CS, FL, QZ, XL, CW, GW, GL and YY participated in the management of the experimental animals and sample collection. FL, XW, QZ, JJ and WZ contributed to the measurement of biochemical traits and validation experiments. FL, QZ, XL and CS conducted the bioinformatics and statistical analyses. NY, CS and FL designed the figures and tables. NY and CS were responsible for critical revisions of the manuscript drafts. All the authors read and approved the final manuscript.

Funding

This work was supported by the National Key Research and Development Program of China (2021YFD1300600 and 2022YFF1000204), the National Natural Science Foundation of China (No. 31930105), China Agriculture Research Systems [CARS-40] and the 2115 Talent Development Program of China Agricultural University (00109015).

Data availability

Whole-genome resequencing data are available on the NCBI Sequence Read Archive (SRA) under accession PRJNA980845, and RNA-Seq data are available on the SRA under the accession PRJNA954522 (liver) and PRJNA1077588 (duodenum). 16S rRNA sequencing data can be accessed on SRA under the accession PRJNA877753 (duodenum), PRJNA877756 (jejunum), PRJNA877757 (ileum), PRJNA877758 (cecum) and SUB-12035409PRJNA877759 (feces).

Declarations

Ethics approval and consent to participate

All experiments involving animals were conducted according to the ethical policies and procedures approved by the Institutional Animal Care and Use Committee of China Agricultural University, China (Issue No. 32303202–1-1).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]State Key Laboratory of Animal Biotech Breeding and Frontier Science Center of Molecular Design Breeding, China Agricultural University, Beijing 100193, China. [2]National Engineering Laboratory for Animal Breeding and Key Laboratory of Animal Genetics, Breeding and Reproduction, Ministry of Agriculture and Rural Affairs, China Agricultural University, Beijing 100193, China. [3]Department of Animal Genetics and Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China. [4]Beijing Engineering Research Centre of Layer, Beijing 101206, China.

Received: 22 February 2024 Accepted: 1 February 2025
Published online: 01 March 2025

