Divergent host genetic architectures drive breed-specific modulation of the caecal microbiome in chickens

Xiaoxia Dai
Royal Veterinary College
Ankit Hinsu
Royal Veterinary College
Christos Dadousis
University of Surrey
Melanie Hay
Royal Veterinary College
Bruno Fosso
University of Bari “Aldo Moro”
Matteo Crotta
Royal Veterinary College
Ramesh Pandit
Government of Gujarat
Javier Guitian
Royal Veterinary College
Fiona Tomley
Royal Veterinary College
Prakash Koringa
Kamdhenu University
Chaitanya Joshi
Government of Gujarat
Damer P. Blake
Royal Veterinary College
Androniki Psifidi
apsifidi@rvc.ac.uk
Royal Veterinary College

Research Article

Keywords: host genetics, caecal microbiome, host-microbiome interaction, GWAS, whole genome sequencing, indigenous chickens, broilers

Posted Date: February 6th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-5953792/v1

Additional Declarations: No competing interests reported.

Abstract

Background:

Enteric microbiota are crucial for animal health and performance, but their role is often challenging to determine. Host genetics appear to influence gut microbial communities, even among individuals in similar environments. This study aimed to identify host genetic variation associated with the caecal microbiota in two chicken breeds: the indigenous Indian Kadaknath and commercial Cobb400 broilers. Blood and caecal contents were collected from 300 chickens per breed in Western India. Genotyping was performed using the 600K Affymetrix Axiom HD single nucleotide polymorphism (SNP) array, and caecal microbiota were characterized using 16S rRNA gene sequencing. A subset of 40 chicken underwent whole genome sequencing for deeper genetic insights. SNP-based heritability estimates and genome-wide association studies (GWAS) were conducted separately for both breeds to explore the genetic background of caecal microbial structure (alpha and beta diversity) and the host's influence on the establishment of different bacterial genera in the caeca, with a focus on potential pathogens.

Results:

The GWAS in Kadaknath chickens identified 108 significant SNPs across 26 *Gallus gallus* chromosomes (GGC1–GGC23, GGC27, and GGC28), associated with heritable caecal microbial traits, including microbiota structure and the abundance of specific genera. SNP-based heritability estimates ranged from 0.12 to 0.76, with the nearest genes predominantly involved in immune response and cell signaling pathways. In Cobb400 chickens, 71 significant SNPs were identified across 22 chromosomes (GGC1–GGC5, GGC7–GGC11, GGC14, GGC17–GGC22, GGC25–GGC28, and GGCZ), linked to heritable caecal microbial traits with heritability estimates from 0.16 to 0.61. Genes nearest to these SNPs were primarily associated with microbial regulation, growth processes, and adaptation under stress conditions. These findings underscore the role of host genetic variation in shaping caecal microbiota composition and diversity in indigenous and commercial chicken breeds and reveal breed-specific genetic architectures underlying caecal microbiota-related traits.

Conclusion:

This study provides valuable insights into the genetic basis of host-microbiome interactions in chickens, highlighting distinct breed-specific genetic influences on caecal microbiota composition. These findings have the potential to inform future genomic selection strategies aimed at enhancing protective or productive gut microbial populations while reducing reliance on antibiotics in poultry production.

Background

Host-microbiota relationships have garnered attention in recent years due to their impact on animal productivity, animal health and welfare. Recent studies have indicated a notable host genetic component controlling microbiota composition, suggesting scope for the identification of genetic markers and application in selective animal breeding strategies to improve gut health [1, 2]. Many enteric microorganisms have established symbiotic relationships with their hosts and contribute to vital functions such as regulating host metabolic and immune pathways, or preventing pathogen colonization, emphasising their importance [3, 4]. The composition of the gut microbiota can be influenced by numerous extrinsic factors including diet and environmental conditions [5], medication [6], and host genetics [7, 8]. Understanding the mechanisms that underpin these interactions can be used to influence enteric microbiota to improve animal health and productivity. However, studying the microbiota at an industrial scale in farmed stock such as chickens is costly and laborious.

A deeper understanding of chicken-microbiota relationships and the host genes that can regulate the colonisation of beneficial microbes can help select individuals with better resistance to diseases, leading to better growth and reduced use of antimicrobials. Improving chicken health and productivity can secure a sustainable food source for the global human population [9]. Chickens are the world's most widely farmed terrestrial animals, their production and consumption having surpassed that of pigs, and they are a major component of human nutrition [10, 11]. Consideration of host-microbiota relationships in chickens has revealed inter- as well as intra-breed variation [8]. Chicken breeds have distinct physiological and anatomical characteristics that can impact the composition and function of their microbiota. For example, indigenous Indian breeds such as the Kadaknath have gut microbial communities that are distinct from other indigenous breeds such as the Aseel, or commercial broiler-type lines like the Cobb400 [8]. These differences may be influenced by diet and digestive physiology [12] but host genetics can also modulate gut microbial communities [13, 14]. Identification of genotypes that determine such within breed genetic variation can be used in selective breeding strategies to enhance enteric microbiota, benefitting chicken health and welfare in balance with regional needs and climates.

In this study, we investigated the genetic factors influencing caecal microbiota in two distinct chicken breeds from Western India, namely the widely used commercial breed with global representation, Cobb400, and the indigenous traditional breed, Kadaknath. Cobb400, known for meat production, was juxtaposed with Kadaknath, recognized for its slow growth, traditional characteristics, high-quality meat, and disease resistance [15]. Blood and caecal contents were collected from 300 individuals of each breed. Genotyping utilized the 600K Affymetrix Axiom HD array, and caecal microbiota were characterized using 16S rRNA gene sequencing. Whole-genome sequencing (WGS) was specifically performed on 18 Cobb400 and 22 Kadaknath chickens sub-sampled from the main cohorts. The resulting single-nucleotide polymorphism (SNP) and microbiota datasets were employed in Genome-Wide Association Studies (GWAS) to systematically assess the role of host genetics in shaping caecal microbial communities. The WGS analysis of 40 subsamples provided finer resolution, aiding in the identification of potential candidate genes and variants and contributing to a more comprehensive understanding of the genetic landscape influencing caecal microbiota. This integrated approach significantly augmented the depth and precision of the study, offering a nuanced insight into the genetic

determinants of host effects on the caecal microbial community. The research offers insights to discern within breed-specific influences on host-microbiota interactions, complementing our recent crossbreed analyses (manuscript under review) and advancing our comprehension of the intricate relationship between genetics and gut health in poultry.

Results

Microbial diversity and caecal bacterial genera abundance in Kadaknath and Cobb400 chickens

The DADA2 pipeline was utilized to generate 9,217 ASVs from 16S rRNA amplicon sequencing. During the initial filtering process in phyloseq, 41 samples were removed due to low read counts (less than 10,000 per sample). The data were then rarefied to 10,000 reads, resulting in 6,326 ASVs from 556 samples for further analysis, including 292 Cobb400 and 264 Kadaknath chicken. The α - diversity (Shannon index) varied from 0.78 to 5.37 (Supplementary Fig. 2). No statistically significant differences were observed in α - diversity between Kadaknath and Cobb400 chickens (Fig. 1A). β - diversity analysis revealed that the first 10 PCoA axes accounted for 38.1% and 45.1% of the total microbial variation for Cobb400 and Kadaknath chickens, respectively (Supplementary Fig. 3), indicating differences in microbial community structure and composition between the two chicken lines (Fig. 1B). ASVs were classified as Bacteria (n = 6,304) and Archaea (n = 9), and further classified into 24 phyla and 319 genera, with 13 ASVs remaining unassigned. The most abundant genera were *Bacteroides*, *Fecalibacterium*, *Alistipes, Clostridiales.vadinBB60.group_X*, and *Ruminococcaceae_X* for both breeds (Supplementary Fig. 4). After filtering out low-abundance genera, 35 genera were retained for GWAS analysis as numerical/continuous phenotypes, and 39 as binary phenotypes. These genera had a prevalence ranging from 0.3 to 0.97 and accounted for at least 93% of the total sequencing reads (Supplementary Table 2). Among these genera *Megamonas, Faecalibacterium, Helicobacter, Erysipelatoclostridium*, and *Oscillibacter* were more highly represented in Kadaknath chickens, while *Phascolarctobacterium, Lactobacillus, Cloacibacillus* and *Akkermansia* were more highly represented in Cobb400 chickens (Supplementary Table 3).

Metadata

Multiple Correspondence Analysis of the recorded metadata (farm practices categorical variables) revealed that the first four MCA components accounted for 75.3% and 71.2% of the overall farm practice variability in Cobb400 and Kadaknath, respectively (Supplementary Fig. 5). Therefore, the first four MCA components of farm practice categorical variables were used as covariates in the ensuing SNP-based heritability and GWAS analyses to account for sources of environmental noise.

GWAS

Heritability of caecal microbiota phenotypes. To estimate the contribution of host genetic variation to the caecal microbiota, SNP-based heritability (h^2) was calculated for each selected phenotype. In Cobb400,

heritability ranged from 0 to 0.73, with 33 out of 85 selected phenotypes exhibiting non-zero estimated h^2 (likelihood ratio test, $P < 0.05$; Fig. 2 and Supplementary Table 2). In Cobb400 chickens, β - diversity (PCoA axes 2, 3, 5, 6, 7, and 9) exhibited significant heritability (heritability ranged from 0.20 to 0.73; $P < 0.05$; Fig. 3), with PCoA axis 6 showing the highest h^2 ($h^2 = 0.73$, $P < 0.01$; Fig. 2; Supplementary Table 2). Moreover, the relative abundance of 26 out of 74 genera in the caeca microbiota were identified as heritable traits. Similarly, in Kadaknath, h^2 ranged from 0 to 0.77, with 57 out of 85 selected phenotypes showing non-zero estimated h^2 (likelihood ratio test, $P < 0.05$; Fig. 2 and Supplementary Table 2). Significant h^2 was observed for α - (Shannon index; $h^2 = 0.34$) and β - diversity (PCoA axes 2, 3, 4, 5, 6, 7, 9, and 10; h^2 ranged from 0.29 to 0.76; $P < 0.01$; Fig. 3), with the highest h^2 was estimated for β - diversity of PCoA axis 9 ($h^2 = 0.76$; $P < 0.01$; Fig. 2; Supplementary Table 2). The relative abundance of 48 out of 74 analyzed genera were identified as heritable traits.

GWAS of α - and β - diversity phenotypes. The Manhattan and QQ plots presenting the GWAS results for Cobb400 and Kadaknath chickens indicated that several SNPs across various chromosomes were significantly associated with α - and β - diversity, and the QQ plots indicated substantial deviations from the expected distribution (Fig. 3). Table 1 summarises the top SNPs associated with the heritable alpha and beta diversity of caecal microbiota in Cobb400 and Kadaknath.

Cobb400 broiler. In the Cobb400 broiler, significant genomic associations were observed specifically with β - diversity (PCoA axes 2, 3, and 4; Fig. 3; Supplementary Table 4; Supplementary Fig. 6). A total of 11 SNPs were significantly associated with caecal microbial β - diversity: 7 SNPs were associated with PCoA axis 2, 3 SNPs with PCoA axis 3, and 1 SNP with PCoA axis 4 (Fig. 3; Supplementary Table 4). Among these top SNPs (Table 1), SNP rs316331412 on *Gallus gallus Chromosome* Z (*GGC* Z) at position 58,109,849 bp exhibited the most significant association with PCoA axis 2 (*P*-value = 1.04×10^{-08}, effect size = -0.496, SE = 0.084), located within the *KIAA0825* gene, also known as *C5orf36*. Additional significant associations were found on GGC 2, including SNPs at positions 110,390,949 bp for PCoA axis 3 (*P*-value = 6.51×10^{-09}, effect size = 0.495, SE = 0.083), which is 990 kb downstream of the mitochondrial ribosomal protein L15 (*MRPL15*) gene, and position 8,757,131 bp for PCoA axis 4 (*P*-value = 1.17×10^{-07}, effect size = 0.734, SE = 0.135) within the DnaJ heat shock protein family (Hsp40) member B6 (*DNAJB6*) gene.

