



Identification of Diagnostic Biomarkers for Compensatory Liver Cirrhosis Based on Gut Microbiota and Urine Metabolomics Analyses

Yingjun Chen¹ · Shaoxian Chen¹ · Chandi Xu¹ · Li Yu¹ · Shanshan Chu¹ · Jianzhi Bao¹ · Jinwei Wang² · Junwei Wang¹

Received: 6 July 2023 / Accepted: 26 September 2023
© The Author(s) 2023

Abstract

Liver cirrhosis is one of the most prevalent chronic liver disorders with high mortality. We aimed to explore changed gut microbiome and urine metabolome in compensatory liver cirrhosis (CLC) patients, thus providing novel diagnostic biomarkers for CLC. Forty fecal samples from healthy volunteers (control: 19) and CLC patients (patient: 21) were undertaken 16S rDNA sequencing. Chromatography-mass spectrometry was performed on 40 urine samples (20 controls and 20 patients). Microbiome and metabolome data were separately analyzed using corresponding bioinformatics approaches. The diagnostic model was constructed using the least absolute shrinkage and selection operator regression. The optimal diagnostic model was determined by five-fold cross-validation. Pearson correlation analysis was applied to clarify the relations among the diagnostic markers. 16S rDNA sequencing analyses showed changed overall alpha diversity and beta diversity in patient samples compared with those of controls. Similarly, we identified 841 changed metabolites. Pathway analysis revealed that the differential metabolites were mainly associated with pathways, such as tryptophan metabolism, purine metabolism, and steroid hormone biosynthesis. A 9-maker diagnostic model for CLC was determined, including 7 microorganisms and 2 metabolites. In this model, there were multiple correlations between microorganisms and metabolites. *Subdoligranulum*, *Agathobacter*, *norank_f_Eubacterium_coprostanoligenes_group*, *Butyricoccus*, *Lachnospiraceae_UCG_004*, and L-2,3-Dihydrodipicolinate were elevated in CLC patients, whereas *Blautia*, *Monoglobus*, and 5-Acetamidovalerate were reduced. A novel diagnostic model for CLC was constructed and verified to be reliable, which provides new strategies for the diagnosis and treatment of CLC.

Keywords Compensatory liver cirrhosis · Gut microbiota · Urine metabolomics · Early diagnosis · Inflammation

Introduction

Liver cirrhosis (LC) is one of the most frequent chronic liver diseases worldwide and a dominant cause of death, accounting for 2% of global mortalities yearly [1, 2]. The main characteristics of LC are altered liver architecture induced by regenerative nodules and diffuse fibrosis,

resulting in intrahepatic vascular change, portal hypertension, and secondary liver failure [3]. Furthermore, LC is a risk factor for hepatocellular carcinoma, and the morbidity of hepatocellular carcinoma in LC patients is 2–7% [4]. The progression of LC can be split into the compensatory (asymptomatic) and decompensated stages [5]. Asymptomatic LC develops into the decompensated stage with the common clinical manifestations of ascites, hemorrhage, encephalopathy, and jaundice, conducting to decreased life quality and high mortality [6]. In China, approximately 3% of cases of compensatory LC (CLC) progress into decompensated LC annually and a 5-year death rate of 85% is estimated in patients with decompensated LC [7]. Despite liver transplantation can enhance survival rates, its clinical application is limited by donor shortage, high expenses, and severe post-transplantation complications [7]. Currently, liver biopsy is the major method available for clinical diagnosis of LC; however, it is not appropriate

Yingjun Chen, Shaoxian Chen were co-first authors and contributed equally to this work.

✉ Junwei Wang
901086@163.com

¹ Department of Infectious Diseases, Tiantai People's Hospital of Zhejiang Province, Taizhou 317200, People's Republic of China

² Department of General Medicine, Tiantai People's Hospital of Zhejiang Province, Taizhou 317200, People's Republic of China

for all patients and 2–3% of patients may undergo multiple complications [8]. Early diagnosis and timely treatment are crucial for improving the outcomes of patients with LC. Therefore, there is an urgent need for the discovery of novel diagnostic markers for early-stage LC.

Increasing studies indicate that gut dysbiosis plays a vital role in the occurrence and development of LC [5, 9]. The gut microbiota is defined as the complicated aggregation of microbes in the gut, including bacteria, archaea, fungi, and viruses [10]. The gut microbiota produces a mass of metabolites that facilitate the interaction between the gut microbiota and the host [11]. Normal microbial composition acts as a natural defense barrier, which has significant implications for the acquisition of nutrients and regulation of the immune system and metabolic ability on the host [12]. The liver is a highly active region of metabolism and immune homeostasis, which impacts the microbiota by secreting and modulating various immunogenic molecules and metabolites into the gut [13]. The abnormalities in the gut microbiota in LC patients are marked by the elevated abundance of latently pathogenic bacteria and descended amounts of advantageous bacteria [14]. Previous research observed a rise in *Streptococcaceae*, *Peptostreptococcaceae*, *Erysipelotrichaceae*, *Clostridiaceae*, and *Pasteurellaceae* and a reduction in *Acidaminococcaceae*, *Porphyromonadaceae*, *Prevotellaceae*, and *Bacteroidaceae* in the fecal samples from LC patients [14]. This suggests that alterations in the gut microbiota might serve as diagnostic signatures of LC, prompting us to investigate the status of the gut microbiota in CLC.

Microbiome analysis combined with metabolomics has been extensively applied as an effective method to figure out the correlation between health outcomes and microbiome [15]. Metabolomics is an assessment of multi-parametric metabolic reactions of multi-cellular systems, designed to determine and quantify different small molecules in multiple biological samples or specific physiological conditions [16]. As an emerging technique, metabolomics can not only be used to identify specific diagnostic biomarkers but also to clarify molecular mechanisms underlying specific pathology [17]. A previous study indicated that disordered gut microbiome conducted to LC progression, which might be associated with changed microbiome–metabolite interactions [8]. Hence, the integration of gut microbiota and metabolomics analyses is expected to obtain more efficient diagnosis biomarkers for CLC.

In this study, 16S ribosomal DNA (rDNA) sequencing and untargeted metabolomics were performed on feces and urine samples, respectively. The findings of our study revealed the differences in the gut microbiome and urine metabolites between CLC patients and normal controls. Besides, we constructed a diagnostic model for CLC by integrating gut microbiome with metabolomics analyses and verified the model. This study may provide novel early

diagnostic biomarkers for LC and furnish a reference for further research into the pathogenesis of LC.

Materials and Methods

Human Subjects and Clinical Samples

Forty-one human individuals aged 40–50 were involved in the study, including 20 healthy volunteers (control) and 21 patients with CLC. These human individuals were enrolled in Tiantai People's Hospital of Zhejiang Province (from 2018 to 2021). Patients were diagnosed with LC in terms of medical history, liver biopsy, clinical symptoms, laboratory tests, imaging tests, histological examinations, and complications. The diagnosis for CLC was based on the criteria described in Chinese guidelines on the management of liver cirrhosis (abbreviated version) [18]. No participants suffered hypertension, diabetes, inflammatory bowel disease, or necrotizing enteritis. Besides, none of these participants took proton pump inhibitors, antibiotics, or probiotics within 2 weeks before sample collection. Feces and urine samples were collected in the morning after an overnight fast, then delivered to the laboratory within 2 h on dry ice, and stored at -80°C until further analyses. We performed the sample collection following the principle of voluntariness. Finally, 40 appropriate feces samples (control: 19; patient: 21) and urine samples (control: 20; patient: 20) were included for subsequent gut microbiota and urine metabolomics analyses, respectively. The sample sizes for these analyses conformed to a previous study [19]. The demographic and clinical data of subjects involved in this study are presented in Table S1. All participants signed written informed consent, and the experimental procedures were authorized by the Ethics Committee of Tiantai People's Hospital of Zhejiang Province (TYLL2019-06). This study was in line with the Declaration of Helsinki.

16S rDNA Sequencing

According to the manufacturer's instructions, the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) was used for DNA extraction from fecal samples. NanoDrop2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) was utilized to measure the concentration and purity of the DNA, and DNA quality was evaluated through 1% agarose gel electrophoresis. Then, the V3–V4 regions of bacterial 16S rDNA were amplified using the universal primers 338F, 5'-ACT CCT ACG GGA GGC AGC A-3' and 806R, 5'-GGA CTA CHV GGG TWT CTA AT-3'. Following PCR amplification, the products were obtained by 2% agarose gel. The purified DNA fragments were then eluted using Tris–HCl and assessed using 2% agarose gel

electrophoresis. The following conditions for amplification were applied: initial denaturation at 95 °C for 3 min, 30 cycles of denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s, and primer elongation at 72 °C for 45 s. All amplicons were purified using Agencourt AMPure XP (Beckman Coulter, USA), which were quantified using Qubit dsDNA HS Assay Kit and Qubit 3.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). TruSeq DNA PCR-Free Sample Preparation Kit (Illumina, USA) was used for sequencing library generation and index code addition. The library construction process involved the following steps: (1) ligation of the “Y”-shaped adapter, (2) removal of self-ligated fragments using magnetic beads, (3) enrichment of the library through PCR amplification, and (4) denaturation of DNA using sodium hydroxide. Agilent 4200 TapeStation (Agilent Technologies, Palo Alto, CA, USA) was applied to determine library quality. Finally, the purified amplification fragment was incorporated into a library of PE 2 × 300 for sequencing on the HiSeq 2500 platform (Illumina, San Diego, CA, USA) following the manufacturer’s directions. Only high-quality reads were retained by removing all terminal bases with low quality (< Q20) using Trimmomatic (version 0.35). The sequence data related to this study have been deposited in the NCBI Short Read Archive (SRA) database (accession number: PRJNA1019460).

