Jiao *et al. BMC Microbiology* (2025) 25:176
https://doi.org/10.1186/s12866-025-03892-x

BMC Microbiology

RESEARCH Open Access

Multi-omics analysis of host-microbiome interactions in a mouse model of congenital hepatic fibrosis

Mengfan Jiao[1], Ye Sun[1], Zixing Dai[1], Xiaoxue Hou[2], Xizhi Yin[1], Qingling Chen[1], Rui Liu[1,3], Yuwen Li[4*] and Chuanlong Zhu[1,3*]

Abstract

Background Congenital hepatic fibrosis (CHF) caused by mutations in the polycystic kidney and hepatic disease 1 (*PKHD1*) gene is a rare genetic disorder with poorly understood pathogenesis. We hypothesized that integrating gut microbiome and metabolomic analyses could uncover distinct host-microbiome interactions in CHF mice compared to wild-type controls.

Methods *Pkhd1*$^{del3–4/del3–4}$ mice were generated using CRISPR/Cas9 technology. Fecal samples were collected from 11 *Pkhd1*$^{del3–4/del3–4}$ mice and 10 littermate wild-type controls. We conducted a combined study using 16 S rDNA sequencing for microbiome analysis and untargeted metabolomics. The gut microbiome and metabolome data were integrated using Data Integration Analysis for Biomarker discovery using Latent cOmponents (DIABLO), which helped identify key microbial and metabolic features associated with CHF.

Results CHF mouse model was successfully established. Our analysis revealed that the genera *Mucispirillum*, *Eisenbergiella*, and *Oscillibacter* were core microbiota in CHF, exhibiting significantly higher abundance in *Pkhd1*$^{del3–4/del3–4}$ mice and strong positive correlations among them. Network analysis demonstrated robust associations between the gut microbiome and metabolome. Multi-omics dimension reduction analysis demonstrated that both the microbiome and metabolome could effectively distinguish CHF mice from controls, with area under the curve of 0.883 and 0.982, respectively. A significant positive correlation was observed between the gut microbiome and metabolome, highlighting the intricate relationship between these two components.

Conclusion This study identifies distinct metabolic and microbiome profiles in *Pkhd1*$^{del3–4/del3–4}$ mice. Multi-omics analysis effectively differentiates CHF mice from controls and identified potential biomarkers. These findings indicate that gut microbiota and metabolites are integral to the pathogenesis of CHF, offering novel insights into the disease mechanism.

Keywords Metabolomics, Microbiome, PKHD1, Congenital hepatic fibrosis

*Correspondence:
Yuwen Li
ywli@jsph.org.cn
Chuanlong Zhu
zhuchuanlong@jsph.org.cn
[1]Department of Infectious Disease, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China
[2]Department of Respiratory and Critical Care Medicine, Shandong Provincial Hospital, Shandong First Medical University, Jinan, China
[3]Department of Infectious and Tropical Diseases, The Second Affiliated Hospital, NHC Key Laboratory of Tropical Disease Control, Hainan Medical University, Haikou 570216, China
[4]Department of Pediatrics, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China

Introduction

The polycystic kidney and hepatic disease 1 (PKHD1), located on human chromosome 6p12.2, is one of the largest human genes identified to date [1]. This gene encodes a 4074-amino-acid protein known as fibrocystin (FPC). FPC is a single-pass transmembrane protein localized to primary cilia and basal bodies, characterized by a large extracellular domain, a transmembrane domain, and a relatively short cytoplasmic tail [2]. In the liver, FPC is primarily expressed in the primary cilia of bile duct epithelial cells and is involved in bile duct differentiation and maturation [3].

Mutations in *PKHD1* can lead to three distinct diseases: autosomal recessive polycystic kidney disease, congenital hepatic fibrosis (CHF), and Caroli disease, with an estimated incidence ranging from 1 in 10,000 to 1 in 40,000 [4–6]. CHF is a rare disease characterized by cystic dilatation of bile ducts and marked periportal fibrosis [7]. CHF is a multisystem disorder frequently overlooked or misdiagnosed clinically, necessitating multiple diagnostic modalities for accurate confirmation [8]. Despite its rarity and differences from those of acquired hepatic fibrosis, the pathogenesis of CHF remains incompletely understood. Some studies suggest that FPC-deficient bile duct cells may contribute to the development of CHF by secreting chemokines, thereby recruiting macrophage and inducing fibrosis, a process that can be mitigated by macrophage depletion [3]. In recent years, genetically engineered mouse models and organoid models generated using various technologies have enabled the recapitulation of CHF development, providing valuable tools for investigating its pathogenesis [9, 10].

The gut and liver share important anatomical and functional connections, and their interactions has been implicated in the pathogenesis of various liver diseases [11, 12]. In hepatic fibrosis, the delicate balance of the gut-liver axis is disrupted, leading to intestinal barrier dysfunction and alterations in the microbial community, which in turn modulate liver immune function and contribute to disease progression [13, 14]. Notably, while previous studies have primarily focused on alterations in the gut microbiota in acquired hepatic fibrosis, few have explored the gut microbiota characteristics in CHF [15, 16].

Multi-omics sequencing technologies have revolutionized new insights into biological research [17, 18]. The gut microbiome is known to interact with host genetics and environmental factors, influencing disease susceptibility and progression [19]. By integrating gut microbiome and metabolomic data, and employing machine learning classifiers and network analysis, we can uncover the relationship between CHF development and the microbiome. This approach enables the identification of microbial and metabolic alterations associated with disease exposure, providing novel insights into the complex progression of genetic diseases.

Methods

Animal and experimental procedures

All animal experiments were approved by the Animal Care and Use Committee of Nanjing Medical University and adhered to the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines. *Pkhd-1*$^{del3-4/del3-4}$ mice were constructed by GemPharmatech on a C57BL/6J background. To generate *Pkhd1*$^{del3-4/del3-4}$ mice, CRISPR/Cas9 was employed to delete exons 3–4 of *PKHD1* gene. The genotype of the resultant mice was confirmed by polymerase chain reaction (PCR) analysis of tail DNA. In the genotyping PCR, primer set 1 amplifies the 311-bp knockout allele, while primer set 2 amplifies the 288-bp wild-type allele. Wild-type homozygotes exhibit a band only with primer set 2, knockout homozygotes display a band only with primer set 1, and heterozygotes present bands with both primer sets, indicating the presence of both alleles. (Primer: F1, GCAATGCCTGGTTTACCCTTG, R1, AACTGGAAAGGGGCATAACGGT; F2, CAAACGCAATATCTTACATTGCCC, R2, CTAAGCAGAGCAAAGACTGTTTGG)

After successful mouse modeling, fresh fecal samples were collected in the morning and stored in sterile freezing tubes. The samples were immediately frozen in liquid nitrogen and then kept at -80 °C until further analysis. All mice were fasted overnight and anesthetized with isoflurane prior tissue sample collection.

Histologic evaluation

Mouse liver tissues were fixed in 4% paraformaldehyde, dehydrated, and embedded in paraffin wax. Section (4 μm) were cut and stained with hematoxylin and eosin (H&E) for histological examination. Additional sections were stained with a combination of potassium dichromate, acid fuchsin, and aniline blue to visualize collagen fibers, which appeared blue, while cytoplasm and nuclei stained red and black/blue, respectively.

Transmission electron microscopy

For transmission electron microscopy (TEM) observation of cilia in intrahepatic bile duct epithelial cells, liver tissues were immediately fixed in 2.5% glutaraldehyde overnight at 4 °C, followed by washing in phosphate-buffered saline (PBS) and post-fixation with 1% osmium tetroxide for 2 h. Samples were dehydrated in graded ethanol and embedded in epoxy resin. Ultrathin Sects. (60–80 nm) were stained with uranyl acetate and lead citrate, and observed under TEM. Ciliary ultrastructure were documented.

DNA extraction and sequencing

Fecal microbial DNA was extracted utilizing the Fecal Genome DNA Extraction Kit (AU46111-96, BioTeke, China), following the manufacturer's protocol. Subsequent quantification of the extracted DNA was performed using a Qubit fluorometer (Invitrogen, USA). To amplify the total DNA, polymerase chain reaction (PCR) was employed with the universal primer pair 341 F/805R (341 F: 5′-CCT ACGGGNGGCWGCAG-3′; 805R: 5′-GACTACHVGGG TATCTAATCC-3′). The PCR cycling conditions consisted of an initial denaturation step at 98 °C for 30 s, followed by 32 cycles of denaturation at 98 °C for 10 s, annealing at 54 °C for 30 s, and extension at 72 °C for 45 s. A final extension step was conducted at 72 °C for 10 min. The resultant PCR amplicons were purified using AMPure XT Beads (Beckman Coulter Genomics, Danvers, MA, USA) and quantified using a Qubit fluorometer (Invitrogen, USA). The quality of the purified PCR products was assessed using an Agilent 2100 Bioanalyzer (Agilent, USA) and Illumina library quantitative kits (Kapa Biosciences, Woburn, MA, USA). Finally, the qualified PCR products were pooled and sequenced on an Illumina NovaSeq 6000 platform (PE250) at LC-Bio Technology Co., Ltd., Hangzhou, China.

Data processing and 16 S rDNA sequencing analysis

To preprocess the raw sequencing data, cutadapt (v1.9) was employed to remove sequencing primers from the de-multiplexed reads. Subsequently, paired-end reads were merged using FLASH (v1.2.8) to generate contiguous sequences. Quality control measures were implemented to filter out low-quality reads, including those with quality scores below 20, short reads (< 100 bp), and reads containing more than 5% ambiguous bases ('N' records), using the sliding-window algorithm method in fqtrim (v0.94). This rigorous filtering process ensured the retention of high-quality clean tags. To further refine the dataset, chimeric sequences were identified and removed using Vsearch software (v2.3.4). Denoising and amplicon sequence variant (ASV) generation were performed using DADA2, a robust algorithm for resolving sequence variation. Taxonomic classification of the ASVs was achieved through sequence alignment using the Quantitative Insights Into Microbial Ecology 2 (QIIME2) plugin feature-classifier, with reference to the SILVA and NT-16 S databases.