Kadaknath. For the Kadaknath chickens, the GWAS analysis identified 15 significant SNPs associated with seven diversity phenotypes (α-diversity (Shannon index), PCoA axes 2, 3, 4, 6, 9, and 10) of the caecal microbiota, located on multiple chromosomes (Fig. 3, Supplementary Table 5; Supplementary Fig. 7). The top SNP on *GGC* 25 at position 2,307,838 bp showed the strongest association with PCoA axis 10 (*P* - value = 4.49×10^{-09}, effect size = 0.596, SE = 0.098), located within the coatomer protein complex subunit alpha (*COPA*) gene (Table 1). Significant associations were also observed on GGC 27 at position 5,121,354 bp for PCoA axis 2 (*P* - value = 1.41×10^{-08}, effect size = 0.718, SE = 0.122), located upstream of the DDB1 and CUL4 associated factor 7 (*DCAF7*) gene, and on GGC 1 at position 167,269,928 bp for PCoA axis 3 (*P* - value = 3.36×10^{-08}, effect size = 0.674, SE = 0.118) near the *ENSGALT00000049953*

gene. Furthermore, a significant SNP (rs15679485) on GGC 13 at position 7,493,276 bp observed the strong association with PCoA axis 4 (*P*- value = 2.16×10^{-09}, effect size = -0.513, SE = 0.083) was located 36 kb upstream of the cyclin G1 (*CCNG1*) gene. Additionally, the top SNP on GGC 20 at position 5,969,268 bp showed significant associations with PCoA axis 6 (*P* - value = 1.60×10^{-08}, effect size = 0.982, SE = 0.168), located within the Sulfatase 2 (*SULF2*) gene. A significant association was also observed between the top SNP rs317996850 on GGC 28 at position 2,138,482 bp and PCoA axis 9 (*P* - value = 3.11×10^{-08}, effect size = 0.871, SE = 0.153), located within the DOT1 Like Histone Lysine Methyltransferase (*DOT1L*) gene.

GWAS of *microbial genera*

In Cobb400, from the identified 26 heritable caecal genera, we identified significant associations between SNP markers with their relative abundance only for 9 genera [*Campylobacter* (h^2 = 0.16), *Cloacibacillus* (h^2 = 0.39), *Desulfovibrio* (h^2 = 0.28), *Enterococcus* (h^2 = 0.61), *Erysipelatoclostridium* (h^2 = 0.19), *Fournierella* (h^2 = 0.33), *Rikenellaceae_RC9_gut_group* (h^2 = 0.32), *Ruminococcaceae_UCG_014* (h^2 = 0.26), and *Sutterella* (h^2 = 0.29)] (Fig. 4; Supplementary Table 6; Supplementary Fig. 8). These 9 genera collectively accounted for a relative abundance of 9.28%, with the highest relative abundance being 2.63% (*Ruminococcaceae.UCG.014;* Supplementary Table 2). From the 48 heritable caecal genera of Kadaknath, we identified significant associations between SNPs with their relative abundance only in 17 genera [*Alistipes* (h^2 = 0.51), *Bacteroides* (h^2 = 0.28), *Blautia* (h^2 = 0.17), *Cloacibacillus* (h^2 = 0.21), *Clostridiales.vadinBB60.group_X* (h^2 = 0.17), *Desulfovibrionaceae_X* (h^2 = 0.49), *Eisenbergiella* (h^2 = 0.17), *Erysipelatoclostridium* (h^2 = 0.36), *Erysipelotrichaceae_X* (h^2 = 0.27), *Helicobacter* (h^2 = 0.36), *Muribaculaceae_X* (h^2 = 0.12), *Parasutterella* (h^2 = 0.30), *Rikenella* (h^2 = 0.36), *Rikenellaceae_X* (h^2 = 0.45), *Ruminiclostridium* (h^2 = 0.35), *Ruminococcaceae.UCG.009* (h^2 = 0.26), and *Ruminococcus.1* (h^2 = 0.24)] (Fig. 5; Supplementary Table 7; Supplementary Fig. 9). These 17 genera collectively accounted for a relative abundance of 32.04%, with the highest relative abundance being 12.41% (*Bacteroides;* Supplementary Table 2).

Cobb400 Breed. For Cobb400, we discovered a total of 60 significant SNPs that were associated with 9 heritable caecal genera (Fig. 4; Supplementary Table 6; Supplementary Fig. 8). Notably, the highest number of these SNPs was found on GGC 1, with a total of 12 SNPs. Following closely, we observed 8 SNPs on GGC 10, 7 SNPs on GGC 3, and 5 SNPs on GGC 2. Most of these SNPs were primarily associated with the genera *Rikenellaceae_RC9_gut_group, Erysipelatoclostridium, Desulfovibrio* and *Campylobacter.*

Among these 60 SNPs, 40 were significantly associated with the genus *Rikenellaceae_RC9_gut_group*, located across various regions of the genome (Fig. 4 and Supplementary Table 6;). The top SNP rs313227007 located on GGC 4 at position 6,663,045 bp exhibited the most significant association with *Rikenellaceae_RC9_gut_group* (*P* - value = 3.95×10^{-10}, effect size = 0.258, SE = 0.004; Table 2), is located within the protocadherin 11 X-linked (*PCDH11X*) gene. Moreover, there was a significant positive

association between the genus *Erysipelatoclostridium* and 6 SNPs (Fig. 4 and Supplementary Table 6). Five of these SNPs were primarily located on GGC 10, specifically in the region spanning positions 16.78 to 16.82 Mb. Notably, this region is approximately 14 kb downstream of the arrestin domain containing 4 (*ARRDC4*) gene. The top SNPs on GGC 10 at position 16,787,954 bp had the strongest association with the genus *Erysipelatoclostridium* (*P* - value = 3.12×10^{-08}, effect size = 0.546, SE = 0.096; Table 2). Additionally, there were five SNPs significantly associated with genus *Desulfovibrio*, two of them on GGC 10 at position 18.34 Mb, located within immunoglobulin superfamily DCC subclass member 3 (*IGDCC3*), also known as Neogenin 1-like (*NEO1- like*). The SNP showing the strongest association (*P* - value = 2.10×10^{-08}, effect size = 0.251, SE = 0.043; Table 2) with the genus *Desulfovibrio* was located on GGC 1 at position 109,193,389 bp, located upstream of ETS proto-oncogene 2, transcription factor (*ETS*) gene. This top SNP also showed the strongest association with the genus *Cloacibacillus* (*P* - value = 1.05×10^{-08}, effect size = 0.33, SE = 0.056; Table 2). We also discovered three SNPs that were significantly and positively associated with *Campylobacter* (Supplementary Table 6). The top SNP on chromosome 27 at position 6,810,316 bp showed the strongest association with *Campylobacter* (*P* - value = 3.42×10^{-9}, effect size = 0.414, SE = 0.068; Table 2). This SNP is located within the *Mixed-Lineage Leukemia Translocated To 6* (*MLLT6*; PHD Finger Containing) gene. In addition, our GWAS analysis revealed a significant positive association between the genus *Ruminococcaceae_UCG_014* and a SNP (rs314860198) on GGC 22 at position 4,186,483 bp (*P* - value = 7.36×10^{-08}, effect size = 0.485, SE = 0.088; Table 2). Notably, this SNP is positioned approximately 14 kb upstream of the transforming growth factor alpha (*TGFA*) gene. In addition, top SNP rs312299864 on GGC 17 at position 6,835,139 bp had the strongest association (*P* - value = 5.44×10^{-08}, effect size = 0.365, SE = 0.065; Table 2) with the genus *Fournierella*, located within cilia and flagella associated protein 77 (*CFAP77*) gene. Top SNP rs315494328 on GGC 1 at position 12,446,509 bp was significantly associated with the genus *Sutterella (P* - value = 4.90×10^{-08}, effect size = 0.253, SE = 0.045; Table 2), located within membrane associated guanylate kinase, WW and PDZ domain containing 2 (*MAGI2*) gene.

Kadaknath breed. GWAS analysis for Kadaknath chickens identified a total of 93 SNPs significantly associated with 17 heritable caecal genera (Fig. 5; Supplementary Table 7; Supplementary Fig. 9). Interestingly, a substantial proportion of these significant SNPs were located on chromosome 1 (20 SNPs), followed by chromosome 2 (16 SNPs). These significant SNPs in Kadaknath were primarily associated with the genera *Bacteroides, Cloacibacillus, Clostridiales.vadinBB60.group_X* and *Helicobacter*.

A total of 25 SNPs showed significant associations with *Bacteroides*, and these SNPs were located across various chromosomes (GGC 1 to 6, GGC 8, GGC11, GGC 14 to 21, and GGC 23; Fig. 5; Supplementary Table 7). The top SNP rs314430640 on GGC 19 at position 954,411 bp had the strongest association with *Bacteroides* (*P* - value = 5.07×10^{-26}, effect size = 1.994, SE = 0.169; Table 2), located within the MAX network transcriptional repressor (*MNT*) gene. We identified 12 significant SNPs that showed associations with *Clostridiales.vadinBB60.group_X* in the caecal microbiome of the Kadaknath (Fig. 5; Supplementary Table 7). The top SNP on GGC 4 at position 1,249,188 bp was associated with

Clostridiales.vadinBB60.group_X (*P* - value = 5.74×10^{-10}, effect size = -1.105, SE = 0.172; Table 2) located within *RAB41* gene (member RAS oncogene family). There were 11 SNPs significantly associated with the genus *Cloacibacillus* (Fig. 5; Supplementary Table 7), with the majority located on GGC 1, including the top SNP rs316303702 at position 38,074,354 bp located upstream of nucleosome assembly protein 1 like 1 (*NAP1L1*) gene (*P* - value = 5.12×10^{-09}, effect size = -0.419, SE = 0.069; Table 2). Meanwhile, a total of 11 significant SNPs were positively associated with the genus *Helicobacter*, with the majority located on GGC 2 (Fig. 5; Supplementary Table 7). The top SNP on GGC 2 at position 116,059,085 bp observed the strongest association with the genus *Helicobacter* (*P* - value = 7.77×10^{-10}, effect size = 0.509, SE = 0.008; Table 2), located upstream of open reading frame (*C8orf34*) gene. Six SNPs significantly associated with the genus *Helicobacter* were on GGC 2 at position ~ 12.2 Mb, located upstream of RALY RNA binding protein-like (*RALYL*) gene. Meanwhile, 8 SNPs located on GGC 1 at position ~ 14 Mb were significantly associated with the genus *Rikenellaceae_X*, all located within the collagen type IV alpha 2 chain (*COL4A2*) gene (Fig. 5; Supplementary Table 7). There were 5 SNPs significantly associated with the genus *Parasutterella*, with 4 SNPs on GGC 9. The top SNP on chromosome 9 at position 13,695,119 bp, located within *ENSGALG00000052927* gene, had the strongest association (*P* - value = 4.59×10^{-12}, effect size = 0.482, SE = 0.066; Table 2) with the genus *Parasutterella.*

Whole genome sequencing

Variant calling using the GATK pipeline revealed the presence of approximately 20 million genetic variants, including SNPs and INDELs, in the chicken WGS data. A total of 222,596 variants spanning 430 protein-coding genes were analyzed from 154 genomic regions (65 from Cobb400 and 89 from Kadaknath). The variants were categorized into several types including intron_variant (47.28% in Cobb400; 57.68% in Kadaknath), intergenic_variant (43.16% in Cobb400; 32.17% in Kadaknath), and non_coding_transcript_variant (4.38% in Cobb400; 5.16% in Kadaknath; Supplementary Fig. 10). Based on the VEP annotations of the variants, we focused our interrogation on missense variants, start_lost variants, and splice_region variants with a predicted moderate to high impact on the encoded proteins (Supplementary Table 8). These variants had SIFT scores indicating deleterious effects, ranging from 0 to 0.05, across several genes in both Cobb400 and Kadaknath chickens.