Bioinformatics Analyses of 16S rDNA Sequencing

Ultra-fast FASTQ preprocessor fastp (version 0.19.4, <https://github.com/OpenGene/fastp>) [20] was used for quality control of raw sequencing data. For quality control, reads with a length of 300 bp were truncated at sites where the average quality score across a sliding window of 50 bp fell below 20 and truncated reads with a length of less than 50 bp were removed. Besides, the original sequences were compared using Burrows–Wheeler Aligner (BWA) software [21], and the contaminated reads with high similarity were removed, resulting in the optimized sequences. Double-ended splicing of original sequences was performed using fast length adjustment of short reads (FLASH; version 1.2.11, <http://www.cbcb.umd.edu/software/flash>) [22]. For splicing, the parameter of maximum mismatch rate between overlaps was limited to 0.2 and the length was longer than 10 bp. DADA2 (version 1.2.1, <https://github.com/benjjneb/dada2>) [23] was applied for denoising to identify amplicon sequence variants (ASVs). Taxonomy of the ASVs and species annotations (domain, kingdom, phylum, class, order, family, genus, and species) were performed using the naïve Bayesian classifier, Ribosomal Database Project (RDP; version 11.1, <http://rdp.cme.msu.edu/>) [24]. Clean tags were clustered into optimal taxonomic units (OTUs) with a similarity threshold of 97% using UPARSE (<http://drive5.com/uparse/>) [25]. The UCHIME (http://www.drive5.com/usearch/manual/uchime_

[algo.html](#)) [26] was used to delete chimeric sequences. Then, a representative sequence was assigned to each OTU to obtain annotation information using the RDP classifier with the alignment threshold limited to 70% by comparison with SILVA database. The bar diagram, heatmap, and Circos diagram were drawn using “ggplot2,” “pheatmap,” and “RCircos” R packages to reflect the differences in microbial community composition between control and patient samples, respectively.

Pan/core-genome analyses of the 16S rDNA sequences were performed using “ggplot2” R package. Alpha diversity indicators, including sobs, ace, chao, shannon, simpson, shannoneven, simpsonseven, coverage, and Faith’s Phylogenetic Diversity (PD) indices, were calculated using QIIME software (version 1.9.1) [27]. Graphs of alpha diversity indexes, rarefaction curves (Sobs and Shannon indices), and rank-abundance curve were drawn via “ggplot2” R package. For beta diversity analysis, the QIIME software was used to calculate the unweighted UniFrac distance, followed by construction of the unweighted pair-group method with arithmetic mean (UPGMA) sample clustering tree using “ape” R package. Principal component analysis (PCA) was performed using “ropls” R package. Furthermore, principal coordinate analysis (PCoA) based on the UniFrac distance and nonmetric multidimensional scaling (NMDS) analysis were carried out using “vegan” R package.

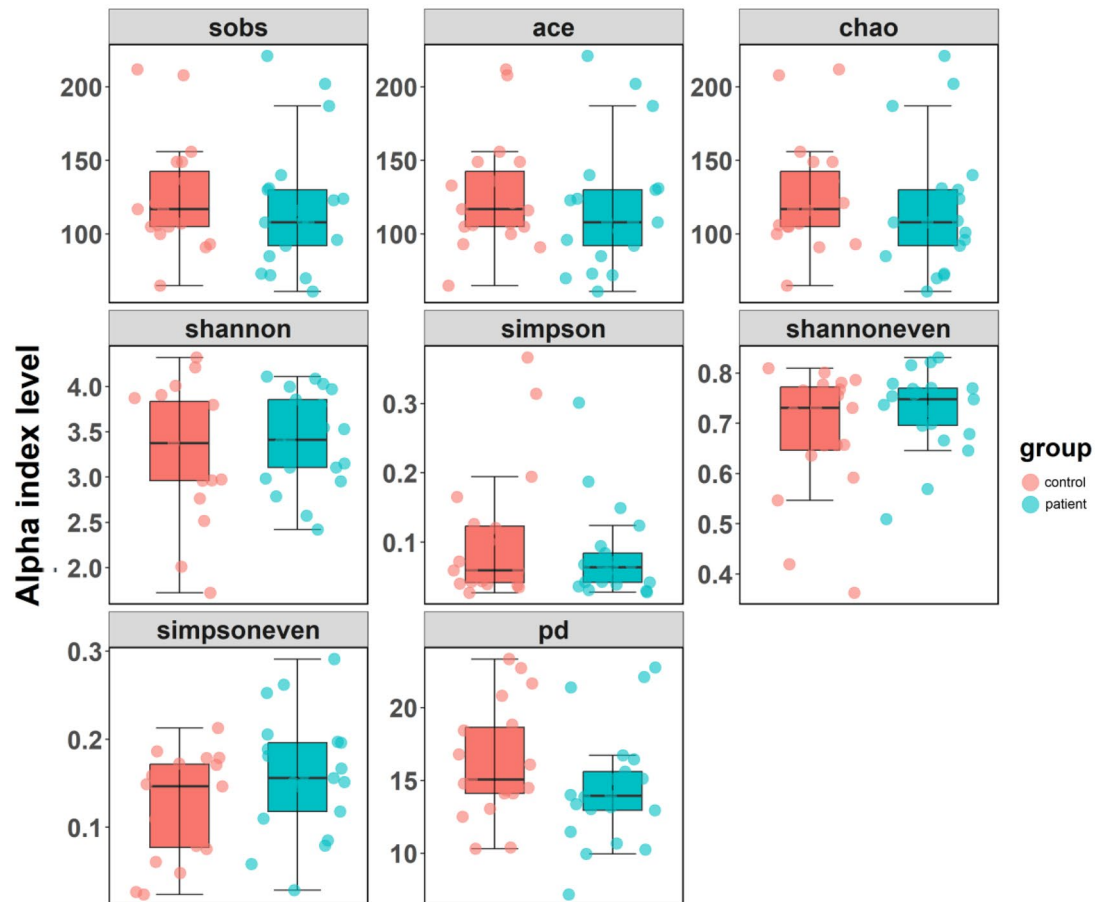
Urine Sample Preparation

First, 400- μ L methanol (containing 1- μ g/mL chlorophenylalanine as internal standard) was added to the urine sample (100 μ L) in the 1.5-mL centrifuge tube. Then, the mixture was vortexed for 30 s and centrifuged for 10 min (14,000 \times g; 4 °C), followed by vacuum drying. The supernatant (200 μ L) was transferred to the autosamplers for subsequent ultra-performance liquid chromatography-mass spectrometry (UPLC-MS) analysis. Additionally, quality control (QC) samples were prepared by blending an equal portion of each urine sample. Prepared QC samples were inserted in every eight urine samples, which were subjected to regular analysis to evaluate the stability and repeatability of the instrumental experiment.

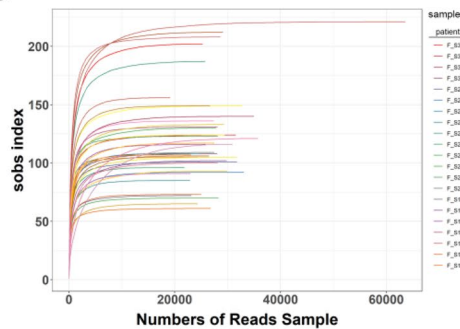
Liquid Chromatography-Mass Spectrometry (LC-MS) Analysis

Sample separation for liquid chromatography analysis was performed using XSelect HSS T3 Column (2.1 mm \times 100 mm; 1.8 μ m; Waters, Milford, MA, USA). The mobile phase was composed of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The conditions of gradient elution were applied as follows: 0–3.5 min, 0–24.5% B, 0.4 mL/min (flow rate); 3.5–5 min,

A



B



C

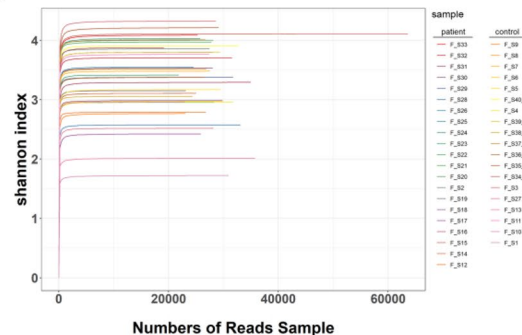


Fig. 1 Changed alpha diversity of gut microbiota in patients with CLC. **A** Box and whisker plots of alpha diversity indexes, including sobs, ace, chao, shannon, simpson, shannoneven, simpsononeven, and PD of control and patient samples. **B** Rarefaction curve of sobs index. **C** Rarefaction curve of shannon index. Alpha diversity indices

were calculated using QIIME software (version 1.9.1) and “ggplot2” R package was used for visualization. Control group, $n=19$; patient group, $n=21$, compensatory liver cirrhosis; PD, Faith’s Phylogenetic Diversity

24.5%–65% B, 0.4 mL/min; 5–5.5 min, 65–100% B; 5.5–7.4 min, 100%B, 0.4–0.6 mL/min; 7.4–7.6 min; 100–51.5%B, 0.6 mL/min; 7.6–7.8 min; 51.5–0%B, 0.6–0.5 mL/min; 7.8–9 min; 0%B, 0.5–0.4 mL/min; 9–10 min; and 0%B, 0.4 mL/min. Other parameters of the

analysis were as follows: injection volume, 2 μ L; column temperature, 40 $^{\circ}$ C; and sample room temperature, 10 $^{\circ}$ C.

For mass spectrometry analysis, positive- and negative-ion modes were applied using the electrospray ionization source of the UHPLC-Q Exactive HF-XQ system (Thermo

Fisher). The mass scan range was set to 50–1000 m/z. The conditions of electrospray ionization were as follows: sheath gas flow rate, 50 arbitrary units; auxiliary gas flow rate, 13 arbitrary units; heater temperature, 425 °C; capillary temperature, 325 °C; ion spray voltage floating, 3500 V (positive mode) and –3500 V (negative mode); collision energy, 20/40/60 eV; full MS resolution, 60,000; and MS/MS resolution, 7500.

Urine Metabolomics Analyses

Raw data were entered into the Progenesis QI software (version 2.2, Waters) for data pre-processing, including baseline filtering, peak identification, alignment, integration, retention time correction, and normalization. Then, a data matrix composed of retention time, mass-to-charge ratio values, and peak intensity was obtained. Metabolite annotation was performed using Human Metabolome Database (HMDB; version 5.0, <http://www.hmdb.ca/>) [28], Kyoto Encyclopedia of Genes and Genomes (KEGG; www.genome.jp/kegg/) [29], and Lipidmaps database (<http://www.lipidmaps.org/>) [30]. Data analysis was conducted using the online platform Majorbio Cloud (<https://cloud.majorbio.com>). Metabolomics features presenting more than 20% of missing values were eliminated from the analysis. After normalization, variables with relative standard deviation > 30% of QC samples were removed. The final data matrix was obtained after log₁₀ transformation and used for subsequent analysis.