Non-targeted metabolomics profiling

Sample preparation involved thawing the stool samples on ice, followed by metabolite extraction using an 80% methanol buffer. Chromatographic separations were performed on an UltiMate 3000 UPLC System (Thermo Fisher Scientific, Bremen, Germany), equipped with an ACQUITY UPLC T3 column (100 mm*2.1 mm, 1.8 μm, Waters, Milford, USA) for reversed-phase separation. The column oven was maintained at a constant temperature of 40 °C. A binary solvent system was employed, consisting of solvent A (5mM ammonium acetate and 5mM acetic acid) and solvent B (acetonitrile), with a flow rate of 0.3 ml/min. Metabolite detection was carried out using a high-resolution tandem mass spectrometer Q-Exactive (Thermo Scientific), operated in both positive and negative ion modes. The Q-Exactive was configured to collect precursor spectra (70–1050 m/z) at a resolution of 70,000, with an automatic gain control (AGC) target of 3e6 and a maximum inject time of 100 ms. Fragment spectra were collected at a resolution of 17,500, with an AGC target of 1e5 and a maximum inject time of 80 ms, using a top 3 configuration in data-dependent acquisition mode. To ensure the stability of the LC-MS system throughout the acquisition process, a quality control sample (pool of all samples) was analyzed after every 10 samples.

Analysis of non-targeted metabolomics profiling data

Mass spectrometry (MS) data preprocessing was conducted using XCMS software, which involved a series of steps including peak detection, peak grouping, retention time alignment, and annotation of isotopes and adducts. To facilitate data analysis, LC-MS raw data files were converted to mzXML format and subsequently processed using the XCMS, CAMERA, and metaX toolbox, all of which were implemented in the R software environment. Normalization of raw protein intensity values was performed using the "medium" method to ensure data consistency. Hierarchical clustering analysis was carried out using the pheatmap package to visualize patterns in the data. Partial least squares discriminant analysis was performed using the ropls package in R, which enabled the calculation of variable importance in projection (VIP) values for each variable.

Microbiome and metabolomics analysis

Differential abundance analysis was performed using the Wilcoxon test (*P* < 0.05) and Linear Discriminant Analysis Effect Size (LEfSe) with an LDA score threshold of ≥ 2.0. Alpha and beta diversity were assessed using the "vegan" and "ineq" packages in R, with Bray-Curtis distance used for beta diversity. Principal Coordinate Analysis (PCoA) was conducted using the "ade4" package to visualize microbial community composition differences. Random Forest models were constructed to identify discriminatory features, with Mean Decrease Gini values calculated for each microbe and metabolite. Biomarkers were selected based on discriminative ability and non-redundancy. Multiple 10-fold cross-validation was performed to assess the robustness and generalizability of

the model [20]. Receiver Operating Characteristic (ROC) curves were generated using the "pRoC" package. The area under the ROC curve (AUC) was calculated to evaluate biomarker performance.

The Data Integration Analysis for Biomarker discovery using Latent cOmponents (DIABLO) framework was applied to integrate microbiome and metabolomic data using the "mixOmics" R package (V6.16.3) [21]. Integration was achieved by identifying shared patterns between the datasets, and the results were validated through permutation tests to assess robustness and significance. A chord diagram was used to visualize the correlations greater than 0.5 between microbiota and metabolites.

Functional prediction using PICRUSt2

Functional profiling of the gut microbiota was performed using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2 (PICRUSt2), with functional predictions based on 16 S ribosomal RNA gene sequencing data from the Kyoto Encyclopedia of Genes and Genomes (KEGG) Level 3 and Clusters of Orthologous Groups (COG) databases [22]. ASV and their relative abundances were used as input to predict hierarchical gene families and pathway abundances, with correction for 16 S rRNA gene copy number QIIME 2variations. Differential analysis was performed using the Wilcoxon rank-sum test, and the top 30 functions with $p < 0.05$ were visualized.

Statistical analysis

All statistical analyses were conducted using R version 4.2.0. The Wilcoxon rank-sum test was used to assess differences in species and metabolic pathways between the two groups. Spearman rank correlation was employed for correlation analysis. A p value < 0.05 was considered statistically significant.

Results

Establishment of the *Pkhd1*$^{del3-4/del3-4}$ mouse model

The *Pkhd1* gene exons 3–4 were deleted using CRISPR/Cas9 technology to generate *Pkhd1*$^{del3-4/del3-4}$ mice (Fig. 1A). Knockout of the *Pkhd1* gene was confirmed by tail genotyping and RT-qPCR analysis of liver tissue (Fig. 1B, C). The FPC protein, encoded by PKHD1, localizes to the apical cilia of biliary epithelial cells. Defects in FPC result in ciliary structural abnormalities. To elucidate ciliary alterations, we performed transmission electron microscopy (Fig. 1D). Wild-type mice exhibited well-defined, regularly arranged cilia. In contrast, knockout (KO) mice displayed significantly reduced numbers of short, disorganized cilia. Heterozygous mice, however, showed relatively normal ciliary morphology and abundance. Ultrasonography revealed heterogeneous echotexture and cystic lesions in the mouse liver (Fig. 1E). Histopathological examination demonstrated a progressive increase in both the number and size of biliary cysts with age, accompanied by an elevated number of renal cysts. At 6 months of age, the mice exhibited pronounced hepatic fibrosis, as evidenced by liver pathology (Fig. 2). These findings confirm the successful establishment of the CHF mouse model. Therefore, 6-month-old mice was further selected for further study, and fecal samples were collected for 16 S rDNA gene sequencing and fecal metabolomic analysis.

Microbial composition and diversity in *Pkhd1*$^{del3-4/del3-4}$ mice

Following DADA2 denoising, non-redundant sequences were identified as amplicon sequence variants (ASVs) or feature sequences [23]. A total of 2651 ASVs were detected, with 948 ASVs were shared between wild-type and knockout mice, 611 ASVs were unique to WT mice and 1,092 were unique to the *Pkhd1*$^{del3-4/del3-4}$ group (Fig. 3A). The rarefaction curves indicated that the richness of ASVs increased with the number of samples and eventually plateaued, suggesting that the sample size was sufficient (Fig. 3B).

We employed the Chao1 index, Shannon index, observed species (obs) index, and Simpson index to evaluated differences in microbial richness and diversity between the two groups (Fig. 3C-F). No significant differences were observed. PCoA was performed to visualize dissimilarities in gut microbial community composition by reducing dimensionality (Fig. 3G). Using Bray-Curtis distances to calculate inter-sample microbial dissimilarities, PCoA revealed no significant differences along the first principal coordinate, but exhibited greater heterogeneity along the second principal coordinate (Fig. 3H, I).

Intestinal microbial differences in the *Pkhd1*$^{del3-4/del3-4}$ group

In this study, we focused on the alteration in gut microbiota at both phylum and genus levels. A Venn diagram revealed that 21 phyla were shared between the two groups, with *Firmicutes* being the most abundant phylum (Fig. 4A, B). To compare the differences in microbial composition at these taxonomic levels, we employed the Wilcoxon rank-sum test to analyze the significance of differences (Fig. 4C). We found significant differences in the abundance of *Deferribacterota* and *Patescibacteria* between the groups. Specifically, *Deferribacterota* was significantly enriched in the *Pkhd1*$^{del3-4/del3-4}$ group, while *Patescibacteria* was significantly reduced in this group.

We next compared the differences at the genus level (Fig. 4D). A bar chart depicted the composition of the top 30 genera across both groups (Fig. 4E). Comparative analysis identified significant differences in 18 genera