References

1. Phillips ML. Gut reaction environmental effects on the human microbiota. Environ Health Persp. 2009;117(5):A198–205. https://doi.org/10.1289/ehp.117-a198.
2. Peters A, Nawrot TS, Baccarelli AA. Hallmarks of environmental insults. Cell. 2021;184(6):1455–68. https://doi.org/10.1016/j.cell.2021.01.043.
3. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature. 2014;505(7484):559–63. https://doi.org/10.1038/nature12820.
4. R B Singh, A K Gupta, J Fedacko, L R Juneja, P Jarcuska, D Pella. Effects of Diet and Nutrients on Epigenetic and Genetic Expressions. Role of Functional Food Security in Global Health. 2019;681–707. https://doi.org/10.1016/B978-0-12-813148-0.00040-2.
5. Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, et al. Moving pictures of the human microbiome. Genome Biol. 2011;12(5):R50. https://doi.org/10.1186/gb-2011-12-5-r50.
6. Yamamoto R, Chung R, Vazquez JM, Sheng HJ, Steinberg PL, Ioannidis NM, et al. Tissue-specific impacts of aging and genetics on gene expression patterns in humans. Nat Commun. 2022;13(1):5803. https://doi.org/10.1038/s41467-022-33509-0.
7. Oliva M, Muñoz-Aguirre M, Kim-Hellmuth S, Wucher V, Gewirtz ADH, Cotter DJ, et al. The impact of sex on gene expression across human tissues. Science. 2020;369(6509):3066. https://doi.org/10.1126/science.aba3066.
8. Hokanson KC, Hernández C, Deitzler GE, Gaston JE, David MM. Sex shapes gut–microbiota–brain communication and disease. Trends Microbiol. 2023. https://doi.org/10.1016/j.tim.2023.08.013.
9. Spor A, Koren O, Ley R. Unravelling the effects of the environment and host genotype on the gut microbiome. Nat Rev Microbiol. 2011;9(4):279–90. https://doi.org/10.1038/nrmicro2540.
10. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. Plos Genet. 2007;3(9):1724–35. https://doi.org/10.1371/journal.pgen.0030161.
11. Xiao LL, Liu SY, Wu YL, Huang YQ, Tao SW, Liu YJ, et al. The interactions between host genome and gut microbiome increase the risk of psychiatric disorders: Mendelian randomization and biological annotation. Brain Behav Immun. 2023;113:389–400. https://doi.org/10.1016/j.bbi.2023.08.003.
12. Mallott EK, Amato KR. Host specificity of the gut microbiome. Nat Rev Microbiol. 2021;19(10):639–53. https://doi.org/10.1038/s41579-021-00562-3.
13. Asnicar F, Berry SE, Valdes AM, Nguyen LH, Piccinno G, Drew DA, et al. Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. Nat Med. 2021;27(2):321–32. https://doi.org/10.1038/s41591-020-01183-8.
14. A Visconti, C I Le Roy, F Rosa, N Rossi, T C Martin, R P Mohney, et al. Interplay between the human gut microbiome and host metabolism. Nat Commun. 2019;10 https://doi.org/10.1038/s41467-019-12476-z.
15. Benson AK, Kelly SA, Legge R, Ma FR, Low SJ, Kim J, et al. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. P Natl Acad Sci USA. 2010;107(44):18933–8. https://doi.org/10.1073/pnas.1007028107.
16. Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, et al. Genetic determinants of the gut microbiome in UK Twins. Cell Host Microbe. 2016;19(5):731–43. https://doi.org/10.1016/j.chom.2016.04.017.
17. Lopera-Maya EA, Kurilshikov A, van der Graaf A, Hu SX, Andreu-Sánchez S, Chen LM, et al. Effect of host genetics on the gut microbiome in 7,738 participants of the dutch microbiome project. Nat Genet. 2022;54(9):143–51. https://doi.org/10.1038/s41588-022-01164-2.