PCA and orthogonal partial least-squares discrimination analysis (OPLS-DA) were carried out to determine the distinctions among variables using “ropls” R package. The reliability of OPLS-DA was evaluated via five-fold cross-validation and permutation test with the criteria “R²Y (goodness of fit indicator) and Q² (predictive capacity indicator) > 0.5” [31]. On the basis of OPLS-DA analysis, metabolites meeting the criterion of “variable importance in projection (VIP) > 1” were selected as key variables. The following criteria were applied to identify differentially expressed metabolites: VIP > 1; $p < 0.05$ [32]. The top 50 differential metabolites were screened based on the values of \log_2 fold change from high to low. The heatmap reflecting the distribution of these metabolites in each sample was drawn by “pheatmap” R package. Besides, Pearson correlation analysis was conducted to assess the links among the top 50 differential metabolites, which was visualized by “corrplot” R package. Then, all the differential metabolites were subjected to pathway enrichment analysis using MetaAnalyst (version 4.0, <https://www.metabanalyst.ca>) [33], followed by network analysis using Metscape [34].

Construction and Verification of Diagnostic Model for CLC

Data of gut microbiota and metabolomics analyses were merged and a multi-factor classification model was built based on the least absolute shrinkage and selection operator (LASSO) regression algorithm using “caret” R package (<https://github.com/topepo/caret/>) [35]. Then, the merged gut microbiota and metabolomics data were randomly divided into training (29 samples) and testing (10 samples) sets to avoid overfitting of the model. Through five-fold cross-validation, the optimal diagnostic markers were identified according to the lambda.min value. The final diagnostic model was generated, and the area under the curve (AUC) of the receiver operating characteristic (ROC) was calculated using SPSS software (version 26.0, IBM, Armonk, NY, USA) and visualized using “pROC” R package. Otherwise, Pearson correlation analysis was applied to clarify the relations among the diagnostic markers and visualized by “corrplot” R package.

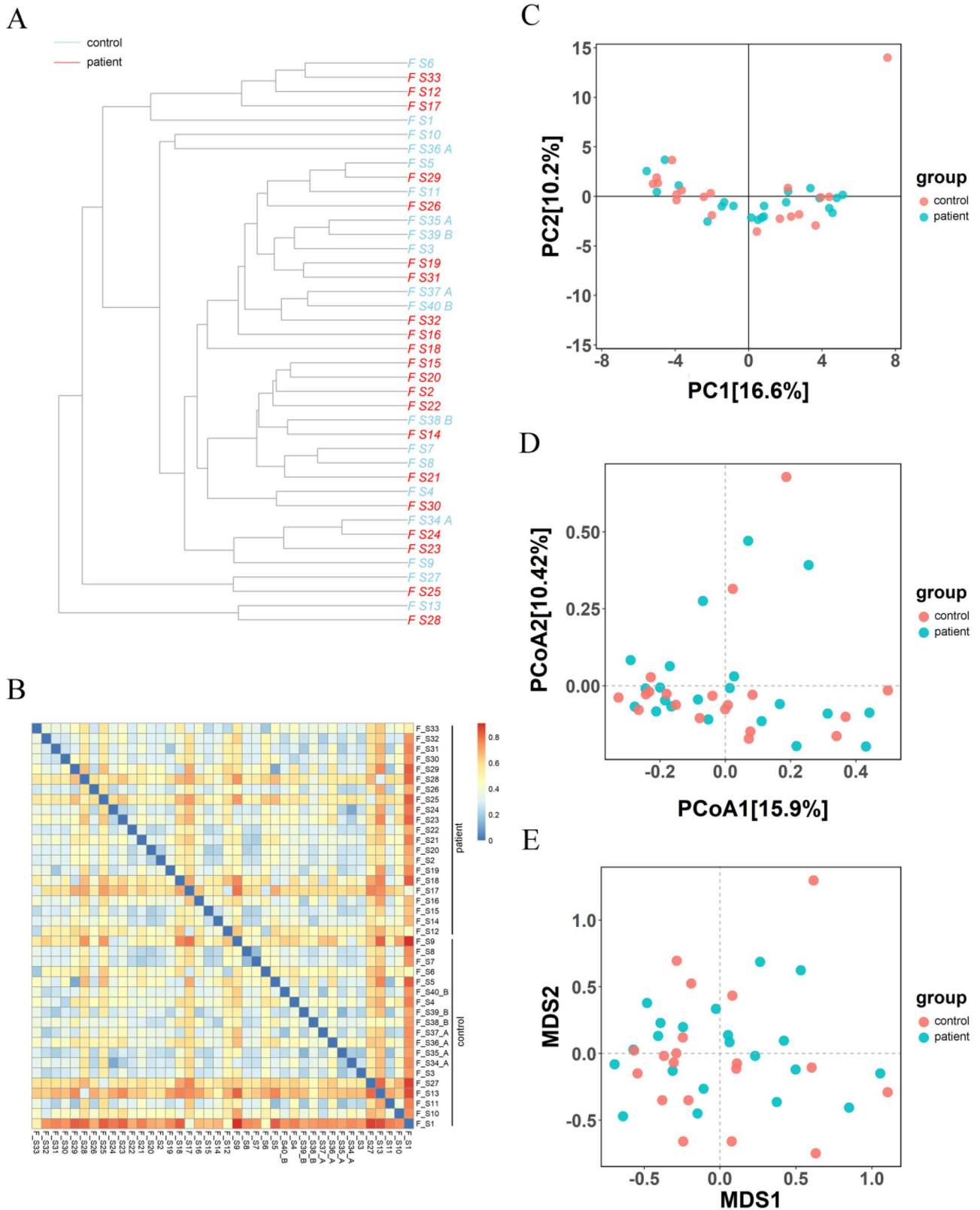
Statistical Analysis

Statistical analyses were undertaken using SPSS software (version 26.0, IBM) and R software (version 4.2.2). Data are exhibited as means \pm standard deviation (SD) or medians with ranges. Fisher’s test was applied to identify significantly enriched pathways related to differential metabolites. Student’s t test was utilized for comparisons of the differences between two groups. The criterion for statistical significance is $p < 0.05$. For controlling multiple hypothesis testing, we employed Bonferroni correction method to adjust statistical significance.

Results

16S rDNA Sequencing Data Analyses

In this study, 16S rDNA sequencing was initially performed on 40 feces samples, including 19 from healthy volunteers (control) and 21 from patients with CLC, resulting in 2,910,360 raw sequences. Quality control was performed using fastp with the following criteria: reads with a length of 300 bp were truncated at sites where the average quality score across a sliding window of 50 bp fell below 20; truncated reads with a length of less than 50 bp were deleted. Furthermore, The raw sequences were compared using BWA software [21], and the contaminated reads with high similarity were deleted, generating the optimized sequences. In this study, a total of 2,362,245 high-quality sequences with the mean length of 412.57 bp were obtained (Table S2).



Pan- and core-genome analyses were utilized to determine whether the current sequencing sample size is sufficient by evaluating alterations in total, and core ASVs as the

sample number were elevated. Pan-genome analysis revealed that the total number of ASVs tended to be maximized with the increasing number of samples (Figure S1A). Moreover,

Fig. 2 Altered beta diversity of gut microbiota in patients with CLC. **A** UPGMA clustering tree of the control and patient samples. The unweighted UniFrac distance was calculated using QIIME software (version 1.9.1) and visualized by “ape” R package. **B** Heatmap of the unweighted UniFrac distance among the control and patient samples, visualized by “pheatmap” R package. **C** PCA analysis of control and patient samples using “ropis” R package. **D** PCoA analysis of control and patient samples using “vegan” R package. **E** NMDS analysis of control and patient samples using “vegan” R package. Control group, $n=19$; patient group, $n=21$. CLC, compensatory liver cirrhosis; UPGMA, unweighted pair-group method with arithmetic mean; PCA, principal component analysis; PCoA, principal coordinate analysis; NMDS, nonmetric multidimensional scaling

core-genome analysis showed that the number of key ASVs descended and then remained unchanged with the growing number of samples (Figure S1B). Furthermore, the rank-abundance curves tended to be smooth, demonstrating that most of the microbial diversity on genus level in each sample had already been acquired at the present sequencing depth (Figure S1C).

Alpha and Beta Diversity Analyses

Based on the ASVs, alpha diversity analyses were performed to evaluate the richness and diversity. In this study, we assessed 9 alpha diversity indicators, including Sobs, ace, chao, shannon, simpson, shannoneven, simpsonseven, coverage, and PD indexes. Sobs, ace, and chao indices reflect community richness [36]; shannon, simpson, and PD reflect community diversity [36, 37]; shannoneven and simpsonseven reflect community evenness [37]; and coverage index reflects community coverage [38]. Our data showed that sobs, ace, chao, simpson, and PD indexes of patient samples were lower than those of control samples, while simpsonseven index was higher in the patient samples (Fig. 1A). These alpha diversity indices of each sample are listed in Table S3. Furthermore, rarefaction curves were built using alpha diversity indices (Sobs and Shannon) for each sample to reflect the microbial diversity of each sample at different sequencing amounts. Here, the rarefaction curves tended to be flat, indicating that the amount of present sequencing data was sound and large enough to mirror the majority of microbial information in the feces samples (Fig. 1B, C).