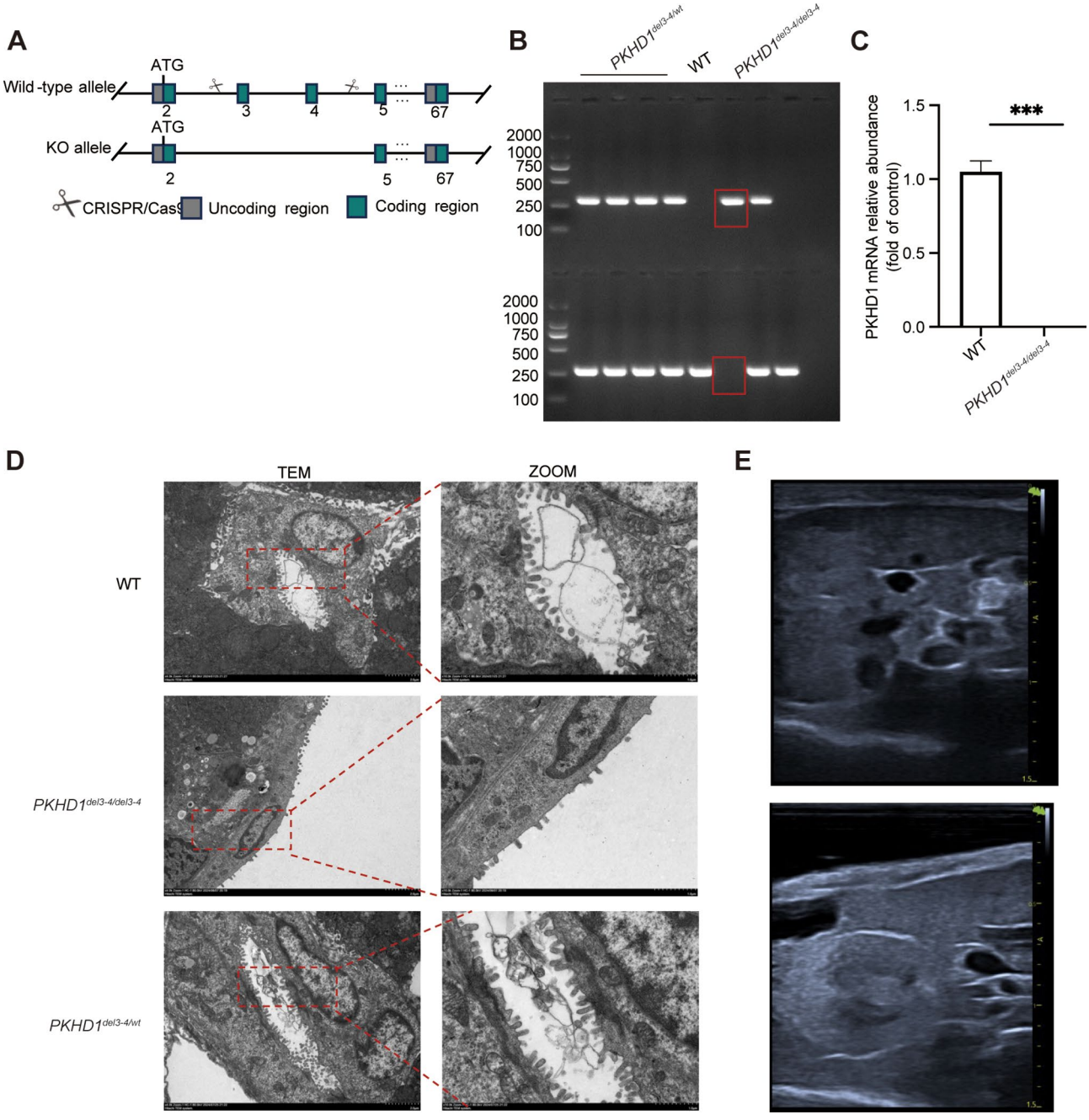

Fig. 1 Genetic and phenotypic analysis of *Pkhd1*$^{del3–4/del3–4}$ mice. (**A**) Schematic representation of the *Pkhd1* gene structure in wild-type and knockout alleles. The KO allele illustrates the CRISPR/Cas9-mediated deletion of exons 3–4 in the coding region. (**B**) *Pkhd1*$^{del3–4/del3–4}$ mice were confirmed by PCR analysis of tail DNA. (**C**) RT-qPCR analysis of *Pkhd1* mRNA relative abundance. (**D**) Electron microscopy images of intrahepatic cholangiocytes from WT, KO, and heterozygous mice. (**E**) Ultrasound imaging of the liver from KO mice. ***$P < 0.0001$

between the two groups, with 13 genera showing marked enrichment in the *Pkhd1*$^{del3–4/del3–4}$ group (Fig. 4F). Among these, *Mucispirillum*, *Eisenbergiella*, and *Oscillibacter* were the most abundant. In contrast, the genera that were significantly reduced in the *Pkhd1*$^{del3–4/del3–4}$ group included *Parvibacter*, *Streptococcus*, and *Candidatus_Saccharimonas*. The results of the LefSe analysis were corroborated these observations, highlighting notable shifts in gut microbiota composition in CHF mice (Fig. 4G, H).

To further elucidate microbial community dynamics, we conducted a Spearman correlation analysis on the top 30 genera (Fig. 5A, B). The results showed significant positive correlations among *Mucispirillum*, *Eisenbergiella*, and *Oscillibacter*, of which were significantly enriched in the *Pkhd1*$^{del3–4/del3–4}$ group.

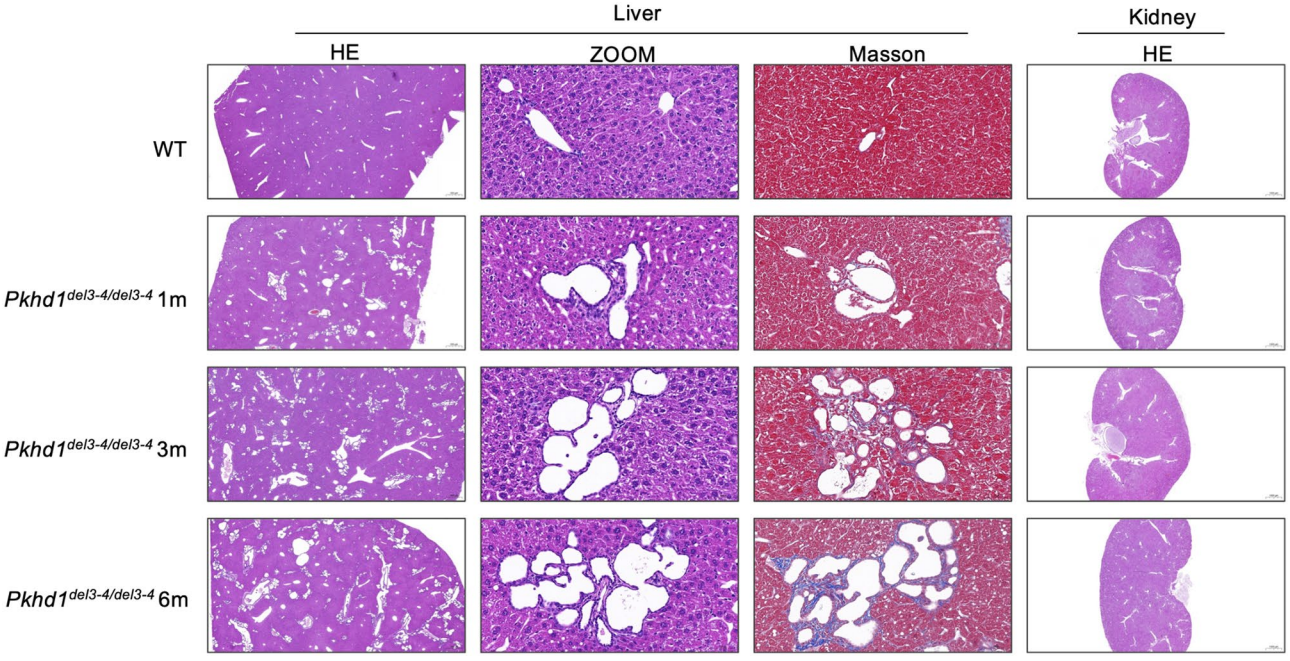

Fig. 2 Histopathological analysis of liver and kidney in *Pkhd1*$^{del3–4/del3–4}$ mice and WT mice. HE staining, zoomed HE images, and Masson's trichrome staining of liver and HE staining of kidney sections from WT and *Pkhd1*$^{del3–4/del3–4}$ mice at 1, 3, and 6 months (m) of age are shown. With increasing age, the number and size of hepatic cysts in *Pkhd1*$^{del3–4/del3–4}$groups progressively enlarged, accompanied by aggravated periductal fibrosis

Functional prediction of intestinal microbiota in *Pkhd1*$^{del3–4/del3–4}$ Mice

The 16 S rDNA sequencing data of gut microbiota were analyzed via PICRUSt2 for functional prediction based on the KEGG and Clusters of Orthologous Group (COG) database. The top 11 KEGG pathways with the most significant differences in the prediction results were shown. Among these, three pathways were downregulated in the *Pkhd1*$^{del3–4/del3–4}$ group: ethylbenzene degradation, caprolactam degradation, and butyrate metabolism. In contrast, the pathways upregulated in the *Pkhd1*$^{del3–4/del3–4}$ group were primarily associated with disease development, including colorectal cancer, influenza A, Parkinson's disease, and small cell lung cancer (Fig. 5C). Notable variations were observed in functions related to DNA metabolism and FHIP-associated peptidases between the groups (Fig. 5D).

Metabolomic characteristics of *Pkhd1*$^{del3–4/del3–4}$ mice

Untargeted metabolomic analysis was performed using MS. Group-wise Partial Least Squares Discriminant Analysis was conducted to visualize the separation trend between the groups (Fig. 6A, B). The results showed a clear separation between the WT and *Pkhd1*$^{del3–4/del3–4}$ groups. The volcano plot displayed the differential metabolites (Fig. 6C). A total of 145 differential metabolites were identified, with 99 metabolites significantly upregulated in the *Pkhd1*$^{del3–4/del3–4}$ group, indicating an active metabolic state in these mice. Gene set enrichment analysis (GSEA) revealed significant differences in multiple metabolic pathways, signal transduction pathways, and disease-related pathways (Fig. 6D). The upregulated genes in the *Pkhd1*$^{del3–4/del3–4}$ group were primarily enriched in taurine and hypotaurine metabolism, biosynthesis of cofactors, and pantothenate and coa biosynthesis. Notably, given that the disease originates from bile duct cells, there was an enhancement in the bile acid metabolism pathway.

Integration of multi-omics and development of diagnostic models

The dimensionality reduction analysis revealed distinct clustering patterns between the study groups (Fig. 7A, B). Analysis using DIABLO demonstrated a strong correlation coefficient ($r = 0.93$) between microbial and metabolic profiles, indicating potential biological interactions between these two omics layers (Fig. 7C). The key correlations between them are illustrated in the figure (Fig. 7D). Significant features were identified using the Block Rank algorithm (Fig. 7E). Among the microbiome features, several bacterial taxa emerged as primary contributors, including *Desulfovibrionaceae unclassified*, *Oscillospiraceae unclassified*, and *Lactobacillaceae unclassified*. The metabolomic analysis identified key metabolites, with oxoproline, 3-hydroxy-4-imino-2,5-pyrrolidinedione, and piccolinic acid showing the strongest associations.