18. Turpin W, Espin-Garcia O, Xu W, Silverberg MS, Kevans D, Smith MI, et al. Association of host genome with intestinal microbial composition in a large healthy cohort. Nat Genet. 2016;48(11):1413–7. https://doi.org/10.1038/ng.3693.
19. Qin YW, Havulinna AS, Liu Y, Jousilahti P, Ritchie SC, Tokolyi A, et al. Combined effects of host genetics and diet on human gut microbiota and incident disease in a single population cohort. Nat Genet. 2022;54(2):134–42. https://doi.org/10.1038/s41588-021-00991-z.
20. Kurilshikov A, Medina-Gomez C, Bacigalupe R, Radjabzadeh D, Wang J, Demirkan A, et al. Large-scale association analyses identify host factors influencing human gut microbiome composition. Nat Genet. 2021;53(2):156–65. https://doi.org/10.1038/s41588-020-00763-1.
21. Rühlemann MC, Hermes BM, Bang C, Doms S, Moitinho-Silva L, Thingholm LB, et al. Genome-wide association study in 8,956 German individuals identifies influence of ABO histo-blood groups on gut microbiome. Nat Genet. 2021;53(2):147–55. https://doi.org/10.1038/s41588-020-00747-1.
22. G Difford, J Lassen, P Løvendahl. Genes and microbes, the next step in dairy cattle breeding, EAAP Scientific Committee, Ed.; ed., Wageningen Academic Publishers, Wageningen, The Netherlands, 2016.
23. Yang H, Wu JY, Huang XC, Zhou YY, Zhang YF, Liu M, et al. ABO genotype alters the gut microbiota by regulating GalNAc levels in pigs. Nature. 2022;606(7913):358–67. https://doi.org/10.1038/s41586-022-04769-z.
24. Introduction to Quantitative Genetics - Falconer,Ds. Population. 1962;17(1) 152–153.
25. Difford GF, Plichta DR, Lovendahl P, Lassen J, Noel SJ, Hojberg O, et al. Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. Plos Genet. 2018;14(10):e1007580. https://doi.org/10.1371/journal.pgen.1007580.
26. Camarinha-Silva A, Maushammer M, Wellmann R, Vital M, Preuss S, Bennewitz J. Host genome influence on gut microbial composition and microbial prediction of complex traits in pigs. Genetics. 2017;206(3):1637–44. https://doi.org/10.1534/genetics.117.200782.
27. S Tang, Y Xin, YL Ma, XW Xu, SH Zhao, JH Cao. Screening of microbes associated with swine growth and fat deposition traits across the intestinal tract. Front Microbiol. 2020;11 https://doi.org/10.3389/fmicb.2020.586776.
28. Khanal P, Maltecca C, Schwab C, Fix J, Tiezzi F. Microbiability of meat quality and carcass composition traits in swine. J Anim Breed Genet. 2021;138(2):223–36. https://doi.org/10.1111/jbg.12504.
29. Wen CL, Yan W, Sun CJ, Ji CL, Zhou QQ, Zhang DX, et al. The gut microbiota is largely independent of host genetics in regulating fat deposition in chickens. Isme J. 2019;13(6):1422–36. https://doi.org/10.1038/s41396-019-0367-2.
30. CL Wen, W Yan, CN Mai, ZY Duan, JX Zheng, CJ Sun, et al. Joint contributions of the gut microbiota and host genetics to feed efficiency in chickens. Microbiome. 2021;9(1) https://doi.org/10.1186/s40168-021-01040-x.
31. A C Nica, S B Montgomery, A S Dimas, B E Stranger, C Beazley, I Barroso, et al. Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. Plos Genet. 2010;6(4) https://doi.org/10.1371/journal.pgen.1000895.