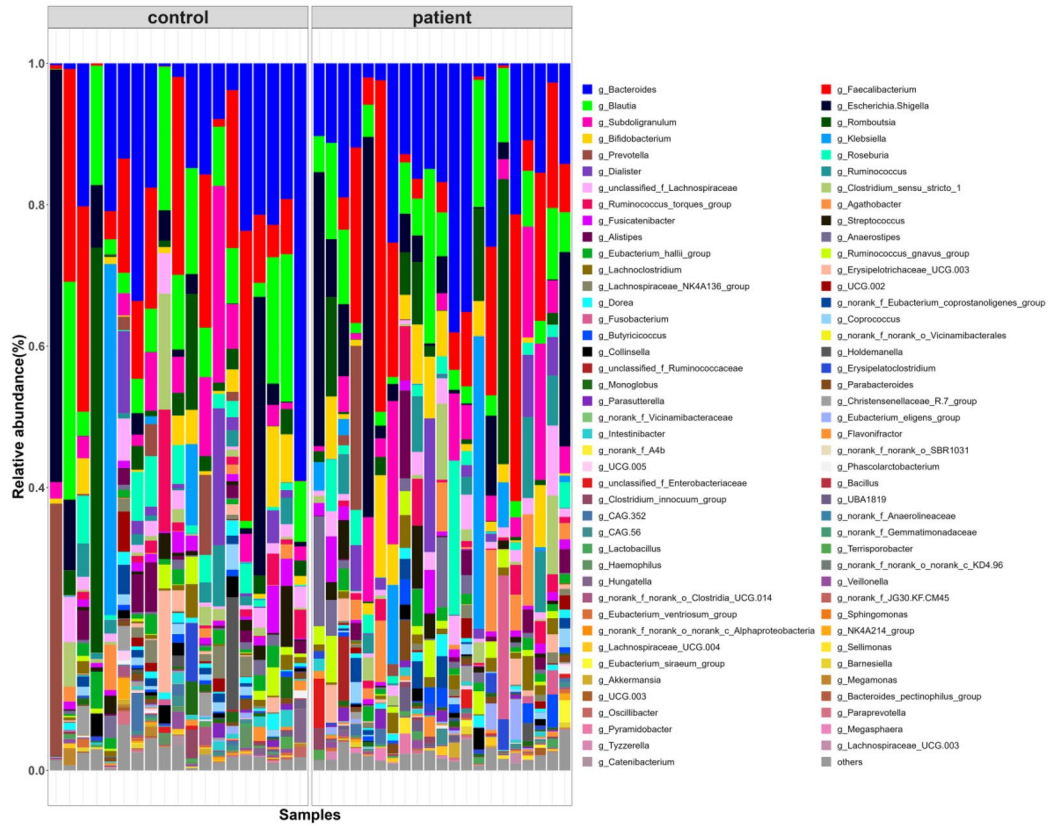
Beta diversity analyses were conducted to figure out the similarity or difference in the community composition between controls and patients. UPGMA clustering tree analysis is a data visualization approach to reflect the degree of variation in microbial evolution in multiple samples [39]. UPGMA clustering tree analysis and related heatmap showed different degrees of discrepancies and similarities in the microbial composition among the samples, as the distances varied between every two samples (Fig. 2A, B). PCA reflects the discrepancies in various datasets on a two-dimensional coordinate plot through

variance decomposition. The higher the similarity in the sample composition between two samples, the shorter the distance between them presented in the PCA plot [40]. According to PCA on ASV level, the vast majority of dots (samples) in patient group were separated from those in control group (PC1 = 16.6%; PC2 = 10.2%; Fig. 2C). PCoA is a data dimensionality reduction analysis method, allowing visualization of discrepancies between samples [41]. In this study, most of the dots in patient group were significantly isolated from those in control group in the pots of PCoA (PCoA1 = 15.9%; PCoA2 = 10.42%; Fig. 2D). NMDS is also a widely used visualization method to study sample differences [42]. NMDS on ASV level further demonstrated the differences in the overall microbial composition between normal subjects and CLC patients (Fig. 2E).

Microbial Community Structure Analyses

The bar diagram, heatmap, Circos diagram, and histogram of microbial average abundance were used to clarify the discrepancies in microbial community structure between control and patient fecal samples. The bar diagram showed variations in the relative abundance of different microbiota among these 40 fecal samples on the genus level (Fig. 3A). *Bacteroides* (a maximum of 59% in control, 38% in patient), *Faecalibacterium* (a maximum of 41% in control, 49% in patient), *Blautia* (a maximum of 31% in control, 24% in patient), and *Escherichia. Shigella* (a maximum of 58% in control, 51% in patient) were the major genera in these samples (Table S4). Besides, the heatmap illustrated that each genus was differently distributed in each sample, which further verified the findings of the bar diagram analysis (Fig. 3B). The Circos diagram directly reflected the distribution of each genus in control and patient samples. In the Circos diagram, the overall abundance of some genera (e.g., *Blautia*) in patient samples was lower than that of control samples, while some genera (e.g., *Escherichia. Shigella*) showed higher overall abundance in patient samples (Figure S2). Through analysis of average microbial abundance, we found changed microbial abundance of some genera in patient samples compared with that of control samples. For instance, the average abundance of *Bifidobacterium*, *Lachnospirillum*, and *Ruminococcus_gnavus_group* was elevated in patient samples compared with that in control samples (Fig. 4A). However, decreased average abundance of *Christensenellaceae_R_7_group*, *Coprococcus*, and *Eubacterium_ventriosum_group* was observed in patient samples by comparison with control samples (Fig. 4B).

A



B



Fig. 3 Changed microbial community composition in patients with CLC. **A** Bar diagram of the relative abundance of different microbial taxa on genus level of each control and patient sample, visualized by “ggplot2” R package. **B** Heatmap of the microbial community composition of each control and patient sample, visualized by “pheatmap” R package. Control group, $n=19$; patient group, $n=21$. CLC, compensatory liver cirrhosis

Urine Metabolomics Analysis

Metabolomics analysis was conducted on 40 urine samples (control: 20; patient: 20). Through UPLC-MS and data pre-processing, 9,351 peaks were picked and 2,637 metabolites with specific names were identified (Table S5).

To assess the metabolic variations in CLC patients compared with healthy individuals, PCA and OPLS-DA analyses were conducted on the patient and control urine samples. According to PCA analysis based on QC and testing samples, PC1 and PC2 accounted for 15.6% and 9.5% variables, respectively; significantly, the majority of patient samples were isolated from control samples (Fig. 5A). Besides, a notable separation tendency was observed between control and patient samples based on the OPLS-DA model (Fig. 5B). Further, the OPLS-DA model exhibited a favorable explanatory ability ($R^2Y=0.876$; p value <0.01) and predictive capacity ($Q^2=0.703$, p value <0.01) through five-fold cross-validation and permutation test (Fig. 5C).

Determination of Differential Metabolites

In the plot of VIP scores of the identified metabolites, the red dots denoted the metabolites with $VIP > 1$ and were regarded as potential biomarkers for CLC based on the OPLS-DA model (Fig. 6A). Notably, 841 differential metabolites were identified under the criteria of “ $VIP > 1$; $p < 0.05$ ” and Bonferroni correction was used to adjust the p values (Table S6). The volcano plot showed 533 upregulated metabolites (red dots) and 308 downregulated metabolites (blue dots) (Fig. 6B). For example, the abundance of 3-Oxovalproic acid and 2-(4'-Methylthio)butylmalate was markedly upregulated in patient samples compared with those in control samples ($p < 0.05$; Fig. 6C). Conversely, 9-(2, 3-Dihydroxypropoxy)-9-Oxononanoic acid and methionine glutamate exhibited notable downregulated abundance in patient samples by comparison with those in control samples ($p < 0.01$; Fig. 6C). The distribution of the top 50 differential metabolites in all samples was displayed in the heatmap (Figure S3A).

Correlation and Pathway Analyses of Differential Metabolites

Also, Pearson correlation analysis was performed to figure out the correlations among these 841 differential

metabolites. Results showed multiple relations among the top 50 differential metabolites (Figure S3B). Pathway enrichment analysis showed that the differential metabolites were concentrated in 36 pathways. The top 5 significantly enriched pathways included tryptophan metabolism, nicotinate and nicotinamide metabolism, purine metabolism, steroid hormone biosynthesis, and histidine metabolism ($p < 0.05$; Fig. 7). The corresponding networks of the top 5 pathways are displayed in Figure S4.

Establishment and Validation of a Diagnostic Model for CLC

Through five-fold cross-validation, a 9-marker diagnostic model for CLC was built. This model consisted of 7 intestinal microorganisms (*Blautia*, *Subdoligranulum*, *Agathobacter*, *norank_f_Eubacterium_coprostanoligenes_group*, *Butyricicoccus*, *Monoglobus*, and *Lachnospiraceae_UCG_004*) and 2 metabolites (L-2,3-Dihydrodipicolinate and 5-Acetamidovalerate) (Table 1). Notably, ROC curves showed that the 9-marker model had optimal performance with favorable AUC values in both training set (AUC = 0.924, 95% CI = 0.834–1) and testing set (AUC = 0.95, 95% CI = 0.811–1) (specificity $> 90\%$, sensitivity $> 90\%$; Fig. 8A). Furthermore, the heatmap showed various relations among these intestinal floras and metabolites (Fig. 8B). For example, *Agathobacter* was positively related to L-2,3-Dihydrodipicolinate ($r=0.42$), while negatively related to 5-Acetamidovalerate ($r=0.18$). *Norank_f_Eubacterium_coprostanoligenes_group* was positively correlated with L-2,3-Dihydrodipicolinate ($r=0.2$) and 5-Acetamidovalerate ($r=0.05$). In this model, the relative abundance of *Subdoligranulum*, *Agathobacter*, *norank_f_Eubacterium_coprostanoligenes_group*, *Butyricicoccus*, and *Lachnospiraceae_UCG_004* was increased in CLC patients, whereas *Blautia* and *Monoglobus* were decreased (Fig. 8C). Otherwise, the abundance of L-2,3-Dihydrodipicolinate was markedly elevated in CLC patients, while 5-Acetamidovalerate was reduced ($p < 0.05$; Fig. 8D).