To evaluate the diagnostic efficacy of the gut microbiome for CHF, a random forest model was performed to

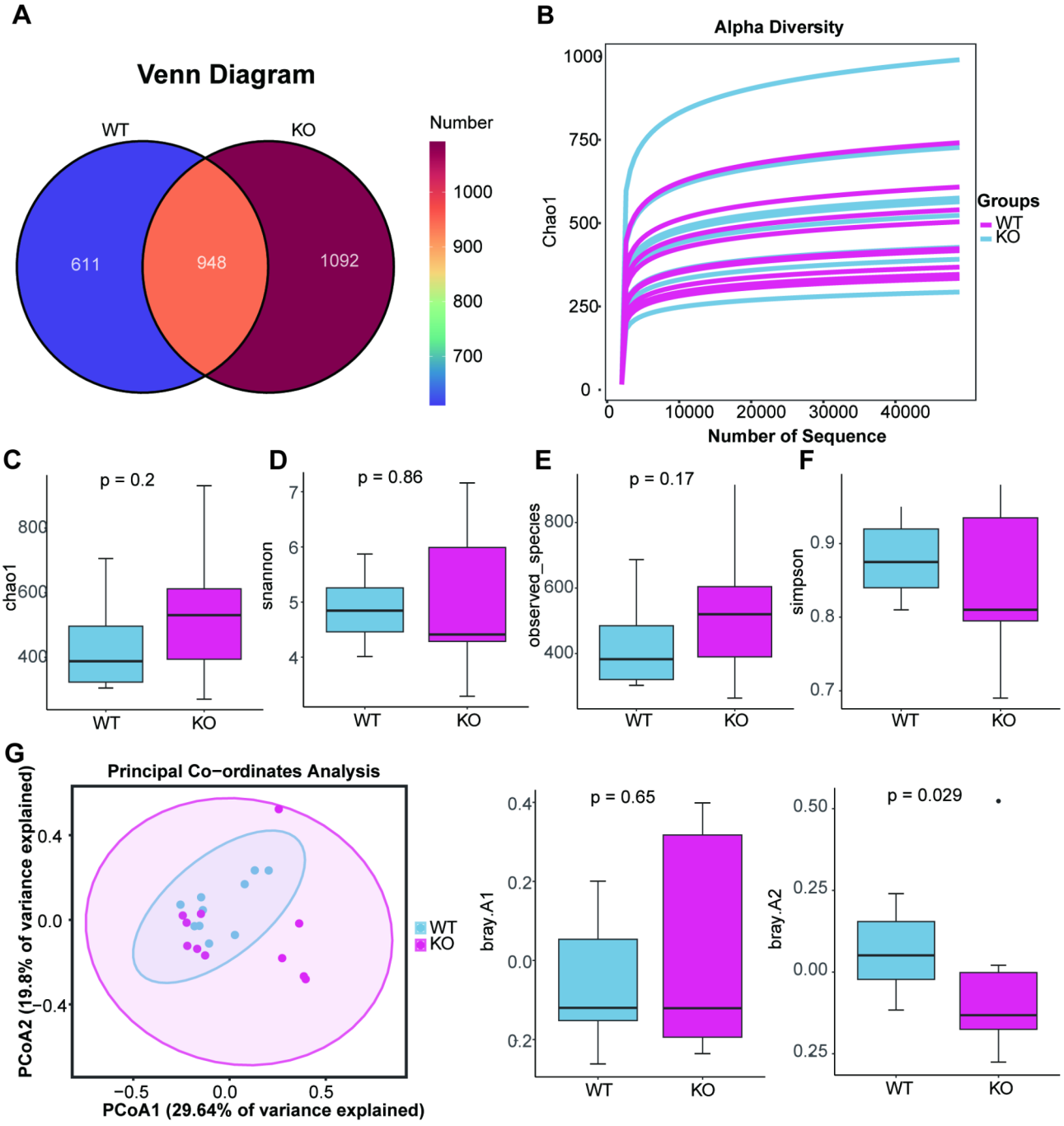

Fig. 3 Analysis of microbial community composition and diversity. (**A**) Venn diagram illustrating the number of unique and shared ASV between two groups. (**B**) Alpha diversity rarefaction curves based on Chao1 index, with each line representing an individual sample. (**C-F**) Box plots comparing alpha diversity indices (Chao1, Shannon, Observed species, and Simpson) between WT and *Pkhd1*$^{del3-4/del3-4}$groups. *P*-values were determined by the Wilcoxon rank-sum test. (**G**) PCoA plot based on Bray-Curtis dissimilarity. The first and second principal coordinates (PCoA1 and PCoA2) explain 29.64% and 19.8% of the variance, respectively. (**H-I**) Boxplots of Bray-Curtis distances to centroids along A1 and A2 components, respectively. Statistical significance was assessed using the Wilcoxon rank-sum test, with *p*-values shown above the plots. WT group is shown in blue, and KO group is shown in pink

identify discriminative features. In this study, we used the Mean Decrease Gini method to evaluate feature importance and selected the top three genera and the top four metabolites as input features for the classifier (genera: *Eubacterium ruminantium group*, *Parvibacter*, and *GCA-900066575*; metabolites: 1-Naphthalenecarboxylic acid, Nonaethylene Glycol, Arg Leu Glu, and Liriodenine). The Mean Decrease Gini values for these genera and metabolites were significantly higher than those of the other variables, indicating their greater contribution

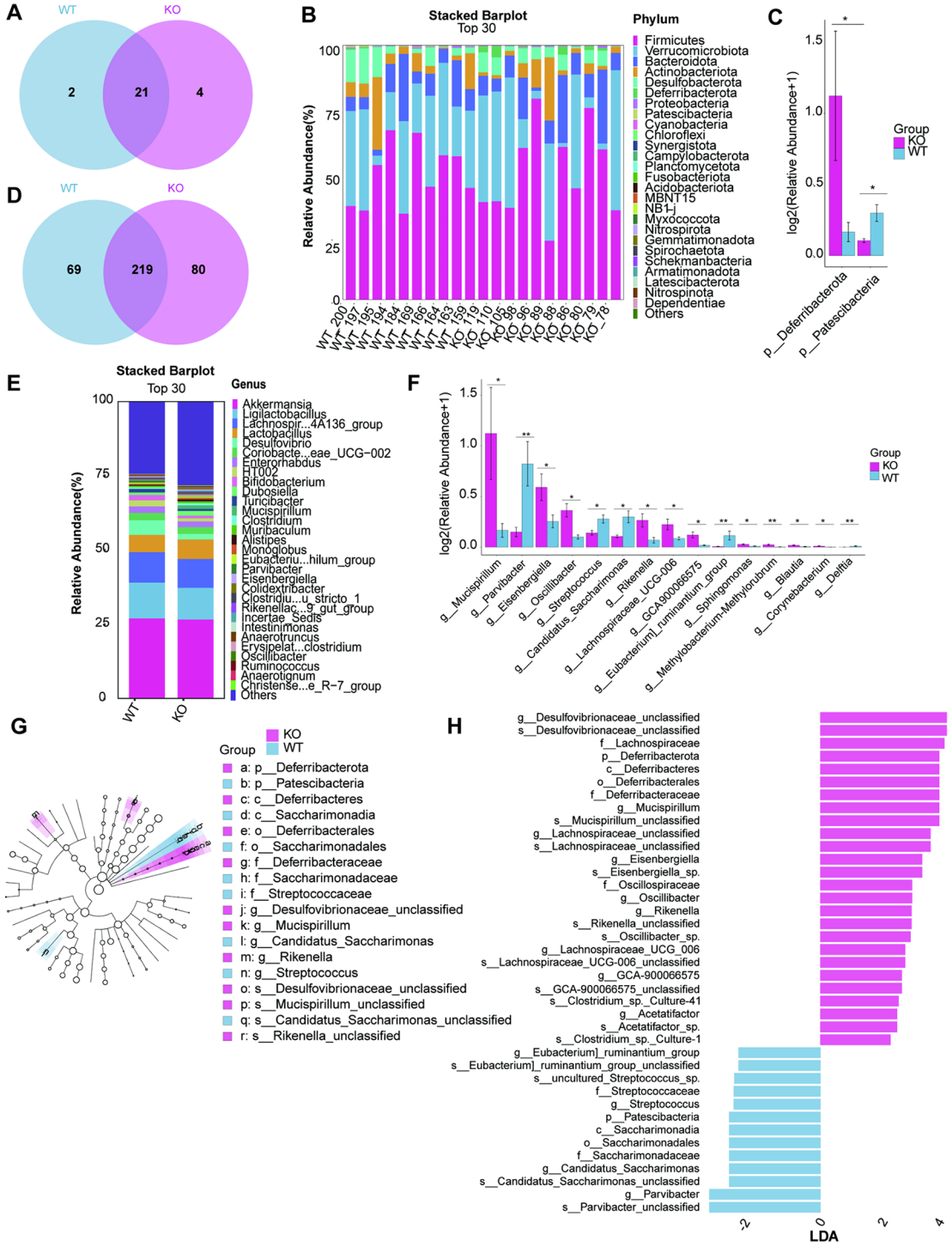

Fig. 4 Comparative analysis of microbial communities. (**A**) Venn diagram showing shared and unique phyla between WT and *Pkhd1*$^{del3-4/del3-4}$groups. (**B**) Stacked bar plot of phyla, illustrating relative abundance differences. (**C**) Bar graph highlighting significant differences in *Firmicutes* and *Bacteroidetes* ratios between groups. (**D**) Venn diagram depicting genus distribution in WT and *Pkhd1*$^{del3-4/del3-4}$ group. (**E**) Stacked bar plot of the top 30 genera, showing relative abundance variations. (**F**) Bar graph indicating significant differences in specific genera. (**G**) Cladogram representing phylogenetic relationships, with significant clades marked. (**H**) The differential abundance of gut microbiota taxa between the KO and WT groups, analyzed using LDA. Taxa enriched in the *Pkhd1*$^{del3-4/del3-4}$ group are represented in pink, while those enriched in the WT group are shown in blue