According to the VEP analysis, we identified several missense variants with moderate predicted impact that overlapped with genes in close proximity to the SNPs significantly associated with phenotypes of interest in the GWAS analysis (Table 3). Specifically, in Cobb400 chickens, four genes—*KIAA0100, ITPR2, PSMG1*, and *MIA3*—were prioritized by VEP analysis. These genes overlapped with the loci closest to significant SNPs associated with traits of interest. Each of these genes harbored significant missense variants linked to traits related to the *Rikenellaceae_RC9_gut_group* caecal genus. The genomic locations of the variants were as follows: *KIAA0100* (GGC 19: 5906905–5908070), *ITPR2* (GGC 1: 67693634), *PSMG1* (GGC 1: 109268092), and *MIA3* (GGC 3: 18147884). Each of these variants involved different alleles (A, T) and were classified as missense variants with moderate predicted impacts on gene function. SIFT analysis consistently predicted these variants to be deleterious, with scores ranging

from 0 to 0.05, indicating a high likelihood of adverse effects on the protein functions encoded by these genes. In Kadaknath chickens, nine genes—*ESYT2* (Extended Synaptotagmin 2), *LRBA* (Lipopolysaccharide Responsive Beige-Like Anchor Protein), *BSN* (Bassoon Presynaptic Cytomatrix Protein), *ATP6V1H* (ATPase H + Transporting V1 Subunit H), *FBXW8* (F-box and WD Repeat Domain Containing 8), *FAM161A* (Family With Sequence Similarity 161 Member A), *FGF12* (Fibroblast Growth Factor 12), *IQCB1* (IQ Motif Containing B1), and *UTRN* (Utrophin)—were prioritised. These genes were the closest to significant SNPs associated with traits related to bacterial taxa such as *Bacteroides, Cloacibacillus, Clostridiales, Erysipelotrichaceae, Helicobacter, Parasutterella*, and *Ruminococcaceae* abundance. The genomic positions of the significant variants were: *ESYT2* (GGC 2: 9557745), *LRBA* (GGC 4: 32872969), *BSN* (GGC 12: 2920447–2946803), *ATP6V1H* (GGC 2: 110272741), *FBXW8* (GGC 15: 11677524), *FAM161A* (GGC 3: 15445199–15451241), *FGF12* (GGC 9: 13611602), *IQCB1* (GGC 7: 26302335–26316863), and *UTRN* (GGC 3: 46127091). The identified variants involved different alleles (A, C, G, T) and were classified as either missense variants or combined missense and splice region variants, all with a moderate predicted impact. SIFT analysis consistently indicated these variants as deleterious, with scores ranging from 0 to 0.05, suggesting potential deleterious effects on the protein function of these genes

Fine-mapping analysis was performed to identify potential causal variants within the GWAS-associated candidate regions. Variants from the WGS data with predicted high or moderate functional impact on the encoded proteins by VEP were tested for a significant association with the respective phenotypes of interest using a linear model. This analysis identified 36 significant associations across 5 phenotypes in Cobb400 (Supplementary Table 9) and 60 significant associations across 6 phenotypes in Kadaknath (Supplementary Table 10). The significant variants spanned 15 genes in Cobb400 and 14 genes in Kadaknath. Notably, the largest number of significant associations were observed with PCoA axis 4 in Kadaknath (n = 52) and with *Campylobacter* in Cobb400 (n = 28). In Cobb400, variants associated with *Campylobacter* were located on GGC 4, spanning multiple genes, including *BCORL1, ZDHHC9, SASH3, ERCC6L*, and *NHSL2*. The strongest associations were observed in *ERCC6L* (adjusted $R^2 = 1$; *P*-value = 3.10×10^{-62}) and *SASH3* (adjusted $R^2 = 0.995$; *P*-value = 5.50×10^{-7}). Additionally, two highly significant variants in *PCGF2* were identified on GGC 27 (adjusted $R^2 = 1$; *P*-value = 1.96×10^{-7}). For *Fournierella*, variants on GGC 17 were mapped to *SETX* and *TTF1*, with significant associations (adjusted $R^2 = 1$; *P*-values as low as 9.45×10^{-88}). Variants associated with microbial diversity, represented by *PCoA2*, were identified on GGC Z in *ENSGALG00000000264* (adjusted $R^2 = 0.948$; *P*-value = 0.0065). Variants linked to *Rikenellaceae_RC9_gut_group* were found on GGC 3, 19, and 28, with the most significant associations located in *ASXL2* on GGC 3 (adjusted $R^2 = 1$; *P*-value = 1.48×10^{-88}). Associations with *Sutterella* were observed on GGC 1, mapped to *ENSGALG00000048789* (adjusted $R^2 = 0.839$; *P*-value = 0.0395). In Kadaknath, genetic associations were observed with *Alistipes* on GGC 6, with the gene *ENSGALG00000045754* (adjusted $R^2 = 0.674$, *P* - value = 0.0489), and with *Cloacibacillus* on GGC 1 and 2, linked to genes *ENSGALG00000051395* and *ASXL3*, respectively. Significant loci for *Helicobacter* abundance were found on GGC 15, associated with *RNFT2* (adjusted $R^2 = 0.73$, *P* - value = 0.0115). Multiple loci were associated with *Parasutterella* on GGC 3, within the genes *FAM161A* and *SEC23B*.

Additionally, loci on GGC 2, 13, and 17 were significantly linked to microbial diversity as represented by PCoA4 and PCoA6, involving genes such as *DNAJB6*, *HMMR*, *SETX, TTF1*, and *SULF2*, with the most significant associations occurring in *SETX* (adjusted R^2 up to 0.928, *P* -value = 6.55×10^{-5}).

Pathway and Network analyses

Based on the significant heritability estimates and the substantial genetic variance explained by the identified SNPs, we hypothesized that the candidate regions pinpointed by the GWAS might harbor genes involved in common pathways related to caecal microbiome composition or the abundance of individual genera. To explore this, we identified sets of annotated genes located within these candidate regions and investigated potential gene set enrichment.

For the β-diversity of the caecal microbiota in both Cobb400 and Kadaknath, several significant KEGG-enriched pathways (adjusted p-value < 0.1) were identified based on annotated genes located within 50 kb upstream or downstream of significant associated SNPs (Supplementary Fig. 11). In Cobb400, the enriched pathways were primarily associated with signal transduction, including the GnRH signaling pathway, MAPK signaling pathways, and RIG-I-like receptor signaling pathways, as well as pathways related to cell junctions (Adherens junction and Tight junction) and ubiquitin-mediated proteolysis. In Kadaknath, enriched pathways were mainly related to cytoskeletal regulation (regulation of actin cytoskeleton), cell adhesion and extracellular interactions (cell adhesion molecules, ECM-receptor interaction), cell cycle control (p53 signaling pathway), and ubiquitin-mediated proteolysis.

For individual caecal genera, significant pathways (adjusted *P* -value < 0.1) were identified based on annotated genes located within 50 kb upstream and downstream of significant associated GWAS SNPs in both Cobb400 (Supplementary Fig. 12) and Kadaknath (Supplementary Fig. 13). In Cobb400 chickens, significant KEGG-enriched pathways were identified for the specific genera *Campylobacter, Desulfovibrio, Enterococcus*, *Rikenellaceae_RC9_gut_group* and *Ruminococcaceae_UCG_014.* These pathways encompassed signalling processes (Phosphatidy signaling system), epigenetic regulation (Polycomb repressive complex) and RNA processing (Spliceosome). Additionally, pathways related to metabolic and biosynthesis processes, such as Inositol phosphate metabolism, Arachidonic acid metabolism, Folate biosynthesis, Metabolism of xenobiotics by cytochrome P450, Retinol metabolism, Other glycan degradation, Sphingolipid metabolism, Amino sugar and nucleotide sugar metabolism, Terpenoid backbone biosynthesis, were enriched.

In Kadaknath chickens, significant KEGG-enriched pathways (adjusted *P*-value < 0.1) were identified for the genera *Cloacibacillus*, *Desulfovibrionaceae_X*, *Eisenbergiella*, *Helicobacter*, *Muribaculaceae_X*, *Parasutterella*, *Rikenellaceae_X*, *Ruminococcaceae.UCG.009*, and *Ruminococcus.1*. These pathways included Nicotinate and nicotinamide metabolism, Purine metabolism, and Ubiquitin-mediated proteolysis. Pathways related to adhesion and junctions were also significant, such as Adherens junction, Cell adhesion molecules, Focal adhesion, and Tight junction. In terms of the immune response, pathways like Cytokine-cytokine receptor interaction and TGF-beta signalling pathways were enriched. Additionally, several signalling pathways were identified, including the Wnt signalling pathway, Cytosolic

DNA-sensing pathway, and ECM-receptor interaction. Other significant pathways included Peroxisome, Protein processing in the endoplasmic reticulum, Vascular smooth muscle contraction, and Phototransduction, indicating their potential roles in metabolism, protein folding, muscle function, and light response mechanisms.

Discussion

Understanding host-microbiota interactions in chickens can support improved health, disease resistance, and nutrient utilization, thereby enhancing poultry productivity and reducing environmental impact. Insights gained can guide breeding programs and reduce antibiotic reliance, promoting sustainable and efficient poultry farming including utilization of low-value diets and resistance to pathogen colonization. In this study, we focused on identifying host genetic variation associated with caecal microbial communities in two chicken breeds — the indigenous Indian Kadaknath and the commercial Cobb400 broiler.

Building on a previous study that demonstrated the significant influence of environmental factors and chicken breeds on caecal microbiome composition [8], we aimed to study the host's genetic contribution by controlling for environmental and management variables. To achieve this, we co-raised chickens of the two breeds in most of the studied farms and we incorporated the first four MCA components of environmental variables as covariates in the GWAS and SNP-based heritability analysis of microbial traits. By controlling for these factors, our analysis more precisely attributed differences in the caecal microbiome to genetic factors. Nevertheless, as a field study, our findings should be interpreted with the consideration of the limitations and advantages of field-based genome-wide association studies [41]. Compared to controlled challenge experiments, the unknown and uncontrolled exposure to non-genetic factors in a field study may reduce its power, but this does not undermine its ability to demonstrate host genetic differences [18]. Moreover, demonstrating heritable resistance in field studies that simulate commercial practice is highly relevant to the production system into which selectively-bred chicken would be introduced.

Good enteric health is of key importance to poultry production, influencing performance, health and welfare. Recent studies of microbial population structures in Cobb400 and Kadaknath chickens revealed three distinct population types, termed enterotypes, one of which presented a dysbiotic aspect with low alpha diversity and high proliferation of *Campylobacter* [41]. Identification of a ‘good’ or optimal enterotype is controversial and likely to differ between breeds and management systems, however, evidence of heritability and SNPs associated with distinct PCoAs offers opportunities for selective breeding towards defined bacterial populations.

Our GWAS analysis identified significant genetic associations with caecal microbiome diversity in both Cobb400 and Kadaknath chickens, highlighting the influence of host genetics on microbial community structure. In Cobb400 chickens, significant associations were specifically observed with β - diversity, particularly involving SNPs on GGC 2 associated with PCoA axes 3 and 4. These SNPs, located several

megabases apart, suggest the involvement of multiple loci within this chromosome in shaping microbial composition. There are 7 SNPs located on GGCZ associated with PCOA axes 2, and the strongest association was identified with SNP rs316331412 within the *KIAA0825* (*C5orf36*) gene. Although *KIAA0825*'s function in chickens is uncharacterized, studies on its mouse ortholog, show expression in developing limbs, with knock-out models exhibiting skeletal irregularities, such as altered growth and bone density [42]. Genes closest to the SNPs associated with PCoA axis 2 were enriched in pathways such as Adherens Junction, Ubiquitin-Mediated Proteolysis, and MAPK Signaling, all of which play critical roles in regulating growth processes. Moreover, the enrichment of the MAPK signaling pathway in PCoA2 of Cobb400 has been previously linked to growth factors and body weight, as identified in the largest GWAS conducted in broiler chickens to date [43]. Fine-mapping association analysis identified SNP rs317417966 as significantly associated with PCoA axis 2, located on GGC Z, suggesting that this genomic region may act as a potential regulator of microbial diversity.The *MRPL15* gene encodes a mitochondrial ribosomal protein involved in protein synthesis within the mitochondria. While not previously linked to growth in chickens, it has been identified as a candidate gene associated with growth in beef cattle [44]. Its association with PCoA axis 3 in Cobb400 suggests a potential connection between mitochondrial function and the regulation of gut microbial communities, possibly influencing cellular energy metabolism and ultimately impacting the growth of Cobb400.