32. Torres JM, Gamazon ER, Parra EJ, Below JE, Valladares-Salgado A, Wacher N, et al. Cross-tissue and tissue-specific eQTLs: partitioning the heritability of a complex trait. Am J Hum Genet. 2014;95(5):521–34. https://doi.org/10.1016/j.ajhg.2014.10.001.
33. Li YI, van de Geijn B, Raj A, Knowles DA, Petti AA, Golan D, et al. RNA splicing is a primary link between genetic variation and disease. Science. 2016;352(6285):600–4. https://doi.org/10.1126/science.aad9417.
34. Aguet F, Barbeira AN, Bonazzola R, Brown A, Castel SE, Jo B, et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020;369(6509):1318–30. https://doi.org/10.1126/science.aaz1776.
35. Liu SL, Gao YH, Canela-Xandri O, Wang S, Yu Y, Cai WT, et al. A multi-tissue atlas of regulatory variants in cattle. Nat Genet. 2022;54(9):1438–47. https://doi.org/10.1038/s41588-022-01153-5.
36. Teng J, Gao Y, Yin H, Bai Z, Liu S, Zeng H, et al. A compendium of genetic regulatory effects across pig tissues. Nat Genet. 2024;56:112–23. https://doi.org/10.1038/s41588-023-01585-7.
37. G Dailu, B Zhonghao, Z Xiaoning, Z Conghao, H Yali, G C The Chicken, et al. The ChickenGTEx pilot analysis: a reference of regulatory variants across 28 chicken tissues. bioRxiv. 2023; 2023.06.27.546670. https://doi.org/10.1101/2023.06.27.546670.
38. G Consortium. Genetic effects on gene expression across human tissues. Nature. 2018;553(7689):530–530. https://doi.org/10.1038/nature25160.
39. Vosa U, Claringbould A, Westra HJ, Bonder MJ, Deelen P, Zeng B, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet. 2021;53(9):1300–10. https://doi.org/10.1038/s41588-021-00913-z.
40. D Dutta, Y He, A Saha, M Arvanitis, A Battle, N Chatterjee. Aggregative trans-eQTL analysis detects trait-specific target gene sets in whole blood. Nat Commun. 2022;13(1) https://doi.org/10.1038/s41467-022-31845-9.
41. QQ Zhou, FR Lan, S Gu, GQ Li, GQ Wu, YY Yan, et al. Genetic and microbiome analysis of feed efficiency in laying hens. Poultry Sci. 2023;102(4) https://doi.org/10.1016/j.psj.2022.102393.
42. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60. https://doi.org/10.1093/bioinformatics/btp324.
43. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9. https://doi.org/10.1093/bioinformatics/btp352.
44. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–303. https://doi.org/10.1101/gr.107524.110.
45. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75. https://doi.org/10.1086/519795.
46. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81(5):1084–97. https://doi.org/10.1086/521987.
47. Chen SF, Zhou YQ, Chen YR, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):884–90. https://doi.org/10.1093/bioinformatics/bty560.
48. Kim D, Landmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357–60. https://doi.org/10.1038/Nmeth.3317.
49. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30. https://doi.org/10.1093/bioinformatics/btt656.
50. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290–5. https://doi.org/10.1038/nbt.3122.
51. MD Robinson, A Oshlack. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11(3) https://doi.org/10.1186/gb-2010-11-3-r25.
52. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq illumina sequencing platform. Appl Environ Microb. 2013;79(17):5112–20. https://doi.org/10.1128/Aem.01043-13.
53. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(9):852–7. https://doi.org/10.1038/s41587-019-0252-6.
54. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581–3. https://doi.org/10.1038/Nmeth.3869.
55. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. Isme J. 2017;11(12):2639–43. https://doi.org/10.1038/ismej.2017.119.
56. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41(D1):D590–6. https://doi.org/10.1093/nar/gks1219.
57. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci. 2003;14(6):927–30. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x.
58. Yang JA, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88(1):76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.
59. PM Visscher, G Hemani, AAE Vinkhuyzen, GB Chen, SH Lee, NR Wray, et al. Statistical power to detect genetic (Co)variance of complex traits using SNP data in unrelated samples. Plos Genet. 2014;10(4) https://doi.org/10.1371/journal.pgen.1004269.
60. Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, et al. Environment dominates over host genetics in shaping human gut microbiota. Nature. 2018;555(7695):210–5. https://doi.org/10.1038/nature25973.
61. Zierer J, Jackson MA, Kastenmüller G, Mangino M, Long T, Telenti A, et al. The fecal metabolome as a functional readout of the gut microbiome. Nat Genet. 2018;50(6):790–5. https://doi.org/10.1038/s41588-018-0135-7.
62. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012;44(7):821–4. https://doi.org/10.1038/ng.2310.
63. Gao XY, Stamier J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol. 2008;32(4):361–9. https://doi.org/10.1002/gepi.20310.
64. Jiang LD, Zheng ZL, Qi T, Kemper KE, Wray NR, Visscher PM, et al. A resource-efficient tool for mixed model association analysis of large-scale data. Nat Genet. 2019;51(12):1749–55. https://doi.org/10.1038/s41588-019-0530-8.
65. Barbeira AN, Bonazzola R, Gamazon ER, Liang YY, Park Y, Kim-Hellmuth S, et al. Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. Genome Biol. 2021;22(1):49. https://doi.org/10.1186/s13059-020-02252-4.
66. Zhu ZH, Zhang FT, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481–7. https://doi.org/10.1038/ng.3538.
67. Hurley JR, Cattell RB. The procrustes program - producing direct rotation to test a hypothesized factor structure. Behav Sci. 1962;7(2):258–62. https://doi.org/10.1002/bs.3830070216.
68. Priya S, Burns MB, Ward T, Mars RAT, Adamowicz B, Lock EF, et al. Identification of shared and disease-specific host gene-microbiome associations across human diseases using multi-omic integration. Nat Microbiol. 2022;7(6):780–95. https://doi.org/10.1038/s41564-022-01121-z.
69. MB Burns, J Lynch, TK Starr, D Knights, R Blekhman. Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment. Genome Med. 2015;7 https://doi.org/10.1186/s13073-015-0177-8.
70. Zheng DP, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. Cell Res. 2020;30(6):492–506. https://doi.org/10.1038/s41422-020-0332-7.