Discussion

LC is a primary cause of death and a common outcome of multiple progressive liver disorders, leading to more than one million deaths worldwide each year [43, 44]. In the present study, 16S ribosomal DNA (rDNA) sequencing and untargeted metabolomics were used to investigate the status of the gut microbiota and metabolome in patients with CLC. We observed differences in the intestinal microorganisms and urinary metabolites between healthy participants and patients with CLC. Finally, a 9-signature diagnostic model for CLC based on the differential gut and metabolites was

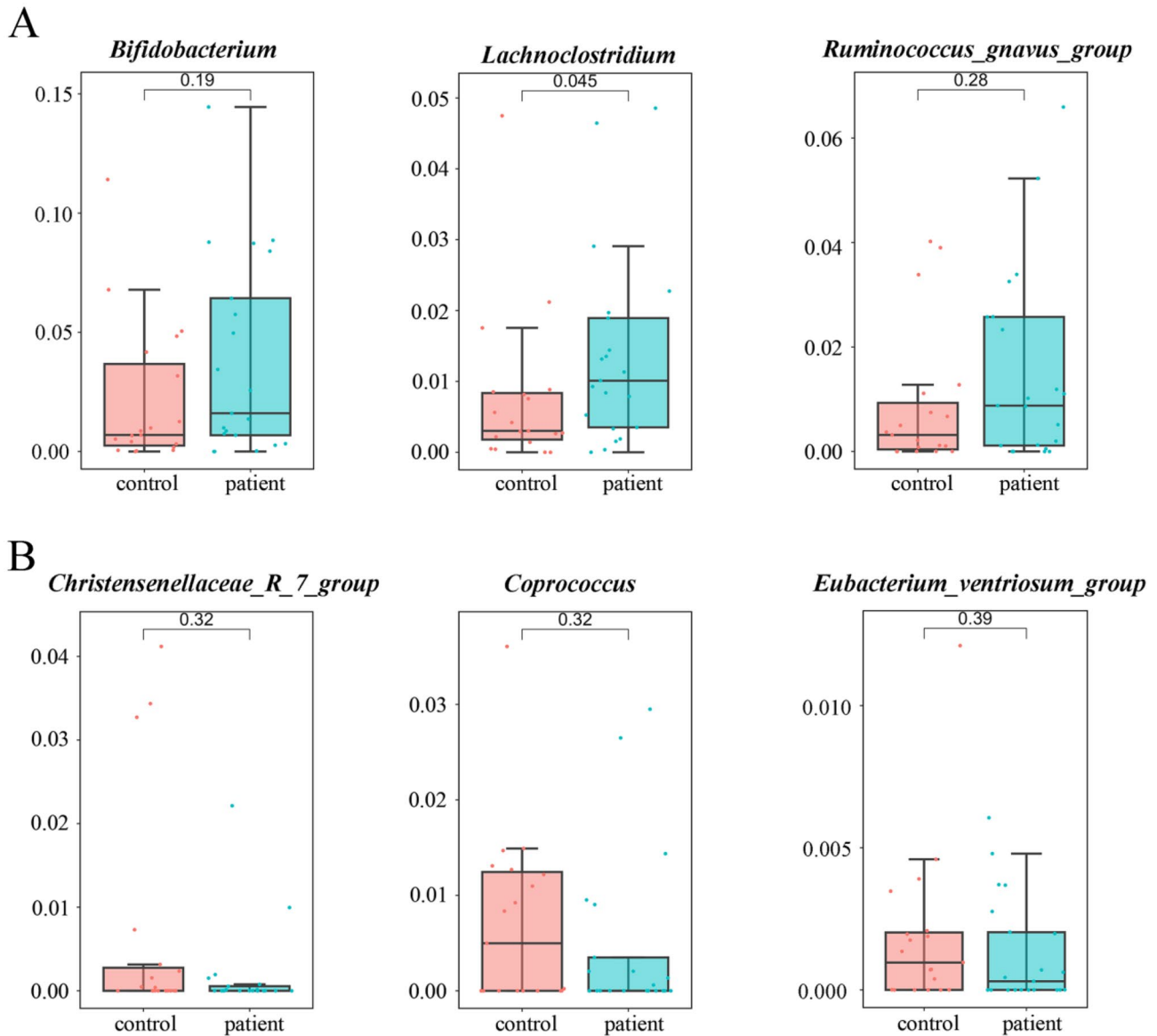


Fig. 4 Altered microbial community composition in patients with CLC. **A** Box and whisker plots of elevated average abundance of microbial taxa in patient samples compared with that in control samples. *Bifidobacterium*, *Lachnocostridium*, and *Ruminococcus_gnavus_group* were displayed. **B** Box and whisker plots of reduced aver-

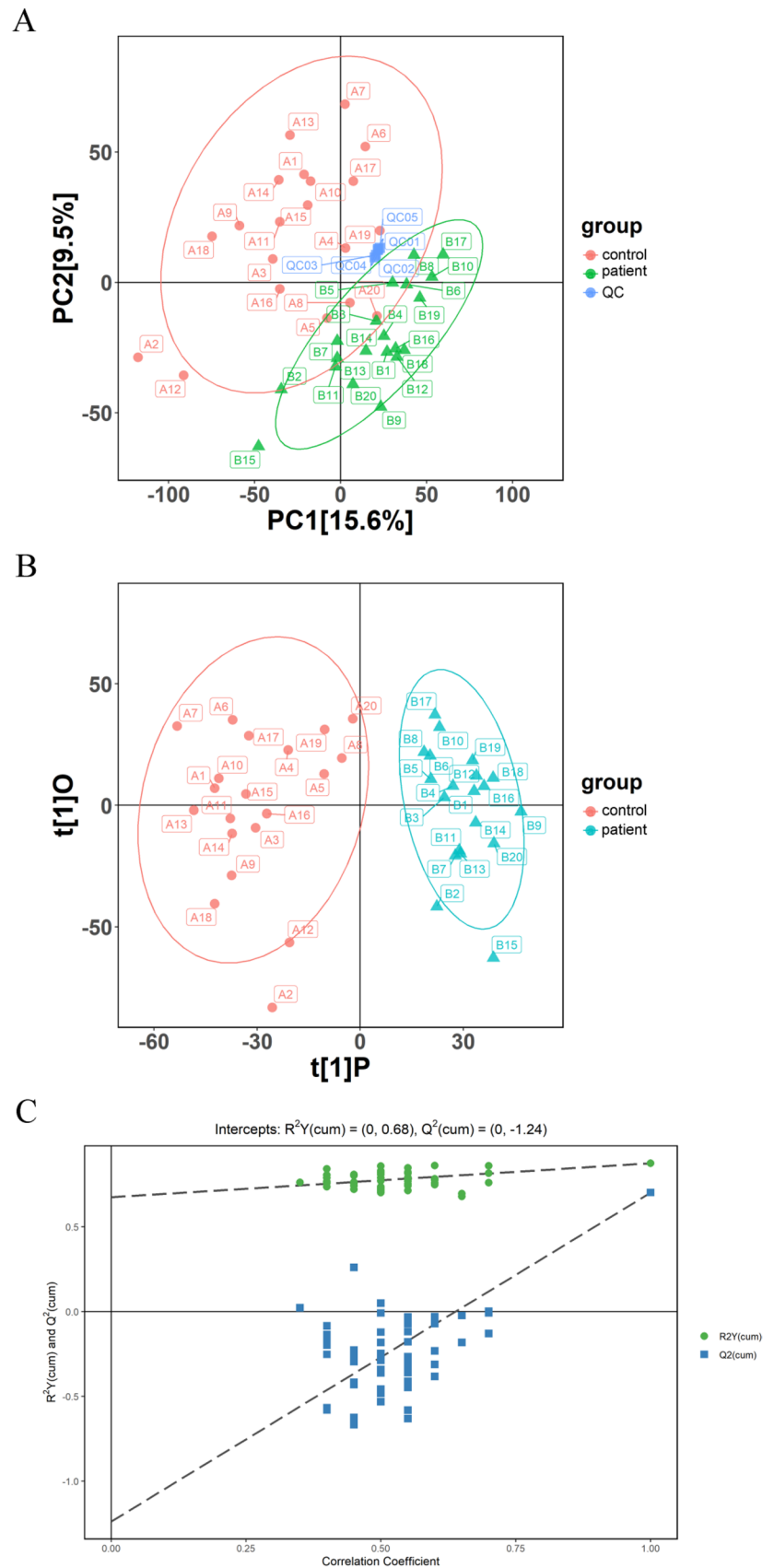
age abundance of microbial taxa in patient samples compared with that in control samples. Box and whisker plots were visualized by “ggplot2” R package. Control group, $n=19$; patient group, $n=21$. *Christensenellaceae_R_7_group*, *Coprococcus*, and *Eubacterium_ventriosum_group* were displayed. CLC, compensatory liver cirrhosis

constructed and verified to be reliable. The diagnostic model comprised 7 intestinal microorganisms (*Blautia*, *Subdoligranulum*, *Agathobacter*, *norank_f_Eubacterium_coprostanoligenes_group*, *Butyricoccus*, *Monoglobus*, and *Lachnospiraceae_UCG_004*) and 2 metabolites (L-2,3-Dihydrodipicolinate and 5-Acetamidovalerate).