DNAJB6, a highly conserved co-chaperone in the heat shock protein (HSP) family that binds and refolds misfolded proteins [45], associated with PCoA axis 4 in both Cobb400 and Kadaknath and was confirmed by fine-mapping analysis using WGS data in Kadaknath. This gene is significantly upregulated in the chicken HD11 cell line during acute heat stress combined with lipopolysaccharide stimulation [46]. The studies above indicated that *DNAJB6* is capable of being involved in various types of stress through transcriptional up-regulation. Additionally, circRNA derived from *DNAJB6* (exon 2 to exon 5 on GGC2) is speculated to play a critical role in the chicken immune response [47]. This suggests that both Cobb400 and Kadaknath chickens may rely on *DNAJB6* to maintain cellular function and resilience under environmental stress or disease, supporting their growth and adaptation.

Meanwhile, in Kadaknath chickens, the identified significant SNPs associated with both α- and β-diversity were also distributed across various chromosomes. Among these, the most significant SNP associated with PCoA axis 10 was located within the *COPA* gene, and was part of a prominent association peak in the Manhattan plot, reinforcing its statistical robustness and potential biological significance. Gene *COPA* plays a critical role in intracellular protein transport and vesicular trafficking, and expression of mutant *COPA* results in endoplasmic reticulum stress and the upregulation of cytokines priming for a T helper type 17 (TH-17) response [48]. In addition, genes such as *DCAF7*, *CCNG1*, *SULF2*, and *DOT1L*, identified through GWAS associations with β-diversity of caecal microbiota in Kadaknath, are involved in key biological processes, including cell cycle regulation, gene expression, immune signalling, and extracellular matrix (ECM) modulation [49–51], all of which can directly or indirectly influence the gut environment and microbial populations. Fine-mapping revealed that most SNPs associated with PCoA axis 4 were located within the *SETX* genes. The gene *SETX* has been reported to exert an inhibitory effect on the transcriptional response to viral infection [52].

In this study we have shown that caecal microbiota structure was breed-specific, differentiating between Kadaknath and Cobb400, consistent with our previous findings [8]. Our GWAS analysis further highlighted these differences, revealing significant but distinct genetic associations with microbial diversity in both breeds. In Cobb400 chickens, genes were associated with β - diversity, primarily related to microbial regulation and growth adaptation under stress conditions, as previously stated. In contrast, Kadaknath chickens exhibited genetic associations with both α - and β - diversity, primarily impacting gut health through immune response and cell signalling pathways. These associations were further supported by enriched pathways such as ECM-receptor interaction, the p53 signalling pathway, and the regulation of the actin cytoskeleton, all of which were linked to the β - diversity of caecal microbiota in Kadaknath. These findings underscore the importance of selective breeding strategies tailored to each breed's unique genetic makeup.

Colonization by *Cloacibacillus* was relatively higher in Cobb400 chickens compared to Kadaknath, consistent with previous studies reporting a higher relative abundance of *Cloacibacillus* in fast-growing compared to slow-growing broilers [53, 54]. The relatively high heritability of *Cloacibacillus* (h^2 = 0.39) in Cobb400 suggests a strong genetic component in its colonization. GWAS, identified one significant SNP 49 kb upstream of the *ETS2* gene that was positively associated with the abundance of *Cloacibacillus*. *ETS2* has previously been reported to be essential for normal progression of the adipocyte differentiation program *in vitro*, and as a functionally important transcription factor in adipogenesis [55]. This SNP associated with *Cloacibacillus* was also significantly and positively associated with *Desulfovibrio* in Cobb400 broilers, with a moderate heritability (h^2 = 0.28). *Desulfovibrio* utilizes free hydrogen to reduce sulfate, thereby aiding in the removal of hydrogen produced during anaerobic fermentation in the gut [56]. Hydrogen removal is critical for sustaining anaerobic fermentation, ensuring a steady production of short-chain fatty acids, key energy sources that support the host's growth. These findings suggest that *ETS2* and its associated genetic variants may play a crucial role in modulating the caecal microbiome, specifically influencing the abundance of *Cloacibacillus* and *Desulfovibrio* in commercial broiler chickens. In addition to the top SNP significantly associated with *Desulfovibrio*, another significant SNP on GGC10 was identified with a negative association to *Desulfovibrio* abundance. This SNP, located within the *IGDCC3* gene (also known as NEO1-like),, a member of the immunoglobulin superfamily, that encodes a protein involved in cell adhesion and interaction. This gene is thought to regulate critical cellular processes such as growth, migration, and development by facilitating intercellular communication, potentially linking host genetic regulation to microbial composition in the gut [57].

We identified five SNPs on GGC10 that were positively associated with *Erysipelatoclostridium* abundance in Cobb400. These SNPs were mapped to the *ARRDC4* gene, whose expression is controlled through carbohydrate-response elements by a MondoA-dependent mechanism [58–60]. High glucose levels induce *ARRDC4* expression in cultured human beta cells through MondoA, which is found to be important for insulin resistance and lipid metabolism [59, 61, 62]. The abundance of *Erysipelatoclostridium* was higher in Cobb400 chickens compared to Kadaknath, consistent with

previous findings in fatty-type chicken breeds with fast growth rates [63], and *Erysipelatoclostridium* is positively associated with lipid metabolism [63]. Our findings suggest that *ARRDC4* variants may impact lipid metabolism and create a gut environment that promotes *Erysipelatoclostridium* abundance. *ARRDC4* may serve as a key regulator of the host-microbiome-lipid metabolism axis, influencing traits such as growth rate and fat deposition in chickens. Further studies are needed to confirm the role of *ARRDC4.*

Conversely, for *Cloacibacillus* (h^2 = 0.21) in Kadaknath, we identified 11 significant SNPs associated with its abundance a. Notably, some of these SNPs were located on GGC 1, similar to findings for Cobb400, but at different positions and within different genes. This suggests breed-specific effects and distinct genetic background on *Cloacibacillus* colonisation between the two breeds. Intriguingly, the GWAS results identified one significant SNP negatively associated with *Cloacibacillus* abundance located within the *LRBA* gene. VEP analysis revealed that this gene contains deleterious missense variants. *LRBA* plays a crucial role in cellular and immune processes. In humans, mutations in the *LRBA* gene result in severe immunodeficiency, characterized by hypogammaglobulinemia and recurrent infections, as well as immune dysregulation with a wide range of autoimmune manifestations [64]. *LRBA* deficiency leads to T-cell dysfunction and mislocalization of the immune receptor CTLA-4 [65], while other immune cell types are also affected, potentially contributing to conditions such as inflammatory bowel disease [66, 67]. Deleterious variants within the *LRBA* gene suggest that functional changes in the encoded proteins could impact cellular processes and immune responses, influencing *Cloacibacillus* colonization. This might result in lower *Cloacibacillus* abundance in Kadaknath, which may contribute to a slower growth rate in this breed when compared to Cobb400. This also emphasizes breed-specific host effects and distinct genetic influences on *Cloacibacillus* colonization between Kadaknath and Cobb400.

The genus *Campylobacter*, a significant foodborne zoonotic pathogen and potential threat to poultry health, was identified as a breed-specific biomarker for Kadaknath chickens in our previous study [8]. In the present study, no specific SNPs were associated with *Campylobacter* abundance in Kadaknath. However, in Cobb400 chickens, SNPs on GGC4, GGC1, and GGC27 were positively associated with *Campylobacter* abundance, highlighting differences in genetic architecture between the breeds. This was further confirmed by fine-mapping association analysis, where 26 SNPs were identified in Cobb400 chickens located between 1,599,253 and 1,684,226 bp on GGC4. This region also includes the SNP (1,646,099 bp on GGC4) previously identified by GWAS, further supporting its significance in influencing *Campylobacter* colonization. Genes located in this associated region, are *BCORL1* (transcriptional regulation) [68], *ZDHHC9* (protein palmitoylation) [69], *SASH3* (T-cell signaling) [70], *ERCC6L* (DNA repair and genomic stability) [71], *NHSL2* (cytoskeletal organization), and *APLN* (angiogenesis and metabolism) [72]. Enrichment analysis revealed that these genes are involved in pathways such as inositol phosphate metabolism, the phosphatidylinositol signaling system, and polycomb repressive complex. Dysregulation in these pathways could impair the host's ability to control pathogen colonization, leading to increased abundance of *Campylobacter* in the gut. While there is some evidence for genetic control of *Campylobacter jejuni* colonization in inbred chickens [75], the overall heritability of *Campylobacter* in the studied population was relatively low at 0.16, in line with previous studies [76–78],

suggesting that environmental and other non-genetic factors contribute more significantly in chickens. Future studies should focus on functional validation of these variants and exploring their roles in broader populations and under different environmental conditions.

Conversely, in Kadaknath we found 11 SNPs mainly located on GGC 2 that were positively associated with *Helicobacter* (h^2 = 0.36). *Helicobacter* is an emerging foodborne pathogen that commonly colonizes the gastrointestinal tract of poultry, where it can cause gastroenteritis in human [79]. The moderate heritability indicates that an important component of the variation in *Helicobacter* colonization can be attributed to host genetic factors. Six of these SNPs were located within the gene *RALYL*, implying that these variants in *RALYL* might influence *Helicobacter* colonization or abundance. *RALYL* is known for its involvement in RNA binding and regulation of gene expression, processes that are crucial for maintaining cellular homeostasis and immune responses [80]. Additionally, one SNP located within the *FBXW8* gene, with deleterious effect (missense and splice region variant), suggests a potential role of *FBXW8*, an F-box protein, that plays a critical role in ubiquitination and protein degradation and is implicated in both the MAPK signaling pathway [81] and ubiquitin-mediated proteolysis pathways.

Parasutterella, a genus of Betaproteobacteria, is part of the healthy fecal core microbiome and has been identified in various host species [82]. Its abundance is negatively correlated with high-fat diet (HFD)-induced metabolic phenotypes, including hypothalamic inflammation [83, 84], and it responds to dietary and antibiotic interventions, highlighting its role in maintaining gut and metabolic health [85, 86]. In this study, we identified a strong association peak on GGC 9 linked to *Parasutterella* abundance, with a moderate heritability (h^2 = 0.30), as well as three SNPs on GGC 3 by fine-mapping, suggesting that its colonization is influenced by host genetics across multiple genomic regions. Notably, the significant SNPs on GGC9 mapped to the *FGF12* gene, and those on GGC3 located within the *FAM161A* gene were missense variants with predicted deleterious effects suggesting a potential role on protein activity. *FGF12* (Fibroblast Growth Factor 12) plays critical roles in embryo development, and immune responses to infectious diseases [87], while *FAM161A* has been implicated in cellular structural maintenance and retinal function [88]. These findings suggest that genetic variants in these genes may affect host processes that influence the gut environment and microbial colonization.