71. FZ Xu, YQ Fu, TY Sun, ZL Jiang, ZL Miao, ML Shuai, et al. The interplay between host genetics and the gut microbiome reveals common and distinct microbiome features for complex human diseases. Microbiome. 2020;8(1) https://doi.org/10.1186/s40168-020-00923-9.
72. Cuevas-Sierra A, Ramos-Lopez O, Riezu-Boj JI, Milagro FI, Martinez JA. Diet, gut microbiota, and obesity: links with host genetics and epigenetics and potential applications. Adv Nutr. 2019;10:S17–30. https://doi.org/10.1093/advances/nmy078.
73. Morais LH, Schreiber HLT, Mazmanian SK. The gut microbiota-brain axis in behaviour and brain disorders. Nat Rev Microbiol. 2021;19(4):241–55. https://doi.org/10.1038/s41579-020-00460-0.
74. Sgritta M, Dooling SW, Buffington SA, Momin EN, Francis MB, Britton RA, et al. Mechanisms underlying microbial-mediated changes in social behavior in mouse models of autism spectrum disorder. Neuron. 2019;101(2):246-259 e6. https://doi.org/10.1016/j.neuron.2018.11.018.
75. Buffington SA, Di Prisco GV, Auchtung TA, Ajami NJ, Petrosino JF, Costa-Mattioli M. Microbial reconstitution reverses maternal diet-induced social and synaptic deficits in offspring. Cell. 2016;165(7):1762–75. https://doi.org/10.1016/j.cell.2016.06.001.
76. Grice EA, Segre JA. The human microbiome: our second genome. Annu Rev Genom Hum G. 2012;13:151–70. https://doi.org/10.1146/annurev-genom-090711-163814.
77. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006;444(7122):1027–31. https://doi.org/10.1038/nature05414.
78. Tremaroli V, Bäckhed F. Functional interactions between the gut microbiota and host metabolism. Nature. 2012;489(7415):242–9. https://doi.org/10.1038/nature11552.
79. Rubino F, Carberry C, Waters SM, Kenny D, McCabe MS, Creevey CJ. Divergent functional isoforms drive niche specialisation for nutrient acquisition and use in rumen microbiome. Isme J. 2017;11(6):932–44. https://doi.org/10.1038/ismej.2017.34.
80. Dodd D, Spitzer MH, Van Treuren W, Merrill BD, Hryckowian AJ, Higginbottom SK, et al. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. Nature. 2017;551(7682):648–52. https://doi.org/10.1038/nature24661.
81. Zhao LP, Zhang F, Ding XY, Wu GJ, Lam YY, Wang XJ, et al. Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. Science. 2018;359(6380):1151–6. https://doi.org/10.1126/science.aao5774.
82. Fanning S, Hall LJ, Cronin M, Zomer A, MacSharry J, Goulding D, et al. Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. P Natl Acad Sci USA. 2012;109(6):2108–13. https://doi.org/10.1073/pnas.1115621109.
83. Stanley D, Hughes RJ, Moore RJ. Microbiota of the chicken gastrointestinal tract: influence on health, productivity and disease. Appl Microbiol Biot. 2014;98(10):4301–10. https://doi.org/10.1007/s00253-014-5646-2.
84. Young KT, Davis LM, DiRita VJ. Campylobacter jejuni: molecular biology and pathogenesis. Nat Rev Microbiol. 2007;5(9):665–79. https://doi.org/10.1038/nrmicro1718.
85. Kolodziejczyk AA, Zheng DP, Elinav E. Diet-microbiota interactions and personalized nutrition. Nat Rev Microbiol. 2019;17(12):742–53. https://doi.org/10.1038/s41579-019-0256-8.
86. Ecklu-Mensah G, Gilbert J, Devkota S. Dietary selection pressures and their impact on the gut microbiome. Cell Mol Gastroenter. 2022;13(1):7–18. https://doi.org/10.1016/j.jcmgh.2021.07.009.
87. Barreto HC, Gordo I. Intrahost evolution of the gut microbiota. Nat Rev Microbiol. 2023;21(9):590–603. https://doi.org/10.1038/s41579-023-00890-6.
88. Barroso-Batista J, Pedro MF, Sales-Dias J, Pinto CJG, Thompson JA, Pereira H, et al. Specific eco-evolutionary contexts in the mouse gut reveal escherichia coli metabolic versatility. Curr Biol. 2020;30(6):1049-+. https://doi.org/10.1016/j.cub.2020.01.050.
89. LP Henry, M Bruijning, SKG Forsberg, JF Ayroles. The microbiome extends host evolutionary potential. Nat Commun. 2021;12(1) https://doi.org/10.1038/s41467-021-25315-x.
90. Elena SF, Lenski RE. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat Rev Genet. 2003;4(6):457–69. https://doi.org/10.1038/nrg1088.
91. Petersen C, Hamerich IK, Adair KL, Griem-Krey H, Oliva MT, Hoeppner MP, et al. Host and microbiome jointly contribute to environmental adaptation. Isme J. 2023;17(11):1953–65. https://doi.org/10.1038/s41396-023-01507-9.
92. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, et al. Human genetics shape the gut microbiome. Cell. 2014;159(4):789–99. https://doi.org/10.1016/j.cell.2014.09.053.
93. M Beaumont, JK Goodrich, MA Jackson, I Yet, ER Davenport, S Vieira-Silva, et al. Heritable components of the human fecal microbiome are associated with visceral fat. Genome Biol. 2016;17 https://doi.org/10.1186/s13059-016-1052-7.
94. Liu C, Du MX, Xie LS, Wang WZ, Chen BS, Yun CY, et al. Gut commensal Christensenella minuta modulates host metabolism via acylated secondary bile acids. Nat Microbiol. 2024;9(2):434–50. https://doi.org/10.1038/s41564-023-01570-0.
95. Li X, Li Z, He Y, Li P, Zhou H, Zeng N. Regional distribution of Christensenellaceae and its associations with metabolic syndrome based on a population-level analysis. PeerJ. 2020;8:e9591. https://doi.org/10.7717/peerj.9591.
96. AJ Obregon-Tito, RY Tito, J Metcalf, K Sankaranarayanan, JC Clemente, LK Ursell, et al. Subsistence strategies in traditional societies distinguish gut microbiomes. Nat Commun. 2015;6(1) https://doi.org/10.1038/ncomms7505.
97. Org E, Blum Y, Kasela S, Mehrabian M, Kuusisto J, Kangas AJ, et al. Relationships between gut microbiota, plasma metabolites, and metabolic syndrome traits in the METSIM cohort. Genome Biol. 2017;18(1):70. https://doi.org/10.1186/s13059-017-1194-2.
98. Lim MY, You HJ, Yoon HS, Kwon B, Lee JY, Lee S, et al. The effect of heritability and host genetics on the gut microbiota and metabolic syndrome. Gut. 2017;66(6):1031–8. https://doi.org/10.1136/gutjnl-2015-311326.
99. Ayeni FA, Biagi E, Rampelli S, Fiori J, Soverini M, Audu HJ, et al. Infant and adult gut microbiome and metabolome in rural bassa and urban settlers from Nigeria. Cell Rep. 2018;23(10):3056–67. https://doi.org/10.1016/j.celrep.2018.05.018.
100. HL Barrett, LF Gomez-Arango, SA Wilkinson, HD McIntyre, LK Callaway, M Morrison, et al. A vegetarian diet is a major determinant of gut microbiota composition in early pregnancy. Nutrients. 2018;10(7) https://doi.org/10.3390/nu10070890.
101. XW Sun, HJ Huang, XM Wang, RQ Wei, HY Niu, HY Chen, et al. Christensenella strain resources, genomic/metabolomic profiling, and association with host at species level. Gut Microbes. 2024;16(1) https://doi.org/10.1080/19490976.2024.2347725.
102. Waters JL, Ley RE. The human gut bacteria Christensenellaceae are widespread, heritable, and associated with health. BMC Biol. 2019;17(1):83. https://doi.org/10.1186/s12915-019-0699-4.
103. Fu J, Bonder MJ, Cenit MC, Tigchelaar EF, Maatman A, Dekens JA, et al. The gut microbiome contributes to a substantial proportion of the variation in blood lipids. Circ Res. 2015;117(9):817–24. https://doi.org/10.1161/CIRCRESAHA.115.306807.
104. Hansen EE, Lozupone CA, Rey FE, Wu M, Guruge JL, Narra A, et al. Pan-genome of the dominant human gut-associated archaeon, Methanobrevibacter smithii, studied in twins. Proc Natl Acad Sci U S A. 2011;108 Suppl(Suppl 1):4599–606. https://doi.org/10.1073/pnas.1000071108.
105. Upadhyaya B, McCormack L, Fardin-Kia AR, Juenemann R, Nichenametla S, Clapper J, et al. Impact of dietary resistant starch type 4 on human gut microbiota and immunometabolic functions. Sci Rep. 2016;6:28797. https://doi.org/10.1038/srep28797.
106. Kim H, Lichtenstein AH, Ganz P, Miller ER 3rd, Coresh J, Appel LJ, et al. Associations of circulating proteins with lipoprotein profiles: proteomic analyses from the OmniHeart randomized trial and the Atherosclerosis Risk in Communities (ARIC) Study. Clin Proteomics. 2023;20(1):27. https://doi.org/10.1186/s12014-023-09416-x.
107. GT Uhr, L Dohnalová, CA Thaiss. The dimension of time in host-microbiome interactions. Msystems. 2019;4(1) https://doi.org/10.1128/mSystems.00216-18.
108. O Kolodny, H Schulenburg. Microbiome-mediated plasticity directs host evolution along several distinct time scales. Philos T R Soc B. 2020;375(1808) https://doi.org/10.1098/rstb.2019.0589.
109. V Woo, T Alenghat. Epigenetic regulation by gut microbiota. Gut Microbes. 2022;14(1) https://doi.org/10.1080/19490976.2021.2022407.