Accumulating evidence indicates that disordered gut microbiome is implicated in the pathogenesis of LC [5, 9]. Currently, impaired gut-liver axis has been recognized as

perturbed intestinal microbial structure, disrupted intestinal barrier, and enhanced permeability of intestinal barrier and found in various liver disorders [10, 45]. In this study, we found that the microbial alpha diversity indexes (sobs, ace, chao, simpson, and PD) were decreased in patients with CLC compared with those in healthy participants. Moreover, the beta diversity analysis showed that the microbial composition in patients with CLC significantly varied from that in healthy participants. These results suggest that abnormalities

Fig. 5 Altered urine metabolome in patients with CLC. **A** PCA analysis of urine samples in control and patient groups based on QC samples using “ggplot2” R package. Control group, $n=20$; patient group, $n=20$; QC, $n=5$. **B** OPLS-DA analysis of urine samples in control and patient groups using “ropls” R package. Control group, $n=20$; patient group, $n=20$. **C** OPLS-DA permutation test in the negative and positive ion modes. CLC, compensatory liver cirrhosis; PCA, principal component analysis; QC, quality control; OPLS-DA, orthogonal partial least-squares discrimination analysis



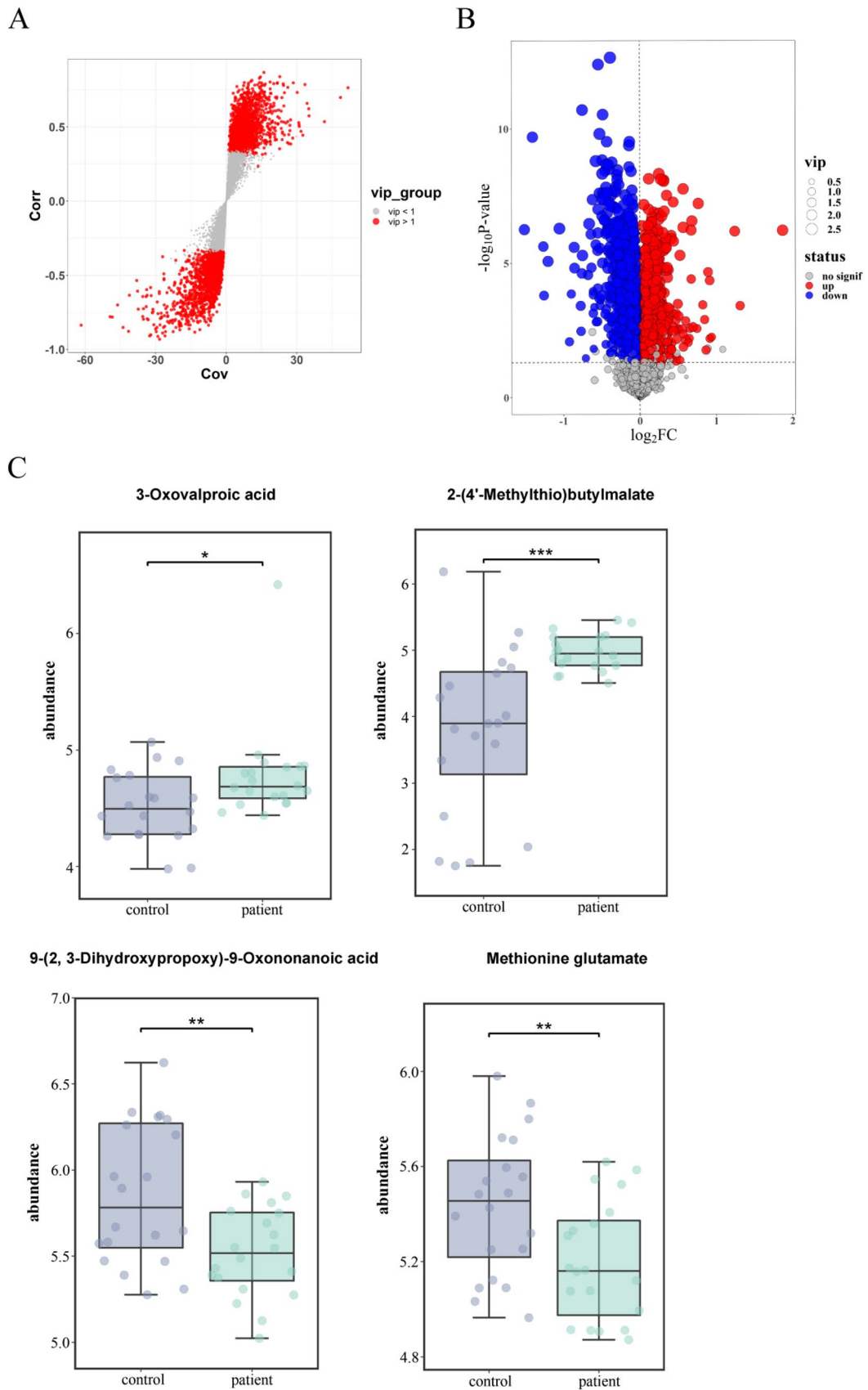


Fig. 6 Identification of differential metabolites in patients with CLC. **A** Plot of VIP scores of the identified metabolites. Red dots denote the metabolites with $VIP > 1$; gray dots denote the metabolites with $VIP < 1$. **B** Volcano plot of differential metabolites. Red dots denote upregulated differential metabolites; blue dots denote downregulated differential metabolites. $VIP > 1$; $p < 0.05$. **C** Box and whisker plots of the abundance of differential metabolites in control and patient samples. “ggplot2” R package was used for visualization. Control group, $n = 20$; patient group, $n = 20$. 3-Oxovalproic acid, 2-(‘methylthio)butylmalate, 9-(2, 3-dihydroxypropoxy)-9-oxononanoic acid, and methionine glutamate were displayed. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$ vs control group. CLC, compensatory liver cirrhosis; VIP, variable importance in projection

in gut microbiome occurred in patients with CLC, which may conduce to the onset of LC. We found elevated abundance of the key microorganisms, *Subdoligranulum*, *Agathobacter*, *norank_f_Eubacterium_coprostanoligenes_group*, *Butyricoccus*, and *Lachnospiraceae_UCG_004* in patients with CLC compared with those in healthy participants. Oppositely, the abundance of *Blautia* and *Monoglobus* was reduced in patients with CLC.

Subdoligranulum, *Butyricoccus*, and *Agathobacter* are the common producers of butyrate, a short-chain fatty acid (SCFA) [46, 47]. SCFAs, including acetate, propionate, and butyrate, are the major energy source for colonocytes, which have been proven to enhance the function of gut barrier [48, 49]. The liver is exposed to inflammatory signaling by impaired gut barrier, resulting in liver injury [49]. Usually, SCFAs are thought to be advantageous to the body. However, a previous study indicated that increased SCFAs could enhance liver inflammation and fibrosis, conducting to the progression of hepatocellular carcinoma [50]. It was previously revealed that increased *Agathobacter* induced by protein intake could promote the development of hepatic encephalopathy in LC [51]. In terms of *norank_f_Eubacterium_coprostanoligenes_group*, little is known about its role in LC development. Nevertheless, a recent study showed that *norank_f_Eubacterium_coprostanoligenes_group* level was notably elevated in rheumatoid arthritis (RA) [52]. Further, it was revealed that lowering *norank_f_Eubacterium_coprostanoligenes_group* and the overall level of SCFAs could alleviate RA [52]. Thus, we can hypothesize the crosstalk between *norank_f_Eubacterium_coprostanoligenes_group* and SCFAs in the inflammatory mechanism in LC. *Lachnospiraceae* species are also involved in the production of butyrate. Previous data observed a rise in *Lachnospiraceae_UCG_004* in human immunodeficiency virus (HIV) patients [53]. Moreover, HIV infection has been reported to contribute to LC [54]. *Blautia* are anaerobic bacteria with a probiotic feature, extensively found in mammalian feces and guts and they can produce acetate [55, 56]. Lately, a reduction in the abundance of *Blautia* has been found in patients with liver injury [57]. *Monoglobus* have been demonstrated to be implicated in dietary fiber

fermentation and related to healthy communities [58]. A recent study has indicated that *Monoglobus* are specialized pectin-degrading bacteria linked with neutrophilic inflammation and serious liver damage [59]. Downregulation of *Monoglobus* promotes systemic inflammation, as evidenced by recent studies [59, 60]. Taken together, these findings suggest that these potential bacteria markers may participate in the pathogenesis of early-stage LC by regulating SCFA production and inflammatory responses.

Increasing data indicate that aberrant gut microbiome is closely associated with metabolic alterations in LC development [8, 11]. Abnormal gut-liver axis results in the transposition of microorganisms and their products, such as lipopolysaccharide and other metabolites into the portal bloodstream, straightly directed at the liver [45]. Through PCA and OPLS-DA analyses, we observed notable differences in the urine metabolome between CLC patients and healthy subjects. Specifically, we obtained 841 differential metabolites in CLC patients, including 533 metabolites upregulated and 308 downregulated.

Through pathway enrichment analysis, we observed that these differential metabolites were primarily concentrated in pathways, such as tryptophan metabolism, purine metabolism, and steroid hormone biosynthesis. Reduced tryptophan, elevated tryptophan-associated enzymes, and increased downstream metabolites are related to aggravated metabolic inflammation and fibrosis [45]. The tryptophan catabolic responses are mediated by intestinal bacteria, such as *Ruminococcus gnavus* [61]. A previous study demonstrated tryptophan metabolism was impaired in LC patients, which might be implicated in the pathogenesis of LC [62]. Purine metabolism acts as a switch in various biological processes (e.g., energy generation and DNA/RNA synthesis) [63]. Abnormal purine metabolism can conduce to the development of multiple disorders, especially hyperuricemia [64]. Lately, notable alterations in purine metabolic pathways have been observed in LC mice, indicating that dysregulated purine metabolism might give rise to LC [11]. Steroid hormones modulate a variety of biological mechanisms, principally in the reproductive system and multiple metabolic pathways [65]. As the pivotal metabolic organ, the liver acts as a key part in the homeostasis of steroid hormones [65]. Aberrant steroid hormone levels have been demonstrated to be closely linked with several liver conditions. For example, hypoenestrogenism can result in the occurrence of and progression of NAFLD in post-menopausal females [66]. Otherwise, a previous study revealed that lacking of the major androgen, testosterone, might lead to sarcopenia in male patients with LC [67]. To sum up, these data indicate the involvement of various disturbed metabolic pathways in the onset and development of LC.