Lactobacillus, a core caecal lumen bacterial genus, has previously been identified as candidate breed-specific biomarker for the Cobb400 [8] and may contribute to effective food conversion in terms of hydrolysing starch and other macromolecules, and the subsequent formation of short-chain fatty acids via fermentation that are absorbed by the host [8]. The higher abundance of *Lactobacillus* in Cobb400 suggests its potential role in enhancing feed conversion efficiency in this breed. No specific SNPs were associated with *Lactobacillus* in the current study, although we identified a significant heritability for this genus in Cobb400, indicating that its colonization has likely a more complex polygenic architecture and is potentially influenced more by diet or environmental factors. Except for *Lactobacillus*, other breed-specific bacterial biomarkers identified in our previous study were not significantly different between Cobb400 and Kadaknath in the current study. This could be attributed to the adjustment for substantial environmental variation (> 70%) through the inclusion of MCA components in the present analysis, rather

than just adjusting for the location. These findings suggest that the selection of bacterial biomarkers may be largely influenced by environmental factors rather than solely by breed-specific genetic differences.

In the current work, we were able to show through the GWAS the significant association between the respective microbiome trait/phenotype and identify potential variants of interest in nearby regions from the high-coverage WGS data. Even though we identified many significant associations, a smaller number of samples is a limitation for the association, therefore, the results should be interpreted with caution and further studies with larger sample sizes need to validate our findings.

Conclusion

In conclusion, our study highlights the role of host genetics in shaping caecal microbiota composition and its impact on key traits such as growth, immune function, and disease resistance in Kadaknath and Cobb400 chickens. GWAS and fine-mapping analyses identified breed-specific genetic associations with microbial diversity, including key genes such as *KIAA0825*, *ARRDC4*, *ETS2*, *COPA*, *RALYL*, *LRBA, FGF12* and *FAM161A*, which are linked to growth regulation, lipid metabolism, and immune responses. Kadaknath exhibited stronger genet resistance against colonization by pathogens, as well as supporting a healthy fecal core microbiome, including *Parasutterella*. In contrast, Cobb400 displayed genetic associations with traits linked to rapid growth The significant heritability estimates for genera such as *Cloacibacillus* and *Erysipelatoclostridium*, further highlight substantial genetic influence on microbial composition and colonization traits that contribute to rapid growth of chicken. These findings demonstrate the potential for selective breeding strategies targeting microbiome-related traits to enhance productivity, disease resistance, and sustainability in poultry production. Future work with larger sample sizes is needed to validate these associations and their functional implications.

Declarations

Ethics Approval and Consent to Participate This study was carried out in India and adopted welfare standards consistent with those established under the Animals (Scientific Procedures) Act 1986, an Act of Parliament of the United Kingdom. All protocols were approved by the Ethical Review Panel of Anand Agricultural University (AAU) and the Clinical Research Ethical Review Board (CRERB) of the Royal Veterinary College (RVC). Participating farmers were informed of the objectives of the study and written consent was obtained.

Consent for Publication Not applicable.

Availability of Data and Materials The 16S rRNA gene sequence data has been uploaded on EBI-ENA under Project ID PRJEB15343, SRA ID ERP017060. The host genotypic data will be also become publicly available upon acceptance of the manuscript.

Competing Interests The authors declare that they have no competing interests.

Funding The research work was funded by the Department of Biotechnology (DBT, reference BT/IN/Indo-UK/FADH/49/CGJ/2013) India, the Biotechnology and Biological Sciences Research Council (BBSRC, reference BB/L00478X/1) UK and the UKRI Global Challenges Research Fund (Grant No. BB/S011269/1, the One Health Poultry Hub) UK. The funding agencies had no role in sample collection, conducting experiments, data analysis or manuscript writing.

Authors' ContributionsXD analyzed and interpreted 16S sequencing, GWAS, SNP-based heritability, and WGS data, and led manuscript writing. AH conducted 16S sequencing, pathway and network analysis, fine-mapping association analysis, and contributed to manuscript writing. CD supervised the GWAS data analysis and manuscript revisions. MC contributed to the meta data analysis. DPB and AP supervised the project and finalized the manuscript. DPB, AP and FT secured the funding. All remaining authors supported analyses of metadata or microbiota and contributed to manuscript revisions. All authors read and approved the final manuscript.

Acknowledgements The authors are thankful to all the farmers for their permission and support with the sample collection. The authors would like to acknowledge the efforts of other research staff (Subhash Jakhesara, Anand Patel, Mithil Parekh, Shriram Patel, Khushal Patel and Jalpa Thakkar) in their assistance with the sample collection.

Author disclaimer

MC is now employed by the European Food Safety Authority (EFSA). However, the present article 617 is published under the sole responsibility of the authors and may not be considered as an EFSA 618 scientific output. The positions and opinions presented in this article are those of the authors 619 alone and do not necessarily represent the views/any official position or scientific works of EFSA. 620 To know about the views or scientific outputs of EFSA, please consult its website 621 www.efsa.europa.eu.

References

1. Wen C, Yan W, Mai C, Duan Z, Zheng J, Sun C, et al. Joint contributions of the gut microbiota and host genetics to feed efficiency in chickens. Microbiome. 2021;9:126. Available from: https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-021-01040-x
2. Dai X, Hinsu AT, Dadousis C, Pandit RJ, Crotta M, Limon G, et al. 584. Genome-wide association studies of chicken caecal microbiota. Proc 12th World Congr Genet Appl to Livest Prod. The Netherlands: Wageningen Academic Publishers; 2022. p. 2421–4. Available from: https://www.wageningenacademic.com/doi/10.3920/978-90-8686-940-4_584
3. Krajmalnik-Brown R, Ilhan Z-E, Kang D-W, DiBaise JK. Effects of Gut Microbes on Nutrient Absorption and Energy Regulation. Nutr Clin Pract. 2012;27:201–14. Available from: http://doi.wiley.com/10.1177/0884533611436116
4. Pickard JM, Zeng MY, Caruso R, Núñez G. Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. Immunol Rev. 2017;279:70–89. Available from:

https://onlinelibrary.wiley.com/doi/10.1111/imr.12567

5. 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;555:210–5. Available from: http://www.nature.com/articles/nature25973
6. Vangay P, Ward T, Gerber JS, Knights D. Antibiotics, Pediatric Dysbiosis, and Disease. Cell Host Microbe. 2015;17:553–64. Available from: https://linkinghub.elsevier.com/retrieve/pii/S193131281500164X
7. Zhao L, Wang G, Siegel P, He C, Wang H, Zhao W, et al. Quantitative Genetic Background of the Host Influences Gut Microbiomes in Chickens. Sci Rep. 2013;3:1163. Available from: http://www.nature.com/articles/srep01163
8. Pandit RJ, Hinsu AT, Patel N V., Koringa PG, Jakhesara SJ, Thakkar JR, et al. Microbial diversity and community composition of caecal microbiota in commercial and indigenous Indian chickens determined using 16s rDNA amplicon sequencing. Microbiome. 2018;6:115. Available from: https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0501-9
9. 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;444:1027–31. Available from: http://www.nature.com/articles/nature05414
10. OECD/FAO. OECD-FAO Agricultural Outlook 2019-2028 [Internet]. Rome: OECD Publishing, Paris/Food and Agriculture Organization of the United Nations; 2019. Available from: https://www.oecd-ilibrary.org/agriculture-and-food/oecd-fao-agricultural-outlook-2019-2028_agr_outlook-2019-en
11. USDA/United States Department of Agriculture. Livestock and poultry: world markets and trade. 2020;
12. Borda-Molina D, Seifert J, Camarinha-Silva A. Current Perspectives of the Chicken Gastrointestinal Tract and Its Microbiome. Comput Struct Biotechnol J. 2018;16:131–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2001037017301162
13. Kers JG, Velkers FC, Fischer EAJ, Hermes GDA, Stegeman JA, Smidt H. Host and Environmental Factors Affecting the Intestinal Microbiota in Chickens. Front Microbiol. 2018;9. Available from: http://journal.frontiersin.org/article/10.3389/fmicb.2018.00235/full
14. Suzuki TA, Phifer-Rixey M, Mack KL, Sheehan MJ, Lin D, Bi K, et al. Host genetic determinants of the gut microbiota of wild mice. Mol Ecol. 2019;28:3197–207. Available from: https://onlinelibrary.wiley.com/doi/10.1111/mec.15139
15. Haunshi S, Prince LLL. Kadaknath: a popular native chicken breed of India with unique black colour characteristics. Worlds Poult Sci J. 2021;77:427–40. Available from: https://www.tandfonline.com/doi/full/10.1080/00439339.2021.1897918
16. Hinsu AT, Thakkar JR, Koringa PG, Vrba V, Jakhesara SJ, Psifidi A, et al. Illumina Next Generation Sequencing for the Analysis of Eimeria Populations in Commercial Broilers and Indigenous

Chickens. Front Vet Sci. 2018;5. Available from: https://www.frontiersin.org/article/10.3389/fvets.2018.00176/full

17. Lê S, Josse J, Husson F. FactoMineR : An R Package for Multivariate Analysis. J Stat Softw. 2008;25. Available from: http://www.jstatsoft.org/v25/i01/
18. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.; 2013. Available from: http://www.r-project.org/
19. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3. Available from: http://www.nature.com/articles/nmeth.3869
20. 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;37:852–7. Available from: http://www.nature.com/articles/s41587-019-0209-9
21. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome. 2018;6:90. Available from: https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0470-z
22. 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. 2012;41:D590–6. Available from: http://academic.oup.com/nar/article/41/D1/D590/1069277/The-SILVA-ribosomal-RNA-gene-database-project
23. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. Watson M, editor. PLoS One. 2013;8:e61217. Available from: https://dx.plos.org/10.1371/journal.pone.0061217
24. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.; 2013.
25. Oksanen J, Blanchet FG, Kindt R, Legendere P, Minchin PR, O'Hara RB. Vegan: Community Ecology Package. 2015;
26. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8
27. Yu G, Smith DK, Zhu H, Guan Y, Lam TT. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. McInerny G, editor. Methods Ecol Evol. 2017;8:28–36. Available from: https://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12628
28. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007;23:1294–6. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btm108
29. Kranis A, Gheyas AA, Boschiero C, Turner F, Yu L, Smith S, et al. Development of a high density 600K SNP genotyping array for chicken. BMC Genomics. 2013;14:59. Available from:

http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-59

30. 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;81:559–75. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002929707613524
31. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012;44:821–4. Available from: http://www.nature.com/articles/ng.2310
32. Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, et al. rMVP: A Memory-efficient, Visualization-enhanced, and Parallel-accelerated Tool for Genome-wide Association Study. Genomics Proteomics Bioinformatics. 2021;19:619–28. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1672022921000504
33. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20. Available from: https://academic.oup.com/bioinformatics/article/30/15/2114/2390096
34. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009;25:1754–60. Available from: https://academic.oup.com/bioinformatics/article/25/14/1754/225615
35. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9. Available from: https://academic.oup.com/bioinformatics/article/25/16/2078/204688
36. 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;20:1297–303. Available from: http://genome.cshlp.org/lookup/doi/10.1101/gr.107524.110
37. Zhang C, Dong S-S, Xu J-Y, He W-M, Yang T-L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Schwartz R, editor. Bioinformatics. 2019;35:1786–8. Available from: https://academic.oup.com/bioinformatics/article/35/10/1786/5132693
38. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10. Available from: https://academic.oup.com/gigascience/article/doi/10.1093/gigascience/giab008/6137722
39. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biol. 2016;17:122. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-0974-4
40. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov. 2021;2:100141. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2666675821000667
41. Hay MC, Hinsu AT, Koringa PG, Pandit RJ, Liu P-Y, Parekh MJ, et al. Chicken caecal enterotypes in indigenous Kadaknath and commercial Cobb chicken lines are associated with Campylobacter

abundance and influenced by farming practices. Front Microbiomes. 2023;2. Available from: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1301609/full