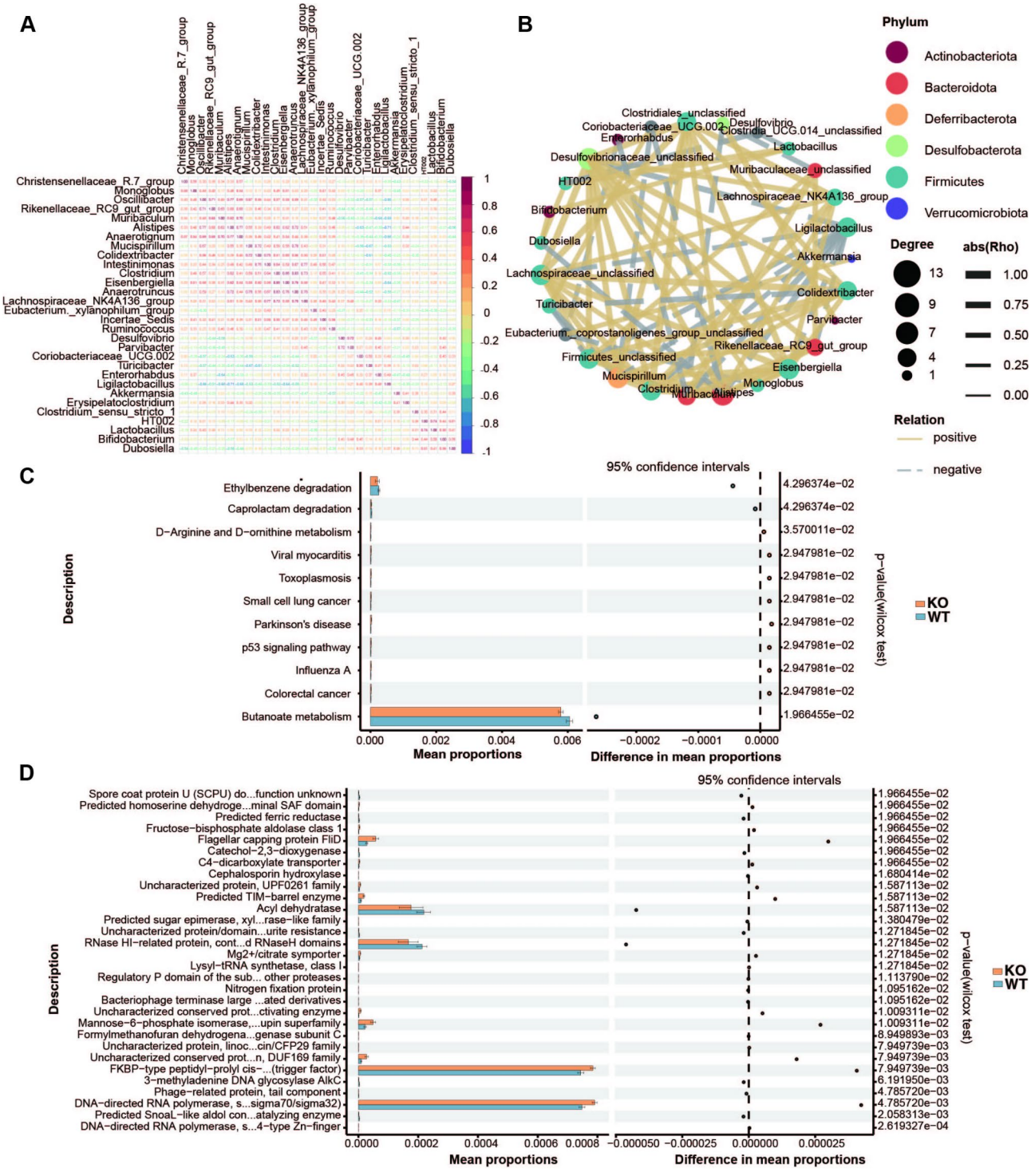

Fig. 5 Microbiome correlation analysis and functional prediction. (**A**) Spearman correlation heatmap of the top 30 genera. The color intensity and size of the circles represent the strength and direction of correlations (red: positive; blue: negative). (**B**) Co-occurrence network of genera, with nodes colored by phylum and sized according to degree. Edge width reflects correlation strength, and edge color indicates positive (yellow) or negative (gray) correlations. (**C**) Differentially abundant microbial taxa between groups at the genus level, visualized using linear discriminant analysis (LDA). Bars represent the mean proportions and differences between groups. (**D**) Functional prediction of microbial communities based on KEGG and COG database. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

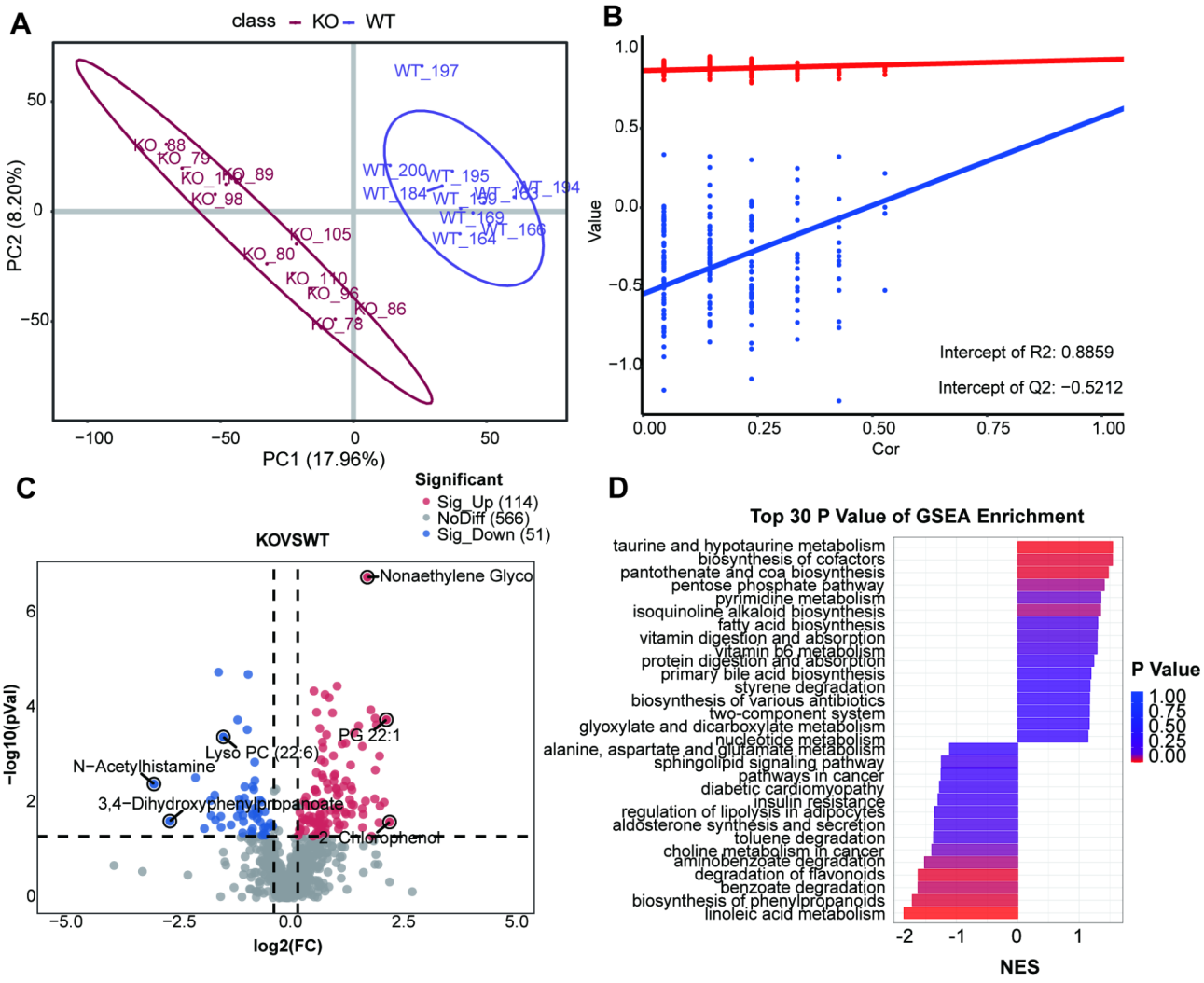

Fig. 6 Non-targeted metabolomics profile of gut microbiota. (**A**) PLS-DA score plot showing separation of WT and *Pkhd1*$^{del3-4/del3-4}$ groups. (**B**) Permutation test. (**C**) A volcano plot was generated for all metabolites from differential expression analysis, with log_2(FC) on the x-axis and -log_{10}(*p*-value) on the y-axis. Red points represent significantly upregulated metabolites, blue points indicate significantly downregulated metabolites, and gray points correspond to metabolites with no significant differential expression. (**D**) Top 30 significantly enriched pathways from GSEA analysis, highlighting up- and down-regulated pathways

to the model's performance (Fig. 8A, B). The metabolite-based model exhibited superior diagnostic performance compared to the microbiome-based model, although both demonstrated robust discriminative capability (microbiome: AUC = 0.883, 95% CI: 0.862–0.905; metabolome: AUC = 0.982, 95% CI: 0.975–0.988) (Fig. 8C, D).