110. Lavelle A, Sokol H. Gut microbiota-derived metabolites as key actors in inflammatory bowel disease. Nat Rev Gastro Hepat. 2020;17(4):223–37. https://doi.org/10.1038/s41575-019-0258-z.
111. T Shock, L Badang, B Ferguson, K Martinez-Guryn. The interplay between diet, gut microbes, and host epigenetics in health and disease. J Nutr Biochem. 2021;95 https://doi.org/10.1016/j.jnutbio.2021.108631.
112. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet. 2016;17(8):487–500. https://doi.org/10.1038/nrg.2016.59.
113. Parker A, Lawson MAE, Vaux L, Pin C. Host-microbe interaction in the gastrointestinal tract. Environ Microbiol. 2018;20(7):2337–53. https://doi.org/10.1111/1462-2920.13926.
114. Levy M, Thaiss CA, Elinav E. Metabolites: messengers between the microbiota and the immune system. Gene Dev. 2016;30(14):1589–97. https://doi.org/10.1101/gad.284091.116.
115. AJ Bilotta, CY Ma, XS Huang, WJ Yang, SX Yao, YZ Cong. Microbiota metabolites SCFA stimulate epithelial migration to promote wound healing through MFGE8 and PAK1. J Immunol. 2019;202(1) https://doi.org/10.4049/jimmunol.202.Supp.191.15.
116. Lee K, Jayaraman A. Interactions between gut microbiota and non-alcoholic liver disease: the role of microbiota-derived metabolites. Pharmacol Res. 2019;142:521–9. https://doi.org/10.1016/j.phrs.2019.02.013.
117. Swann JR, Want EJ, Geier FM, Spagou K, Wilson ID, Sidaway JE, et al. Systemic gut microbial modulation of bile acid metabolism in host tissue compartments. P Natl Acad Sci USA. 2011;108:4523–30. https://doi.org/10.1073/pnas.1006734107.
118. R Blekhman, JK Goodrich, K Huang, Q Sun, R Bukowski, JT Bell, et al. Host genetic variation impacts microbiome composition across human body sites. Genome Biol. 2015;16 https://doi.org/10.1186/s13059-015-0759-1.
119. de Vos WM, Tilg H, Van Hul M, Cani PD. Gut microbiome and health: mechanistic insights. Gut. 2022;71(5):1020–32. https://doi.org/10.1136/gutjnl-2021-326789.
120. Jones RM. The influence of the gut microbiota on host physiology: in pursuit of mechanisms. Yale J Biol Med. 2016;89(3):285–97.
121. Goodrich JK, Davenport ER, Waters JL, Clark AG, Ley RE. Cross-species comparisons of host genetic associations with the microbiome. Science. 2016;352(6285):532–5. https://doi.org/10.1126/science.aad9379.
122. Park W. Gut microbiomes and their metabolites shape human and animal health. J Microbiol. 2018;56(3):151–3. https://doi.org/10.1007/s12275-018-0577-8.
123. Trinh P, Zaneveld JR, Safranek S, Rabinowitz PM. One health relationships between human, animal, and environmental microbiomes: a mini-review. Front Public Health. 2018;6:235. https://doi.org/10.3389/fpubh.2018.00235.
124. Liu XY, Li YI, Pritchard JK. Trans effects on gene expression can drive omnigenic inheritance. Cell. 2019;177(4):1022–34. https://doi.org/10.1016/j.cell.2019.04.014.
125. Zhou X, Cai XD. Joint eQTL mapping and inference of gene regulatory network improves power of detecting both cis- and trans-eQTLs. Bioinformatics. 2022;38(1):149–56. https://doi.org/10.1093/bioinformatics/btab609.
126. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, et al. Genetic analysis of genome-wide variation in human gene expression. Nature. 2004;430(7001):743–7. https://doi.org/10.1038/nature02797.
127. MacLellan WR, Wang Y, Lusis AJ. Systems-based approaches to cardiovascular disease. Nat Rev Cardiol. 2012;9(3):172–84. https://doi.org/10.1038/nrcardio.2011.208.
128. Price AL, Helgason A, Thorleifsson G, McCarroll SA, Kong A, Stefansson K. Single-tissue and cross-tissue heritability of gene expression via identity-by-descent in related or unrelated individuals. Plos Genet. 2011;7(2):e1001317. https://doi.org/10.1371/journal.pgen.1001317.
129. Wright FA, Sullivan PF, Brooks AI, Zou F, Sun W, Xia K, et al. Heritability and genomics of gene expression in peripheral blood. Nat Genet. 2014;46(5):430–7. https://doi.org/10.1038/ng.2951.
130. Grundberg E, Small KS, Hedman ÅK, Nica AC, Buil A, Keildson S, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat Genet. 2012;44(10):1084–9. https://doi.org/10.1038/ng.2394.
131. Ouwens KG, Jansen R, Nivard MG, van Dongen J, Frieser MJ, Hottenga JJ, et al. A characterization of cis- and trans-heritability of RNA-Seq-based gene expression. Eur J Hum Genet. 2020;28(2):253–63. https://doi.org/10.1038/s41431-019-0511-5.
132. FW Albert, JS Bloom, J Siegel, L Day, L Kruglyak. Genetics of trans-regulatory variation in gene expression. Elife. 2018;7 https://doi.org/10.7554/eLife.35471.
133. Battle A, Mostafavi S, Zhu XW, Potash JB, Weissman MM, McCormick C, et al. Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. Genome Res. 2014;24(1):14–24. https://doi.org/10.1101/gr.155192.113.