42. Koscielny G, Yaikhom G, Iyer V, Meehan TF, Morgan H, Atienza-Herrero J, et al. The International Mouse Phenotyping Consortium Web Portal, a unified point of access for knockout mice and related phenotyping data. Nucleic Acids Res. 2014;42:D802–9. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt977
43. Dadousis C, Somavilla A, Ilska JJ, Johnsson M, Batista L, Mellanby RJ, et al. A genome-wide association analysis for body weight at 35 days measured on 137,343 broiler chickens. Genet Sel Evol. 2021;53:70. Available from: https://gsejournal.biomedcentral.com/articles/10.1186/s12711-021-00663-w
44. Grigoletto L, Brito LF, Mattos EC, Eler JP, Bussiman FO, Silva B da CA, et al. Genome-wide associations and detection of candidate genes for direct and maternal genetic effects influencing growth traits in the Montana Tropical® Composite population. Livest Sci. 2019;229:64–76. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1871141319303932
45. Whitley D, Goldberg SP, Jordan WD. Heat shock proteins: A review of the molecular chaperones. J Vasc Surg [Internet]. 1999;29:748–51. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0741521499703290
46. Slawinska A, Hsieh JC, Schmidt CJ, Lamont SJ. Heat Stress and Lipopolysaccharide Stimulation of Chicken Macrophage-Like Cell Line Activates Expression of Distinct Sets of Genes. Shelden EA, editor. PLoS One. 2016;11:e0164575. Available from: https://dx.plos.org/10.1371/journal.pone.0164575
47. Tan J, Sahaer P, Li H, Han W, Sun H. The expression, function, and network regulation of circDNAJB6 in chicken macrophages under lipopolysaccharide (LPS) stimulation. Dev Comp Immunol. 2024;151:105095. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0145305X2301248X
48. Angers S, Thorpe CJ, Biechele TL, Goldenberg SJ, Zheng N, MacCoss MJ, et al. The KLHL12–Cullin-3 ubiquitin ligase negatively regulates the Wnt–β-catenin pathway by targeting Dishevelled for degradation. Nat Cell Biol. 2006;8:348–57. Available from: https://www.nature.com/articles/ncb1381
49. Frendo-Cumbo S, Li T, Ammendolia DA, Coyaud E, Laurent EMN, Liu Y, et al. DCAF7 regulates cell proliferation through IRS1-FOXO1 signaling. iScience. 2022;25:105188. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2589004222014602
50. Chen S, Liu H, Yan C, Li Y, Xiao J, Zhao X. Fecal microbiota transplantation provides insights into the consequences of transcriptome profiles and cell energy in response to circadian misalignment of chickens. Poult Sci. 2024;103:103926. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0032579124005054
51. Nguyen AT, Zhang Y. The diverse functions of Dot1 and H3K79 methylation. Genes Dev. 2011;25:1345–58. Available from: http://genesdev.cshlp.org/lookup/doi/10.1101/gad.2057811

52. Miller MS, Rialdi A, Ho JSY, Tilove M, Martinez-Gil L, Moshkina NP, et al. Senataxin suppresses the antiviral transcriptional response and controls viral biogenesis. Nat Immunol. 2015;16:485–94. Available from: https://www.nature.com/articles/ni.3132
53. Singh KM, Shah T, Deshpande S, Jakhesara SJ, Koringa PG, Rank DN, et al. High through put 16S rRNA gene-based pyrosequencing analysis of the fecal microbiota of high FCR and low FCR broiler growers. Mol Biol Rep. 2012;39:10595–602. Available from: http://link.springer.com/10.1007/s11033-012-1947-7
54. Yan C, Xiao J, Chen D, Turner SP, Li Z, Liu H, et al. Feed Restriction Induced Changes in Behavior, Corticosterone, and Microbial Programming in Slow- and Fast-Growing Chicken Breeds. Animals. 2021;11:141. Available from: https://www.mdpi.com/2076-2615/11/1/141
55. Birsoy K, Berry R, Wang T, Ceyhan O, Tavazoie S, Friedman JM, et al. Analysis of gene networks in white adipose tissue development reveals a role for ETS2 in adipogenesis. Development. 2011;138:4709–19. Available from: https://journals.biologists.com/dev/article/138/21/4709/44684/Analysis-of-gene-networks-in-white-adipose-tissue
56. Vossen E, Goethals S, De Vrieze J, Boon N, Van Hecke T, De Smet S. Red and processed meat consumption within two different dietary patterns: Effect on the colon microbial community and volatile metabolites in pigs. Food Res Int. 2020;129:108793. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0963996919306799
57. Halaby DM, Mornon JPE. The immunoglobulin superfamily: An insight on its tissular, species, and functional diversity. J Mol Evol. 1998;46:389–400. Available from: https://link.springer.com/10.1007/PL00006318
58. Stoltzman CA, Peterson CW, Breen KT, Muoio DM, Billin AN, Ayer DE. Glucose sensing by MondoA:Mlx complexes: A role for hexokinases and direct regulation of thioredoxin-interacting protein expression. Proc Natl Acad Sci. 2008;105:6912–7. Available from: https://pnas.org/doi/full/10.1073/pnas.0712199105
59. Richards P, Rachdi L, Oshima M, Marchetti P, Bugliani M, Armanet M, et al. MondoA Is an Essential Glucose-Responsive Transcription Factor in Human Pancreatic β-Cells. Diabetes. 2018;67:461–72. Available from: https://diabetesjournals.org/diabetes/article/67/3/461/40054/MondoA-Is-an-Essential-Glucose-Responsive
60. Patwari P, Lee RT. An expanded family of arrestins regulate metabolism. Trends Endocrinol Metab. 2012;23:216–22. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1043276012000471
61. Ahn B, Soundarapandian MM, Sessions H, Peddibhotla S, Roth GP, Li J-L, et al. MondoA coordinately regulates skeletal myocyte lipid homeostasis and insulin signaling. J Clin Invest. 2016;126:3567–79. Available from: https://www.jci.org/articles/view/87382
62. Ahn B, Wan S, Jaiswal N, Vega RB, Ayer DE, Titchenell PM, et al. MondoA drives muscle lipid accumulation and insulin resistance. JCI Insight. 2019;4. Available from: https://insight.jci.org/articles/view/129119

63. Liu J, Wang J, Zhou Y, Han H, Liu W, Li D, et al. Integrated omics analysis reveals differences in gut microbiota and gut-host metabolite profiles between obese and lean chickens. Poult Sci. 2022;101:102165. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0032579122004540
64. Lopez-Herrera G, Tampella G, Pan-Hammarström Q, Herholz P, Trujillo-Vargas CM, Phadwal K, et al. Deleterious Mutations in LRBA Are Associated with a Syndrome of Immune Deficiency and Autoimmunity. Am J Hum Genet. 2012;90:986–1001. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002929712002182
65. Lo B, Zhang K, Lu W, Zheng L, Zhang Q, Kanellopoulou C, et al. Patients with LRBA deficiency show CTLA4 loss and immune dysregulation responsive to abatacept therapy. Science (80-). 2015;349:436–40. Available from: https://www.science.org/doi/10.1126/science.aaa1663
66. Sudan R, Fernandes S, Srivastava N, Pedicone C, Meyer ST, Chisholm JD, et al. LRBA Deficiency Can Lead to Lethal Colitis That Is Diminished by SHIP1 Agonism. Front Immunol. 2022;13. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2022.830961/full
67. Wang K, Zhan X, McAlpine W, Zhang Z, Choi JH, Shi H, et al. Enhanced susceptibility to chemically induced colitis caused by excessive endosomal TLR signaling in LRBA-deficient mice. Proc Natl Acad Sci. 2019;116:11380–9. Available from: https://pnas.org/doi/full/10.1073/pnas.1901407116
68. Wong SJ, Senkovich O, Artigas JA, Gearhart MD, Ilangovan U, Graham DW, et al. Structure and Role of BCOR PUFD in Noncanonical PRC1 Assembly and Disease. Biochemistry. 2020;59:2718–28. Available from: https://pubs.acs.org/doi/10.1021/acs.biochem.0c00285
69. Mitchell DA, Vasudevan A, Linder ME, Deschenes RJ. Thematic review series: Lipid Posttranslational Modifications. Protein palmitoylation by a family of DHHC protein S-acyltransferases. J Lipid Res. 2006;47:1118–27. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0022227520332107
70. Delmonte OM, Bergerson JRE, Kawai T, Kuehn HS, McDermott DH, Cortese I, et al. SASH3 variants cause a novel form of X-linked combined immunodeficiency with immune dysregulation. Blood. 2021;138:1019–33. Available from: https://ashpublications.org/blood/article/138/12/1019/475781/SASH3-variants-cause-a-novel-form-of-X-linked
71. Ryan DP, Owen-Hughes T. Snf2-family proteins: chromatin remodellers for any occasion. Curr Opin Chem Biol. 2011;15:649–56. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1367593111001293
72. Fournel A, Drougard A, Duparc T, Marlin A, Brierley SM, Castro J, et al. Apelin targets gut contraction to control glucose metabolism via the brain. Gut. 2017;66:258–69. Available from: https://gut.bmj.com/lookup/doi/10.1136/gutjnl-2015-310230
73. Jung A, Rautenschlein S. Comprehensive report of an Enterococcus cecoruminfection in a broiler flock in Northern Germany. BMC Vet Res. 2014;10:311. Available from: http://bmcvetres.biomedcentral.com/articles/10.1186/s12917-014-0311-7
74. Humphrey S, Chaloner G, Kemmett K, Davidson N, Williams N, Kipar A, et al. Campylobacter jejuni Is Not Merely a Commensal in Commercial Broiler Chickens and Affects Bird Welfare. Maloy S, editor.

MBio. 2014;5. Available from: https://journals.asm.org/doi/10.1128/mBio.01364-14

75. Boyd Y, Herbert EG, Marston KL, Jones MA, Barrow PA. Host genes affect intestinal colonisation of newly hatched chickens by Campylobacter jejuni. Immunogenetics. 2005;57:248–53. Available from: http://link.springer.com/10.1007/s00251-005-0790-6
76. Bailey RA, Kranis A, Psifidi A, Watson KA, Rothwell L, Hocking PM, et al. Colonization of a commercial broiler line by Campylobacter is under limited genetic control and does not significantly impair performance or intestinal health. Poult Sci. 2018;97:4167–76. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0032579119302445
77. Psifidi A, Fife M, Howell J, Matika O, van Diemen PM, Kuo R, et al. The genomic architecture of resistance to Campylobacter jejuni intestinal colonisation in chickens. BMC Genomics. 2016;17:293. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-016-2612-7
78. Psifidi A, Kranis A, Rothwell LM, Bremner A, Russell K, Robledo D, et al. Quantitative trait loci and transcriptome signatures associated with avian heritable resistance to Campylobacter. Sci Rep. 2021;11:1623. Available from: https://www.nature.com/articles/s41598-020-79005-7
79. Perry S, Sanchez M de la L, Yang S, Haggerty TD, Hurst P, Perez-Perez G, et al. Gastroenteritis and Transmission of Helicobacter pylori Infection in Households1. Emerg Infect Dis. 2006;12:1701–8. Available from: https://wwwnc.cdc.gov/eid/article/12/11/06-0086_article
80. Liang Z, Rehati A, Husaiyin E, Chen D, Jiyuan Z, Abuduaini B. RALY regulate the proliferation and expression of immune/inflammatory response genes via alternative splicing of FOS. Genes Immun. 2022;23:246–54. Available from: https://www.nature.com/articles/s41435-022-00178-4
81. Wang H, Chen Y, Lin P, Li L, Zhou G, Liu G, et al. The CUL7/F-box and WD Repeat Domain Containing 8 (CUL7/Fbxw8) Ubiquitin Ligase Promotes Degradation of Hematopoietic Progenitor Kinase 1. J Biol Chem. 2014;289:4009–17. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0021925820441870
82. Willing BP, Dicksved J, Halfvarson J, Andersson AF, Lucio M, Zheng Z, et al. A Pyrosequencing Study in Twins Shows That Gastrointestinal Microbial Profiles Vary With Inflammatory Bowel Disease Phenotypes. Gastroenterology. 2010;139:1844-1854.e1. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0016508510012990
83. Kreutzer C, Peters S, Schulte DM, Fangmann D, Türk K, Wolff S, et al. Hypothalamic Inflammation in Human Obesity Is Mediated by Environmental and Genetic Factors. Diabetes. 2017;66:2407–15. Available from: https://diabetesjournals.org/diabetes/article/66/9/2407/40091/Hypothalamic-Inflammation-in-Human-Obesity-Is
84. Danneskiold-Samsøe NB, Andersen D, Radulescu ID, Normann-Hansen A, Brejnrod A, Kragh M, et al. A safflower oil based high-fat/high-sucrose diet modulates the gut microbiota and liver phospholipid profiles associated with early glucose intolerance in the absence of tissue inflammation. Mol Nutr Food Res. 2017;61. Available from: https://onlinelibrary.wiley.com/doi/10.1002/mnfr.201600528