Discussion

Congenital hepatic fibrosis (CHF) is an autosomal recessive disease commonly characterized by mutations in the *PKHD1* gene, which leads to polycystic kidney and liver disease and contribute to the progression of hepatic fibrosis. Due to its rarity, the gut metabolomic and microbiome in CHF has not been extensively studied. The mouse *Pkhd1* gene contains the human T36M mutation, one of the most common mutation sites, and shares 91% homology with the human gene [24]. *Pkhd1*$^{del3-4/del3-4}$ mice, generated using the CRISPR/Cas9 system, display a spectrum of phenotypes mirroring key aspects of the human disease. Notably, these mice spontaneously develop hepatic fibrosis during growth, establishing them as a valuable model for investigating CHF pathogenesis and disease phenotype [25, 26].

Gut microbiota and metabolites have been linked to numerous hepatic disorders, yet their alterations in *PKHD1* mutation-driven CHF remain uncharacterized [27, 28]. To provide a comprehensive characterization of the multi-omics landscape in *PKHD1* mutant mice, we employed advanced machine learning classification models to integrate both metabolomic and microbiome datasets. Given CHF's complexity, involving genetic mutations and hepatic phenotypic changes, an integrative

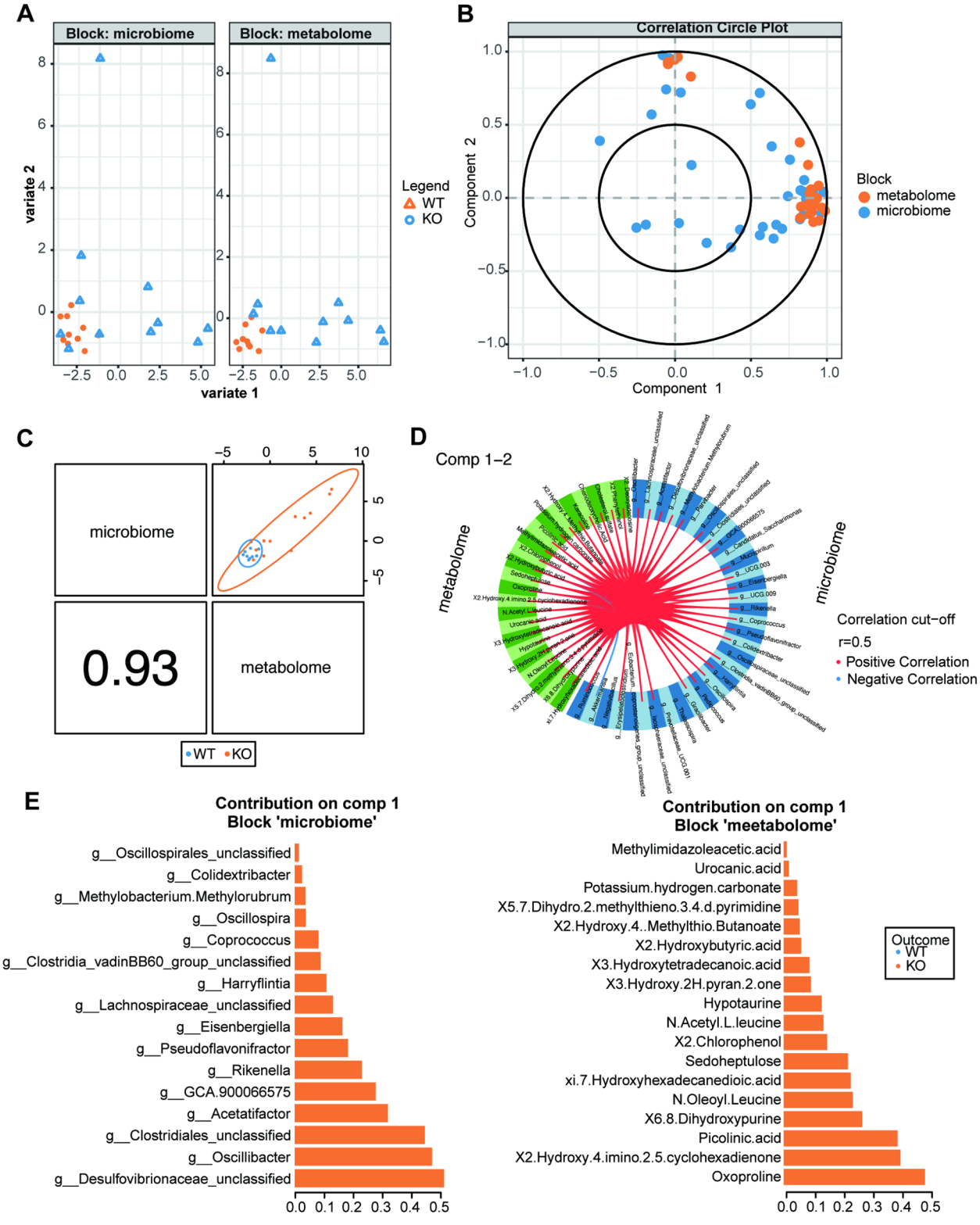

Fig. 7 Multi-omics integration of microbiome and metabolome profiles. (**A**) DIABLO dimension reduction showing variance distribution of microbiome and metabolome blocks in WT and *Pkhd1*$^{del3–4/del3–4}$. (**B**) Sample separation based on multi-omics DIABLO features. (**C**) Multi-omics DIABLO integration demonstrating significant correlations between gut microbiome and metabolome. (**D**) Circular plot demonstrating correlations (cutoff ≥ 0.5) between microbiome and metabolome features. (**F**) Bar plots of top contributing variables to component 1 for microbiome and metabolome, separated by WT and *Pkhd1*$^{del3–4/del3–4}$ groups

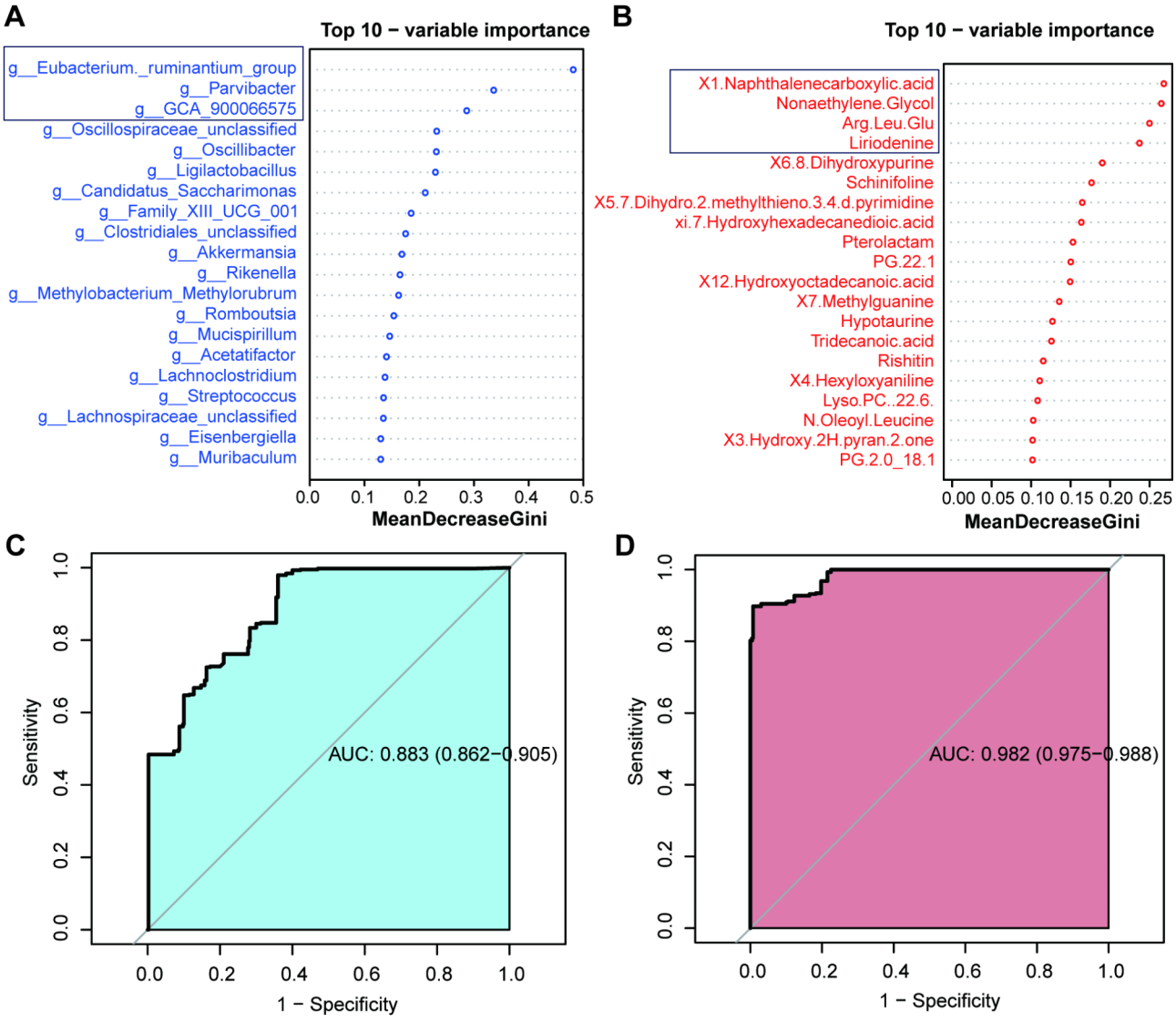

Fig. 8 ROC curves of microbiome and metabolome classifier. (**A**, **B**) Top 20 variables based on the Mean Decrease Gini for microbiome and metabolites, respectively. The top three microbiota and top four metabolites with the highest Mean Decrease Gini values were selected for constructing the diagnostic model. (**C**) ROC of gut microbiome, achieving an AUC of 0.883 (95% CI: 0.862–0.905). (**D**) ROC curve for metabolome, achieving an AUC of 0.982 (95% CI: 0.975–0.988)

multi-omics approach offers a more precise identification of microbial dysbiosis [29]. Using the DIABLO machine learning method, we found a significant positive correlation between the gut microbiome and metabolome, indicating the stability of the multi-omics approach [30]. The independent analysis of the two omics demonstrated strong discriminative power, effectively distinguishing *Pkhd1*$^{del3-4/del3-4}$ mice from WT with high classification accuracy. Notably, metabolomic data outperformed microbiome data in discriminating disease states, suggesting that metabolic alterations may play a more pivotal role in pathogenesis. These findings highlight the complex interplay between gut microbiota and host metabolism, underscoring the potential of metabolomics as robust biomarkers for CHF.