134. Heinig M, Petretto E, Wallace C, Bottolo L, Rotival M, Lu H, et al. A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. Nature. 2010;467(7314):460–4. https://doi.org/10.1038/nature09386.
135. Small KS, Hedman ÅK, Grundberg E, Nica AC, Thorleifsson G, Kong A, et al. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes (vol 43, pg 561, 2011). Nat Genet. 2011;43(10):1040–1040. https://doi.org/10.1038/ng1011-1040c.
136. Brynedal B, Choi J, Raj T, Bjornson R, Stranger BE, Neale BM, et al. Large-scale trans-eQTLs affect hundreds of transcripts and mediate patterns of transcriptional co-regulation. Am J Hum Genet. 2017;100(4):581–91. https://doi.org/10.1016/j.ajhg.2017.02.004.
137. Y Yang, X Song, ZQ Xiong, YJ Xia, GQ Wang, LZ Ai. Complete genome sequence of Lactobacillus salivarius AR809, a probiotic strain with oropharyngeal tract resistance and adhesion to the oral epithelial cells. Curr Microbiol. 2022;79(9) https://doi.org/10.1007/s00284-022-02963-w.
138. Sanchez MG, Passot S, Campoy S, Olivares M, Fonseca F. Ligilactobacillus salivarius functionalities, applications, and manufacturing challenges. Appl Microbiol Biot. 2022;106(1):57–80. https://doi.org/10.1007/s00253-021-11694-0.
139. Wang Z, Zeng XM, Mo YM, Smith K, Guo YM, Lin J. Identification and characterization of a bile salt hydrolase from Lactobacillus salivarius for development of novel alternatives to antibiotic growth promoters. Appl Environ Microb. 2012;78(24):8795–802. https://doi.org/10.1128/Aem.02519-12.
140. Begley M, Hill C, Gahan CGM. Bile salt hydrolase activity in probiotics. Appl Environ Microb. 2006;72(3):1729–38. https://doi.org/10.1128/Aem.72.3.1729-1738.2006.
141. Rimal B, Collins SL, Tanes CE, Rocha ER, Granda MA, Solanki S, et al. Bile salt hydrolase catalyses formation of amine-conjugated bile acids. Nature. 2024;626(8000):859–63. https://doi.org/10.1038/s41586-023-06990-w.
142. Ko CW, Qu J, Black DD, Tso P. Regulation of intestinal lipid metabolism: current concepts and relevance to disease. Nat Rev Gastroenterol Hepatol. 2020;17(3):169–83. https://doi.org/10.1038/s41575-019-0250-7.
143. K R Feingold. Introduction to Lipids and Lipoproteins, in: K.R. Feingold, B. Anawalt, M.R. Blackman, A. Boyce, G. Chrousos, E. Corpas, W.W. de Herder, K. Dhatariya, K. Dungan, J. Hofland, S. Kalra, G. Kaltsas, N. Kapoor, C. Koch, P. Kopp, M. Korbonits, C.S. Kovacs, W. Kuohung, B. Laferrere, M. Levy, E.A. McGee, R. McLachlan, M. New, J. Purnell, R. Sahay, A.S. Shah, F. Singer, M.A. Sperling, C.A. Stratakis, D.L. Trence, D.P. Wilson (Eds.), Endotext, South Dartmouth (MA), 2000.
144. Luo J, Yang H, Song BL. Mechanisms and regulation of cholesterol homeostasis. Nat Rev Mol Cell Biol. 2020;21(4):225–45. https://doi.org/10.1038/s41580-019-0190-7.
145. Zechner R, Kienesberger PC, Haemmerle G, Zimmermann R, Lass A. Adipose triglyceride lipase and the lipolytic catabolism of cellular fat stores. J Lipid Res. 2009;50(1):3–21. https://doi.org/10.1194/jlr.R800031-JLR200.
146. Dumesnil C, Vanharanta L, Prasanna X, Omrane M, Carpentier M, Bhapkar A, et al. Cholesterol esters form supercooled lipid droplets whose nucleation is facilitated by triacylglycerols. Nat Commun. 2023;14(1):915. https://doi.org/10.1038/s41467-023-36375-6.
147. LH Zhu, RR Liao, JW Huang, CF Xiao, YZ Yang, HY Wang, et al. Lactobacillus salivarius SNK-6 regulates liver lipid metabolism partly via the miR-130a-5p/MBOAT2 Pathway in a NAFLD Model of Laying Hens. Cells-Basel. 2022;11(24) https://doi.org/10.3390/cells11244133.
148. X Zhao, XQ Zhong, X Liu, XY Wang, XM Gao. Therapeutic and improving function of Lactobacilli in the prevention and treatment of cardiovascular-related diseases: a novel perspective from gut microbiota. Front Nutr. 2021;8 https://doi.org/10.3389/fnut.2021.693412.

149. Larsen N, Vogensen FK, Gobel RJ, Michaelsen KF, Forssten SD, Lahtinen SJ, et al. Effect of Lactobacillus salivarius Ls-33 on fecal microbiota in obese adolescents. Clin Nutr. 2013;32(6):935–40. https://doi.org/10.1016/j.clnu.2013.02.007.
150. Evers MR, Xia GQ, Kang HG, Schachner M, Baenziger JU. Molecular cloning and characterization of a dermatan-specific N-acetylgalactosamine 4-O-sulfotransferase. J Biol Chem. 2001;276(39):36344–53. https://doi.org/10.1074/jbc.M105848200.
151. Begolli G, Markovic I, Knezevic J, Debeljak Z. Carbohydrate sulfotransferases: a review of emerging diagnostic and prognostic applications. Biochem Med (Zagreb). 2023;33(3):030503. https://doi.org/10.11613/BM.2023.030503.

Publisher's Note