85. Everard A, Lazarevic V, Derrien M, Girard M, Muccioli GG, Neyrinck AM, et al. Responses of Gut Microbiota and Glucose and Lipid Metabolism to Prebiotics in Genetic Obese and Diet-Induced Leptin-Resistant Mice. Diabetes. 2011;60:2775–86. Available from: https://diabetesjournals.org/diabetes/article/60/11/2775/33512/Responses-of-Gut-Microbiota-and-Glucose-and-Lipid
86. Staley C, Kaiser T, Beura LK, Hamilton MJ, Weingarden AR, Bobr A, et al. Stable engraftment of human microbiota into mice with a single oral gavage following antibiotic conditioning. Microbiome. 2017;5:87. Available from: https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-017-0306-2
87. Truong AD, Tran HTT, Chu NT, Nguyen HT, Vu TH, Hong Y, et al. Genome-wide identification, organization, and expression profiles of the chicken fibroblast growth factor genes in public databases and Vietnamese indigenous Ri chickens against highly pathogenic avian influenza H5N1 virus infection. Anim Biosci. 2023;36:570–83. Available from: http://animbiosci.org/journal/view.php?doi=10.5713/ab.22.0277
88. Bandah-Rozenfeld D, Mizrahi-Meissonnier L, Farhy C, Obolensky A, Chowers I, Pe'er J, et al. Homozygosity Mapping Reveals Null Mutations in FAM161A as a Cause of Autosomal-Recessive Retinitis Pigmentosa. Am J Hum Genet. 2010;87:382–91. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002929710003794

Tables

Tables 1 to 3 are available in the Supplementary Files section.

Figures

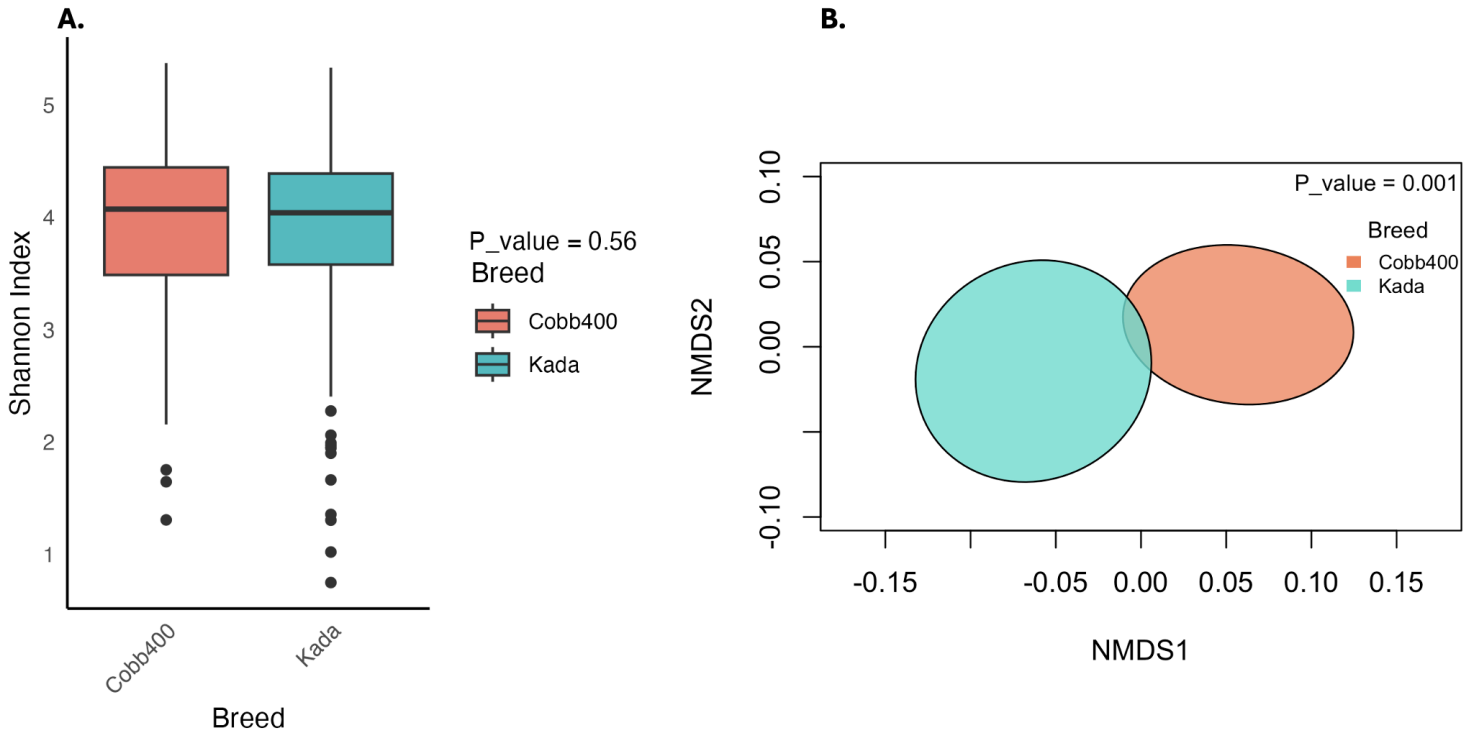

Figure 1

(A) *a* - diversity between Cobb400 and Kadaknath chickens. *a* - diversity was calculated based on the Shannon index, which was not significantly different between Cobb400 and Kadaknath ($P>0.05$). (B) *b* - diversity between Cobb400 and Kadaknath chickens. *b* - diversity was calculated by Bray-Curtis's distance, which was significantly different between Cobb400 and Kadaknath ($P = 0.001$) after taking into account the environmental variation. NMDS = Nonmetric multidimensional scaling.

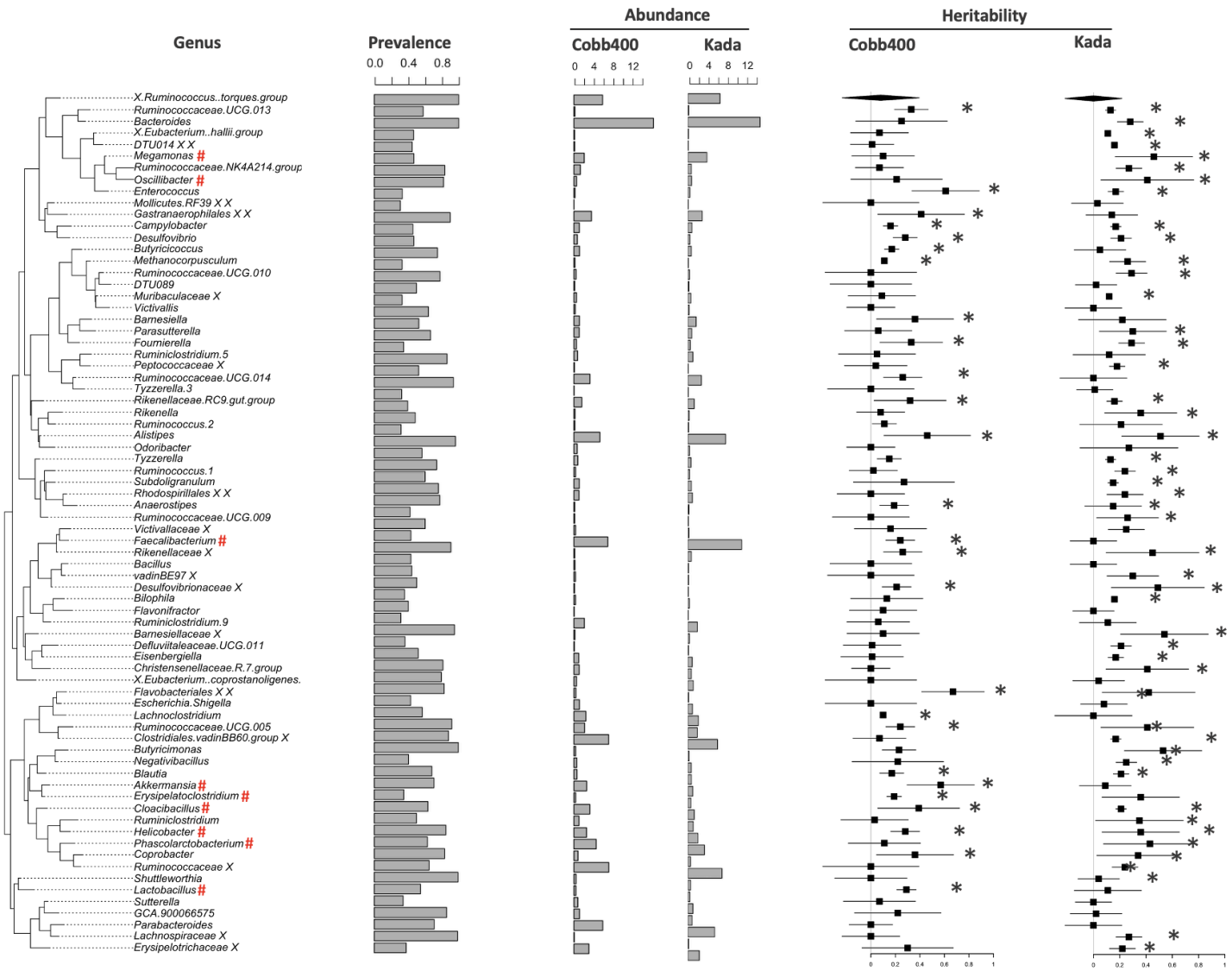

Figure 2

Maximum-likelihood phylogenetic tree of caecal microbial genera used in genome-wide association studies. Bar plots on the right describe the prevalence and the standardized mean abundance of bacterial genera in the whole chicken population, and estimated SNP-based heritability (*highlight the heritability estimatesdetermined to be different from zero using GCTA-GREML). The red # indicates the relative abundance of the genera that were different between Kadaknathand Cobb400.