Recent studies have highlighted the critical role of the gut microbiota in liver disease development through the gut-liver axis, with gut dysbiosis and liver disease mutually reinforcing each other [31]. To further elucidate this relationship, we employed a multi-omics approach to identify characteristic microbes and metabolites in CHF mouse. While no significant differences were observed in overall alpha and beta diversity, we did find notable compositional differences at the genus level. Specifically, the genera *Mucispirillum*, *Eisenbergiella*, and *Oscillibacter* not only showed significant differences but also displayed significant positive correlations with each other,

suggesting that these three genera may be core microbiota associated with the progression of CHF. These findings indicate that even in the absence of significant diversity changes, shifts in microbial composition can still play an important role in the CHF development and progression. Among these, the high expression of *Mucispirillum* in the *Pkhd1*$^{del3-4/del3-4}$ group is particularly noteworthy. *Mucispirillum*, a Gram-negative bacillus belonging to the phylum *Deferribacteres*, has been implicated in changes in liver disease. Previous studies have demonstrated that high-fat diets can increase the abundance of *Mucispirillum* in the gut microbiota of mice [32], with one study noting a gradual rise in *Mucispirillum* levels over 14 months of high-fat feeding, coinciding with liver fibrosis development [27]. The concurrent increase of *Mucispirillum* and liver fibrosis may be due to its colonization in the intestinal mucus layer, where its pathogenic traits trigger gut inflammation through type VI secretion system and putative effector proteins [33]. Its proliferation could compromise the gut barrier, allowing inflammatory factors and bacterial products to translocate to the liver via the portal vein, thereby exacerbating liver disease progression [34]. This potential change warrants further investigation.

Notably, our functional analysis of the gut microbiota revealed a significant decrease in the butyrate metabolism pathway in the *Pkhd1*$^{del3-4/del3-4}$ group compared to WT group. Additionally, the butyrate-producing genus Eubacterium was significantly reduced in the *Pkhd-1*$^{del3-4/del3-4}$ group. Butyrate, a key short-chain fatty acid, is primarily produced by butyrate-producing bacteria through the fermentation of dietary fiber [35]. Butyrate is a primary energy source for intestinal epithelial cells, promoting their proliferation and repair [36, 37], while enhancing the expression of mucin to maintain the integrity of the gut barrier [38]. Additionally, butyrate directly attenuates hepatic stellate cell activation and improves liver fibrosis progression [39]. Impaired butyrate metabolism weakens the gut barrier, enabling harmful substances like lipopolysaccharide to enter the bloodstream and reach the liver, triggering inflammation and damage and accelerating the progression of hepatic fibrosis [40]. Therefore, we hypothesize that the development of fibrosis in CHF mice may be associated with decrease in butyrate levels.

Metabolomic profiling revealed that most of the differential metabolites in the *Pkhd1*$^{del3-4/del3-4}$ group were upregulated. Further GSEA enrichment analysis suggested that the upregulated metabolic pathways were associated with nucleotide and amino acid metabolism, indicating an active metabolic state in the *Pkhd-1*$^{del3-4/del3-4}$ group [41]. Our metabolomic analysis revealed alterations in several differential metabolites in the *Pkhd1*$^{del3-4/del3-4}$ group, including the upregulation of Nonaethylene Glycol and 2-Chlorophenol, and the downregulation of N-Acelyhistamine and 3,4-dihydroxyphenylpropanoate. These findings warrant further investigation to elucidate their potential roles in the development of fibrosis in CHF mice.

CHF mice exhibited both genetic mutations and hepatic fibrosis. While prior omics research on hepatic fibrosis has largely centered on acquired forms, it is critical to recognize that host genetics also shape gut microbiota composition through metabolic alterations [42, 43]. In the case of the CHF mice analyzed here, the observed omics changes likely reflect a combined influence of genetic mutations and hepatic fibrosis. Translating these findings back to the human context, it is important to note that CHF is not solely caused by *PKHD1* gene mutations, and the specific loci and types of mutations can vary [44, 45]. Such genetic variations can significantly impact gut microbiota composition, warranting an assessment of their heritability. This could facilitate the identification of microbial species most associated with different types of CHF and help identify the core of the genetic spectrum [46, 47]. Therefore, larger cohorts may be required to differentiate the heritable and fibrosis-related components of the gut microbiota in CHF. Practically, large-scale sequencing of diverse loci combined with multi-omics integration may offer deeper insights into gut microbiota interactions. This comprehensive approach has the potential to enhance the generalizability of diagnostic models and provide a foundation for targeted microbial therapies in genetic mutation-related diseases [48, 49].

Our study elucidates distinct alterations in gut microbiota composition and metabolic profiles in CHF mice model, highlighting specific host-microbiome interactions implicated in CHF pathogenesis. While the mouse model provides valuable insights, we acknowledge species differences. However, the dysbiosis observed, linked to intestinal inflammation and liver disease progression, mirrors findings in human studies, suggesting that gut microbiota disturbances may also play a role in human CHF. Multi-omics approaches hold promise for CHF diagnosis and screening, and could aid in identifying microbial signatures that serve as non-invasive biomarkers for early detection. Furthermore, therapeutic strategies targeting the gut microbiota, such as microbiome modulation or specific antibiotics, warrant exploration in human trials. Further investigation is needed to fully delineate the clinical implications of these observations.

Author contributions

YWL and CLZ designed this research; MFJ contributed to conceptualization, performed the research and drafted the manuscript; MFJ and YS performed the animal experiments; MFJ, ZXD, XXH and XZY analyzed the data; QLC and RL revised the manuscript; All authors read and approved the final manuscript.

Funding
This work was supported by the National Key Research and Development Program of China (2023YFC2308600), the National Natural Science Foundation of China (No.82472167), Science and Technology Plan of Hainan Province (Clinical Research Center) (LCYX202204, LCYX202306 and LCYX202408), and the Natural Science Foundation of Jiangsu Province (BK20241982).

Data availability
The dataset of 16 S rRNA gene sequencing supporting the conclusions of this article is available in the National Center Biotechnology Information repository with accession code PRJNA1197016 (https://www.ncbi.nlm.nih.gov/bioproject/1197016). The metabolomic data have been deposited in the OMIX, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (https://ngdc.cncb.ac.cn/omix: accession no. OMIX008257).

Declarations

Ethics approval and consent to participate
All animal experiments were approved by the Animal Care and Use Committee of Nanjing Medical University and adhered to the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines (reference number: IACUC-2407024).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 29 November 2024 / Accepted: 13 March 2025
Published online: 31 March 2025