In terms of the metabolite markers, the abundance of L-2,3-Dihydrodipicolinate was increased in CLC patients,

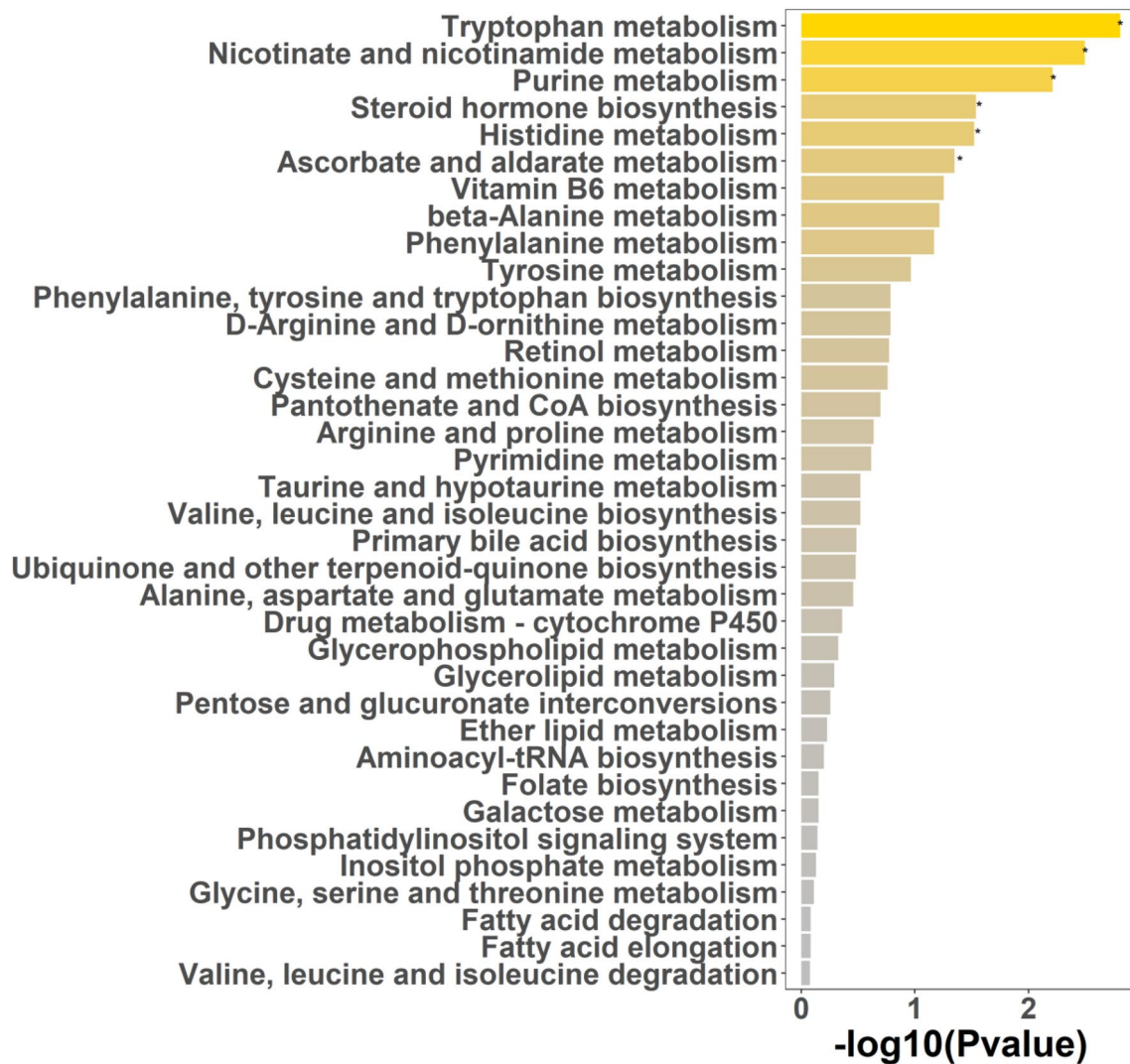


Fig. 7 Pathway enrichment analysis of differential metabolites in CLC patients using MetaboAnalyst software. CLC, compensatory liver cirrhosis

Table 1 Parameters of the optimal diagnostic model for CLC

	Marker	Coef
(Intercept)		0.095743656
g_Blautia	g_Blautia	-3.111189964
g_Subdoligranulum	g_Subdoligranulum	1.135761231
g_Agathobacter	g_Agathobacter	1.495834703
g_norank_f_Eubacterium_coprostanoli-genes_group	g_norank_f_Eubacterium_coprostanoli-genes_group	0.750636129
g_Butyricoccus	g_Butyricoccus	5.608887244
g_Monoglobus	g_Monoglobus	-2.592074156
g_Lachnospiraceae_UCG_004	g_Lachnospiraceae_UCG_004	2.253676077
metab_403	L-2,3-Dihydrodipicolinate	0.136301022
metab_7496	5-Acetamidovalerate	-0.131219474

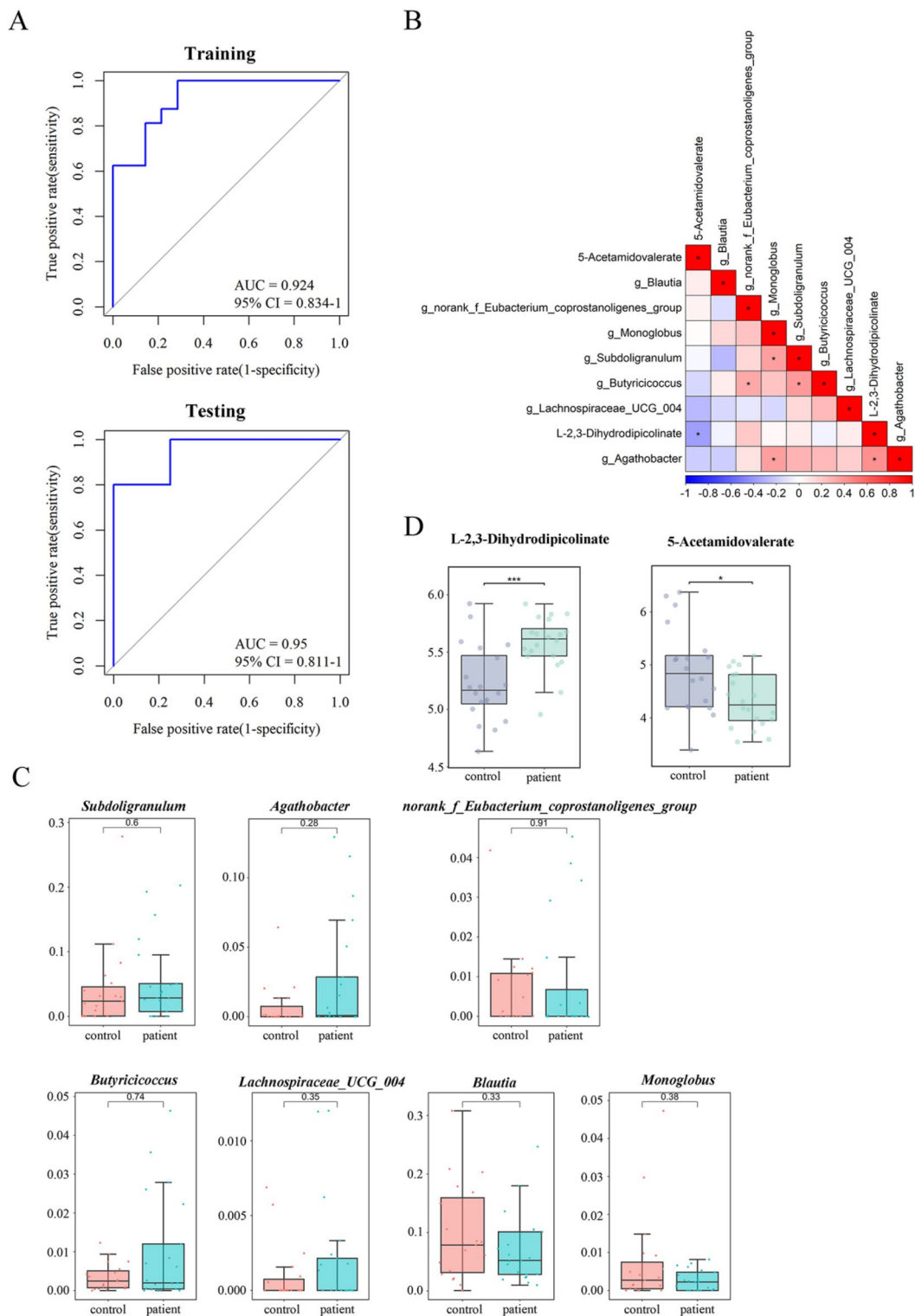


Fig. 8 Construction and validation of the diagnostic model for CLC. **A** ROC analysis in the training and testing sets using “pROC” R package. **B** Heatmap of correlation analysis of the 9 diagnostic biomarkers for CLC, visualized by “corrplot” R package. **C** Box and whisker plots of the average abundance of the 7 gut microbiota biomarkers in control and patient samples, visualized by “ggplot2” R

package. Control group, $n=19$; patient group, $n=21$. **D** Box and whisker plots of the abundance of the 2 metabolite biomarkers in control and patient samples. Control group, $n=20$; patient group, $n=20$. * $p < 0.05$ and *** $p < 0.001$ vs control group. CLC, compensatory liver cirrhosis; ROC, receiver operating characteristic

while 5-Acetamidovalerate was decreased. To our knowledge, L-2,3-Dihydrodipicolinate is a member of the class of alpha amino acids and derivatives and a major metabolite, found in all living organisms, from bacteria to humans. However, few investigations regarding L-2,3-Dihydrodipicolinate have been carried out. Otherwise, 5-Acetamidovalerate belongs to the straight chain fatty acid family, which can be formed by the enzymatic reduction of 5-aminopentanoate or enzymatic oxidation of 2-keto-6-acetamidocaproate (<http://www.hmdb.ca/>). Accumulating evidence indicates that disrupted fatty acid levels conduce to liver diseases, including LC [68, 69]. A previous investigation found that altered 5-Acetamidovalerate was associated with inflammation in aging mice, and it was positively related to *Ruminococcus* [70]. These results suggest that the onset and development of LC may be attributed to the abnormal metabolism of alpha amino acids and straight chain fatty acids.

Nonetheless, this study has some limitations need to be paid attention to. First, a larger number of human subjects are required to validate the sequencing and metabolomics analyses. Second, the crosstalk between microbiome and metabolome in CLC should be thoroughly analyzed. Third, the potential diagnostic model for CLC remains to be clinically confirmed.

Conclusion

In this study, we determined a 9-signature diagnostic model for CLC, including *Blautia*, *Subdoligranulum*, *Agathobacter*, *norank_f_Eubacterium_coprostanoligenes_group*, *Butyricoccus*, *Monoglobus*, *Lachnospiraceae_UCG_004*, L-2,3-Dihydrodipicolinate, and 5-Acetamidovalerate. Besides, abnormal tryptophan, nicotinate and nicotinamide, purine, and histidine metabolism, and steroid hormone biosynthesis were implicated in the pathogenesis of CLC. This study may provide a novel strategy for early diagnosis and treatment of LC and a basis for the discovery of drug targeting microbiome or metabolism for CLC.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12033-023-00922-9>.