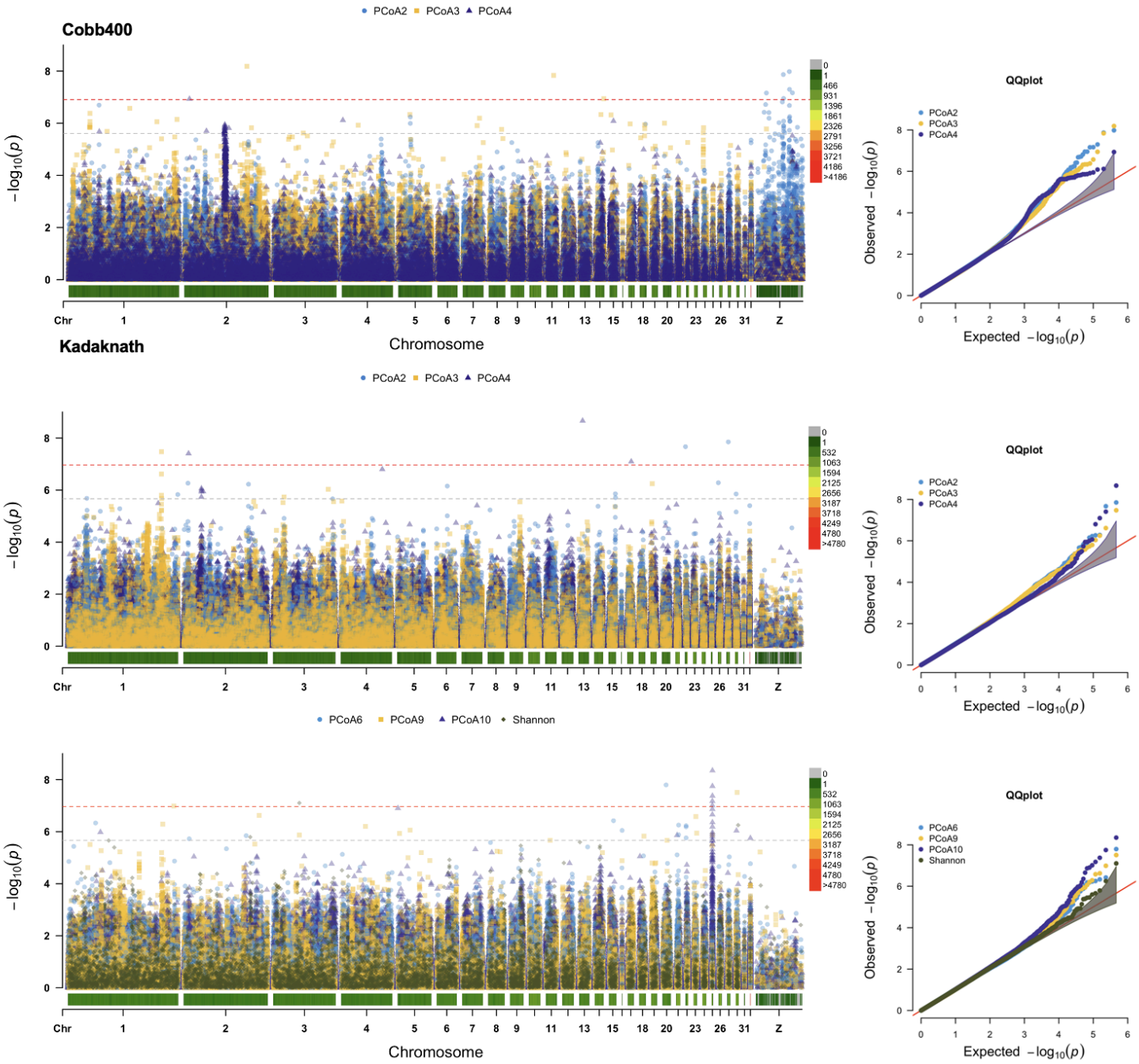

Figure 3

Manhattan plots and QQ plots display the GWAS results for alpha- (Shannon index) and beta-diversity (PCoA axes) in Cobb400 and Kadaknath using the imputed 600K (B) HD arrays. (i) Genomic location is plotted against – log10(P) in the Manhattan plot. Genome-wide ($P < 0.05$) and suggestive genome-wide thresholds are shown as red and grey lines, respectively. (ii) QQ plot of observed P - values against the expected P - values for alpha- (Shannon index) and beta-diversity (PCoA axes) in Cobb400 and Kadaknath.

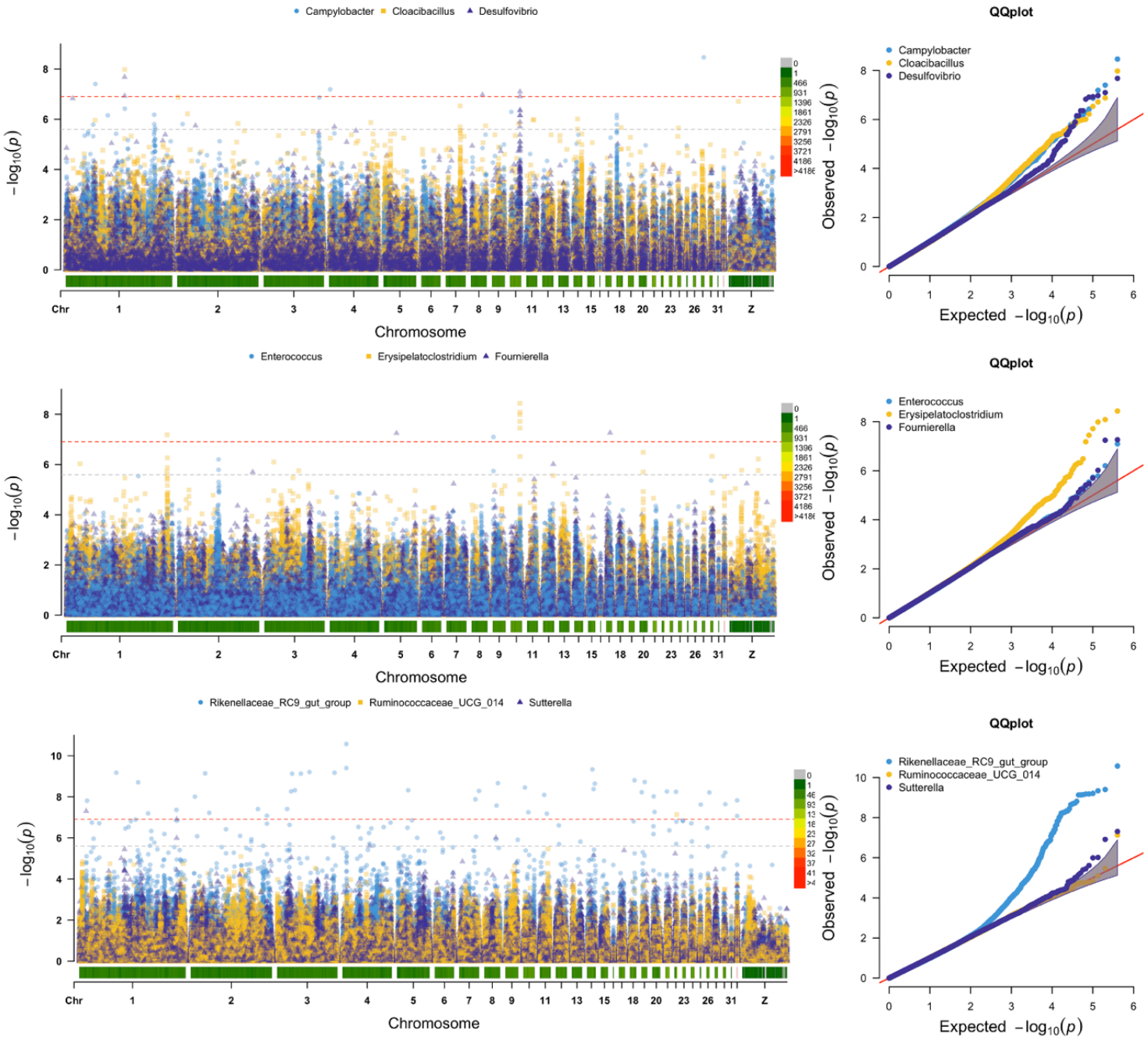

Figure 4

Manhattan plots and QQ plots display the GWAS results for significant SNP-based heritable phenotypes of caecal bacterial genera in Cobb400 using the imputed 600K (B) HD arrays. (i) Genomic location is plotted against – log10(P) in the Manhattan plot. Genome-wide ($P < 0.05$) and suggestive genome-wide thresholds are shown as red and grey lines, respectively. (ii) QQ plot of observed P - values against the expected P - values for significant SNP-based heritable phenotypes of caecal bacterial genera in Cobb400.

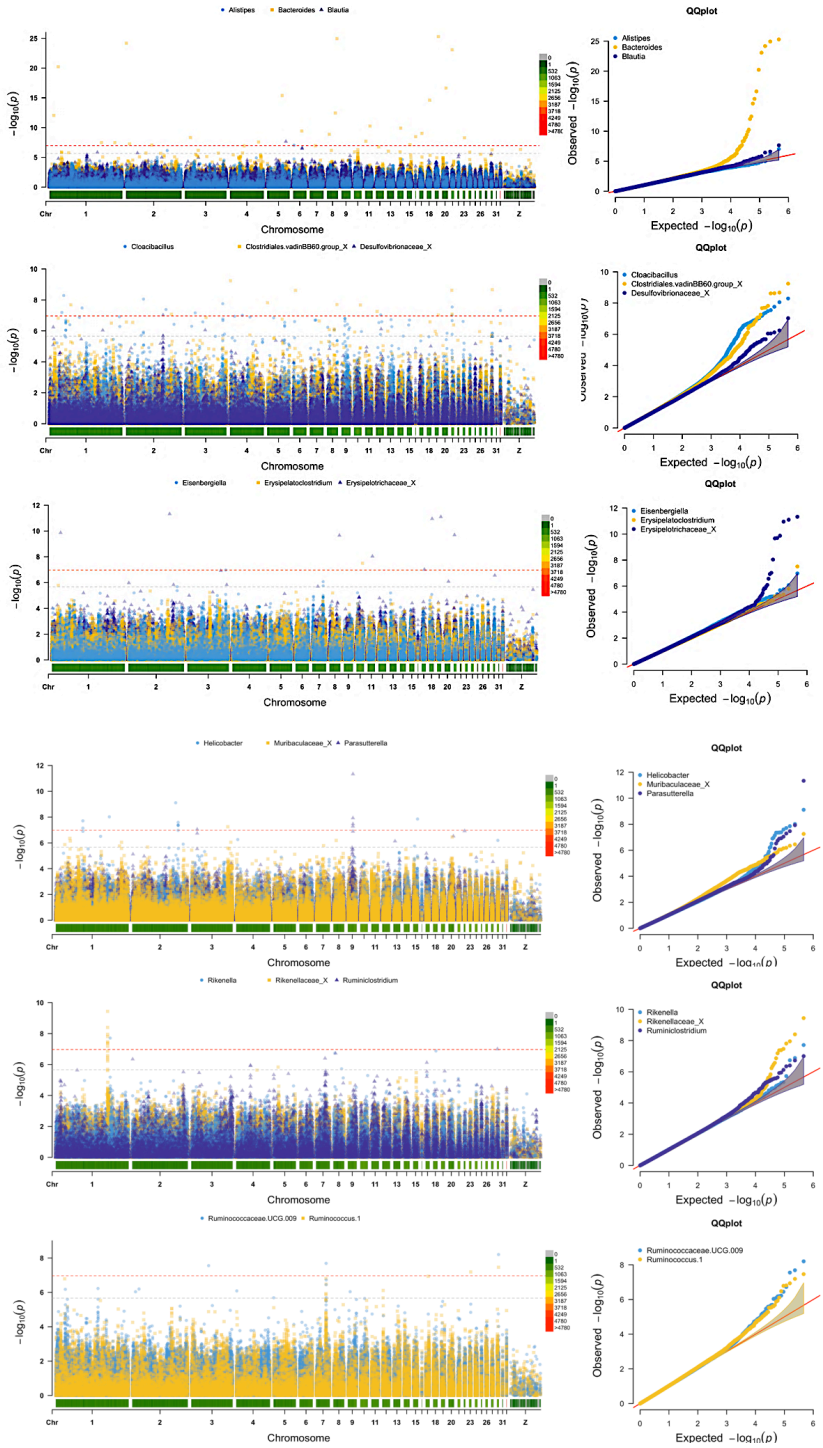

Figure 5

Manhattan plots and QQ plots display the GWAS results for significant SNP-based heritable phenotypes of caecal bacterial genera in Kadaknath using the imputed 600K (B) HD arrays. (i) Genomic location is plotted against – log10(P) in the Manhattan plot. Genome-wide ($P < 0.05$) and suggestive genome-wide thresholds are shown as red and grey lines, respectively. (ii) QQ plot of observed P - values against the

expected P - values for significant SNP-based heritable phenotypes of caecal bacterial genera in Kadaknath.

Supplementary Files

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

- Tables.docx
- SupplementaryTable1.xlsx
- SupplementaryMaterials.docx
- Supplementarytable8.xlsx