References

1. Zhang MZ, Mai W, Li C, Cho SY, Hao C, Moeckel G, Zhao R, Kim I, Wang J, Xiong H, et al. PKHD1 protein encoded by the gene for autosomal recessive polycystic kidney disease associates with basal bodies and primary cilia in renal epithelial cells. Proc Natl Acad Sci U S A. 2004;101(8):2311–6.
2. Menezes LF, Cai Y, Nagasawa Y, Silva AM, Watkins ML, Da Silva AM, Somlo S, Guay-Woodford LM, Germino GG, Onuchic LF. Polyductin, the PKHD1 gene product, comprises isoforms expressed in plasma membrane, primary cilium, and cytoplasm. Kidney Int. 2004;66(4):1345–55.
3. Locatelli L, Cadamuro M, Spirlì C, Fiorotto R, Lecchi S, Morell CM, Popov Y, Scirpo R, De Matteis M, Amenduni M et al. Macrophage recruitment by fibrocystin-defective biliary epithelial cells promotes portal fibrosis in congenital hepatic fibrosis. Hepatol 2016;63(3):965–82.
4. Harris PC, Torres VE. Polycystic kidney disease. Annu Rev Med. 2009;60:321–37.
5. Igarashi P, Somlo S. Genetics and pathogenesis of polycystic kidney disease. J Am Soc Nephrol. 2002;13(9):2384–98.
6. Sweeney WE, Avner ED. Diagnosis and management of childhood polycystic kidney disease. Pediatr Nephrol. 2011;26(5):675–92.
7. Liu R, Dai Z, Zhu C. A rare cause of liver fibrosis in an adolescent woman. Gastroenterology. 2023;164(3):340–3.
8. Zerres K, Rudnik-Schöneborn S, Deget F, Holtkamp U, Brodehl J, Geisert J, Schärer K. Autosomal recessive polycystic kidney disease in 115 children: clinical presentation, course and influence of gender. Arbeitsgemeinschaft für pädiatrische, nephrologie. Acta Paediatr. 1996;85(4):437–45.
9. Gallagher AR, Esquivel EL, Briere TS, Tian X, Mitobe M, Menezes LF, Markowitz GS, Jain D, Onuchic LF, Somlo S. Biliary and pancreatic dysgenesis in mice harboring a mutation in Pkhd1. Am J Pathol. 2008;172(2):417–29.
10. Walker RV, Yao Q, Xu H, Maranto A, Swaney KF, Ramachandran S, Li R, Cassina L, Polster BM, Outeda P, et al. Fibrocystin/Polyductin releases a C-terminal fragment that translocates into mitochondria and suppresses cystogenesis. Nat Commun. 2023;14(1):6513.
11. Hsu CL, Schnabl B. The gut-liver axis and gut microbiota in health and liver disease. Nat Rev Microbiol. 2023;21(11):719–33.
12. Tripathi A, Debelius J, Brenner DA, Karin M, Loomba R, Schnabl B, Knight R. The gut-liver axis and the intersection with the Microbiome. Nat Rev Gastroenterol Hepatol. 2018;15(7):397–411.
13. Lee NY, Suk KT. The role of the gut Microbiome in liver cirrhosis treatment. Int J Mol Sci 2020;22(1).
14. Chopyk DM, Grakoui A. Contribution of the intestinal Microbiome and gut barrier to hepatic disorders. Gastroenterology. 2020;159(3):849–63.
15. Qin N, Yang F, Li A, Prifti E, Chen Y, Shao L, Guo J, Le Chatelier E, Yao J, Wu L, et al. Alterations of the human gut Microbiome in liver cirrhosis. Nature. 2014;513(7516):59–64.
16. Oh TG, Kim SM, Caussy C, Fu T, Guo J, Bassirian S, Singh S, Madamba EV, Bettencourt R, Richards L et al. A universal Gut-Microbiome-Derived signature predicts cirrhosis. Cell Metab 2020;32(5):878–e888876.
17. Ning L, Zhou YL, Sun H, Zhang Y, Shen C, Wang Z, Xuan B, Zhao Y, Ma Y, Yan Y, et al. Microbiome and metabolome features in inflammatory bowel disease via multi-omics integration analyses across cohorts. Nat Commun. 2023;14(1):7135.
18. Gao X, Sun R, Jiao N, Liang X, Li G, Gao H, Wu X, Yang M, Chen C, Sun X, et al. Integrative multi-omics Deciphers the Spatial characteristics of host-gut microbiota interactions in Crohn's disease. Cell Rep Med. 2023;4(6):101050.
19. Babu M, Snyder M. Multi-Omics profiling for health. Mol Cell Proteom. 2023;22(6):100561.
20. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol. 2008;26(8):1364–70.
21. Rohart F, Gautier B, Singh A, KA LC. MixOmics: an R package for 'omics feature selection and multiple data integration. PLoS Comput Biol. 2017;13(11):e1005752.
22. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol. 2020;38(6):685–8.
23. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from illumina amplicon data. Nat Methods. 2016;13(7):581–3.
24. Nagasawa Y, Matthiesen S, Onuchic LF, Hou X, Bergmann C, Esquivel E, Senderek J, Ren Z, Zeltner R, Furu L et al. Identification and characterization of Pkhd1, the mouse orthologue of the human ARPKD gene. J Am Soc Nephrol 2002;13(9):2246–58.
25. Garcia-Gonzalez MA, Menezes LF, Piontek KB, Kaimori J, Huso DL, Watnick T, Onuchic LF, Guay-Woodford LM, Germino GG. Genetic interaction studies link autosomal dominant and recessive polycystic kidney disease in a common pathway. Hum Mol Genet. 2007;16(16):1940–50.
26. Yang C, Harafuji N, Caldovic L, Yu W, Boddu R, Bhattacharya S, Barseghyan H, Gordish-Dressman H, Foreman O, Bebok Z, et al. Pkhd1(cyli/cyli) mice have altered renal Pkhd1 mRNA processing and hormonally sensitive liver disease. J Mol Med (Berl). 2023;101(9):1141–51.
27. Zhang X, Coker OO, Chu ES, Fu K, Lau HCH, Wang YX, Chan AWH, Wei H, Yang X, Sung JJY, et al. Dietary cholesterol drives fatty liver-associated liver cancer by modulating gut microbiota and metabolites. Gut. 2021;70(4):761–74.
28. Shen Y, Wu SD, Chen Y, Li XY, Zhu Q, Nakayama K, Zhang WQ, Weng CZ, Zhang J, Wang HK, et al. Alterations in gut Microbiome and metabolomics in chronic hepatitis B infection-associated liver disease and their impact on peripheral immune response. Gut Microbes. 2023;15(1):2155018.
29. Eddy S, Mariani LH, Kretzler M. Integrated multi-omics approaches to improve classification of chronic kidney disease. Nat Rev Nephrol 2020;16(11):657–68.
30. Singh A, Shannon CP, Gautier B, Rohart F, Vacher M, Tebbutt SJ, LC KA. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. Bioinformatics. 2019;35(17):3055–62.
31. Pabst O, Hornef MW, Schaap FG, Cerovic V, Clavel T, Bruns T. Gut-liver axis: barriers and functional circuits. Nat Rev Gastroenterol Hepatol. 2023;20(7):447–61.
32. Zhang XL, Chen L, Yang J, Zhao SS, Jin S, Ao N, Yang J, Liu HX, Du J. Vitamin D alleviates non-alcoholic fatty liver disease via restoring gut microbiota and metabolism. Front Microbiol. 2023;14:1117644.
33. Caruso R, Mathes T, Martens EC, Kamada N, Nusrat A, Inohara N, Núñez G. A specific gene-microbe interaction drives the development of Crohn's disease-like colitis in mice. Sci Immunol 2019;4(34).
34. Fitzgerald BG. MT Sorbara Fiber-deficient diets reprogram the microbiota. Cell Host Microbe. 2023;31(12):1950–1.
35. Recharla N, Geesala R, Shi XZ. Gut Microbial Metabolite Butyrate and Its Therapeutic Role in Inflammatory Bowel Disease: A Literature Review. Nutrients 2023:15(10).

36. Tang G, Du Y, Guan H, Jia J, Zhu N, Shi Y, Rong S, Yuan W. Butyrate ameliorates skeletal muscle atrophy in diabetic nephropathy by enhancing gut barrier function and FFA2-mediated PI3K/Akt/mTOR signals. Br J Pharmacol. 2022;179(1):159–78.
37. Louis P, Flint HJ. Formation of propionate and butyrate by the human colonic microbiota. Environ Microbiol. 2017;19(1):29–41.
38. Burger-van Paassen N, Vincent A, Puiman PJ, van der Sluis M, Bouma J, Boehm G, van Goudoever JB, van Seuningen I, Renes IB. The regulation of intestinal mucin MUC2 expression by short-chain fatty acids: implications for epithelial protection. Biochem J. 2009;420(2):211–9.
39. Li W, Deng M, Gong J, Zhang X, Ge S, Zhao L. Sodium acetate inhibit TGF-β1-Induced activation of hepatic stellate cells by restoring AMPK or c-Jun signaling. Front Nutr. 2021;8:729583.
40. Wang R, Yang X, Liu J, Zhong F, Zhang C, Chen Y, Sun T, Ji C, Ma D. Gut microbiota regulates acute myeloid leukaemia via alteration of intestinal barrier function mediated by butyrate. Nat Commun. 2022;13(1):2522.
41. Mullen NJ, Singh PK. Nucleotide metabolism: a pan-cancer metabolic dependency. Nat Rev Cancer. 2023;23(5):275–94.
42. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, Beaumont M, Van Treuren W, Knight R, Bell JT, et al. Human genetics shape the gut Microbiome. Cell. 2014;159(4):789–99.
43. Chen L, Zhernakova DV, Kurilshikov A, Andreu-Sánchez S, Wang D, Augustijn HE, Vich Vila A, Weersma RK, Medema MH, Netea MG, et al. Influence of the microbiome, diet and genetics on inter-individual variation in the human plasma metabolome. Nat Med. 2022;28(11):2333–43.
44. Goggolidou P, Richards T. The genetics of autosomal recessive polycystic kidney disease (ARPKD). Biochim Biophys Acta Mol Basis Dis. 2022;1868(4):166348.
45. Bergmann C, Senderek J, Windelen E, Küpper F, Middeldorf I, Schneider F, Dornia C, Rudnik-Schöneborn S, Konrad M, Schmitt CP, et al. Clinical consequences of PKHD1 mutations in 164 patients with autosomal-recessive polycystic kidney disease (ARPKD). Kidney Int. 2005;67(3):829–48.
46. Gunay-Aygun M, Tuchman M, Font-Montgomery E, Lukose L, Edwards H, Garcia A, Ausavarat S, Ziegler SG, Piwnica-Worms K, Bryant J, et al. PKHD1 sequence variations in 78 children and adults with autosomal recessive polycystic kidney disease and congenital hepatic fibrosis. Mol Genet Metab. 2010;99(2):160–73.
47. Bevins CL, Salzman NH. The Potter's wheel: the host's role in sculpting its microbiota. Cell Mol Life Sci. 2011;68(22):3675–85.
48. Borody TJ, Khoruts A. Fecal microbiota transplantation and emerging applications. Nat Rev Gastroenterol Hepatol 2011;9(2):88–96.
49. Xiang Y, Zhang C, Wang J, Cheng Y, Wang L, Tong Y, Yan D. Identification of host gene-microbiome associations in colorectal cancer patients using Mendelian randomization. J Transl Med. 2023;21(1):535.

Publisher's note