Acknowledgements Not applicable.

Author Contributions YC participated in the conceptualization, data curation, formal analysis, investigation, funding acquisition, and writing of the original draft. SC participated in the conceptualization, data curation, methodology, software, visualization, and validation. CX participated in the methodology, data curation, formal analysis, and validation. SC participated in the methodology, data curation, formal analysis, and validation. JB participated in the formal analysis, investigation, supervision, and validation. JW participated in the data curation, formal analysis, and investigation. JW participated in the conceptualization, project administration, resources, supervision, funding

acquisition, and writing, reviewing, and editing of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by [Public Welfare Project of the Science and Technology Agency, Zhejiang Province] under grant [number LGF19H030001] and [Medical and Health Research Project of Zhejiang Province] under grant [number 2021ZH057].

Data Availability The data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of Interest The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethics Approval All participants signed written informed consent, and the experimental procedures were authorized by the Ethics Committee of Tiantai People's Hospital of Zhejiang Province (TYLL2019-06). This study was in line with the Declaration of Helsinki.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Wilson, R., & Williams, D. M. (2022). Cirrhosis. *Medical Clinics of North America*, 106, 437–446.
2. Rashid, A., Gupta, A., Adiamah, A., West, J., Grainge, M., & Humes, D. J. (2022). Mortality following appendicectomy in patients with liver cirrhosis: A systematic review and meta-analysis. *World Journal of Surgery*, 46, 531–541.
3. Rojas-Acuña, D., Polo-Samillan, N., Vasquez-Chavesta, A. Z., Escalante-Arias, C., Rios-Perez, C. J., & Toro-Huamanchumo, C. J. (2022). Morbimortality associated with liver cirrhosis in Peru: an ecological analysis for the period of 2004–2016. *International Journal of Environmental Research Public Health*, 19, 9036.
4. Lurje, I., Hammerich, L., & Tacke, F. (2020). Dendritic cell and T cell crosstalk in liver fibrogenesis and hepatocarcinogenesis: implications for prevention and therapy of liver cancer. *International Journal of Molecular Sciences*, 21, 7378.
5. Fan, Y., Li, Y., Chu, Y., Liu, J., Cui, L., & Zhang, D. (2021). Toll-like receptors recognize intestinal microbes in liver cirrhosis. *Frontiers in Immunology*, 12, 608498.
6. Ou, M., Guo, X., Li, Y., Zhang, H., Liu, T., Liu, Q., et al. (2022). Differences in anxiety among patients with liver cirrhosis with different compensation abilities. *American Journal of Translational Research*, 14, 5187–5194.

7. Xie, Y., He, C., & Wang, W. (2022). A potential novel inflammation biomarker for predicting the prognosis of decompensated liver cirrhosis. *Annals of Medicine*, *54*, 3201–3210.
8. Shao, L., Ling, Z., Chen, D., Liu, Y., Yang, F., & Li, L. (2018). Disorganized gut microbiome contributed to liver cirrhosis progression: A meta-omics-based study. *Frontiers in Microbiology*, *9*, 3166.
9. Lee, N. Y., & Suk, K. T. (2020). The role of the gut microbiome in liver cirrhosis treatment. *International Journal of Molecular Sciences*, *22*, 199.
10. Wang, R., Tang, R., Li, B., Ma, X., Schnabl, B., & Tilg, H. (2021). Gut microbiome, liver immunology, and liver diseases. *Cellular and Molecular Immunology*, *18*, 4–17.
11. Xiong, Y., Wu, L., Shao, L., Wang, Y., Huang, Z., Huang, X., et al. (2021). Dynamic alterations of the gut microbial pyrimidine and purine metabolism in the development of liver cirrhosis. *Frontiers in Molecular Biosciences*, *8*, 811399.
12. Li, J., Cao, Y., Lu, R., Li, H., Pang, Y., Fu, H., et al. (2020). Integrated fecal microbiome and serum metabolomics analysis reveals abnormal changes in rats with immunoglobulin a nephropathy and the intervention effect of Zhen Wu Tang. *Frontiers in Pharmacology*, *11*, 606689.
13. Philips, C. A., Augustine, P., Yerol, P. K., Ramesh, G. N., Ahamed, R., Rajesh, S., et al. (2020). Modulating the intestinal microbiota: therapeutic opportunities in liver disease. *Journal of Clinical and Translational Hepatology*, *8*, 87–99.
14. Liu, Y., Jin, Y., Li, J., Zhao, L., Li, Z., Xu, J., et al. (2018). Small bowel transit and altered gut microbiota in patients with liver cirrhosis. *Frontiers in Physiology*, *9*, 470.
15. Zhao, L., Wang, C., Peng, S., Zhu, X., Zhang, Z., Zhao, Y., et al. (2022). Pivotal interplays between fecal metabolome and gut microbiome reveal functional signatures in cerebral ischemic stroke. *Journal of Translational Medicine*, *20*, 459.
16. Hu, Y., Chen, J., Xu, Y., Zhou, H., Huang, P., Ma, Y., et al. (2020). Alterations of gut microbiome and metabolite profiling in mice infected by *Schistosoma japonicum*. *Frontiers in Immunology*, *11*, 569727.
17. Bauset, C., Gisbert-Ferrández, L., & Cosín-Roger, J. (2021). Metabolomics as a promising resource identifying potential biomarkers for inflammatory bowel disease. *Journal of Clinical Medicine*, *10*, 622.
18. Xu, X. Y., Ding, H. G., Li, W. G., Xu, J. H., Han, Y., Jia, J. D., et al. (2020). Chinese guidelines on the management of liver cirrhosis (abbreviated version). *World Journal of Gastroenterology*, *26*, 7088–7103.
19. Wang, Z., Cai, Z., Ferrari, M. W., Liu, Y., Li, C., Zhang, T., et al. (2021). The correlation between gut microbiota and serum metabolomic in elderly patients with chronic heart failure. *Mediators of Inflammation*, *2021*, 5587428.
20. Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*, i884–i890.
21. Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics*, *26*, 589–595.
22. Magoč, T., & Salzberg, S. L. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, *27*, 2957–2963.
23. Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from illumina amplicon data. *Nature Methods*, *13*, 581–583.
24. Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, *42*, D633–642.
25. Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*, 996–998.
26. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, *27*, 2194–2200.
27. Lawley, B., & Tannock, G. W. (2017). Analysis of 16S rRNA gene amplicon sequences using the QIIME software package. *Methods in Molecular Biology*, *1537*, 153–163.
28. Wishart, D. S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., et al. (2022). HMDB 5.0: the human metabolome database for 2022. *Nucleic Acids Research*, *50*, D622–d631.
29. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, *45*, D353–d361.
30. Wang, X., & Zhi, Y. (2022). Altered urinary metabolomics in hereditary angioedema. *Metabolites*, *12*, 1140.
31. Yang, B., Zhang, C., Cheng, S., Li, G., Griebel, J., & Neuhaus, J. (2021). Novel metabolic signatures of prostate cancer revealed by (1)H-NMR metabolomics of urine. *Diagnostics (Basel)*, *11*, 149.
32. Wang, R., Kang, H., Zhang, X., Nie, Q., Wang, H., Wang, C., et al. (2022). Urinary metabolomics for discovering metabolic biomarkers of bladder cancer by UPLC-MS. *BMC Cancer*, *22*, 214.
33. Chong, J., Wishart, D. S., & Xia, J. (2019). Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. *Current Protocols in Bioinformatics*, *68*, e86.
34. Gao, J., Tarcea, V. G., Karnovsky, A., Mirel, B. R., Weymouth, T. E., Beecher, C. W., et al. (2010). Metscape: A cytoscape plug-in for visualizing and interpreting metabolomic data in the context of human metabolic networks. *Bioinformatics*, *26*, 971–973.
35. López-Díaz, J. Ó. M., Méndez-González, J., López-Serrano, P. M., Sánchez-Pérez, F. J., Méndez-Encina, F. M., Mendieta-Oviedo, R., et al. (2022). Dummy regression to predict dry fiber in *Agave lechuguilla* Torr. in two large-scale bioclimatic regions in Mexico. *PLoS ONE*, *17*, e0274641.
36. Liu, W., Zhang, R., Shu, R., Yu, J., Li, H., Long, H., et al. (2020). Study of the relationship between microbiome and colorectal cancer susceptibility using 16SrRNA sequencing. *BioMed Research International*, *2020*, 7828392.
37. Li, P., Shuai, P., Shen, S., Zheng, H., Sun, P., Zhang, R., et al. (2023). Perturbations in gut microbiota composition in patients with polycystic ovary syndrome: A systematic review and meta-analysis. *BMC Medicine*, *21*, 302.
38. Zhu, D., Ma, Y., Ding, S., Jiang, H., & Fang, J. (2018). Effects of melatonin on intestinal microbiota and oxidative stress in colitis mice. *BioMed Research International*, *2018*, 2607679.
39. Lin, Y.-S., Lin, C.-Y., Hung, C.-L., Chung, Y.-C., & Lee, K.-Z. (2015). GPU-UPGMA: High-performance computing for UPGMA algorithm based on graphics processing units. *Concurrency and Computation: Practice and Experience*, *27*, 3403–3414.
40. Li, C., Zhou, K., Xiao, N., Peng, M., & Tan, Z. (2022). The effect of qiweibaizhu powder crude polysaccharide on antibiotic-associated diarrhea mice is associated with restoring intestinal mucosal bacteria. *Frontiers in Nutrition*, *9*, 952647.
41. Shi, Y., Zhang, L., Do, K. A., Peterson, C. B., & Jenq, R. R. (2020). aPCoA: Covariate adjusted principal coordinates analysis. *Bioinformatics*, *36*, 4099–4101.
42. Gong, W., Zhu, Y., Shi, X., Zhang, W., & Wen, P. (2021). Influence of tissue type on the bacterial diversity and community in pork bacon. *Frontiers in Microbiology*, *12*, 799332.
43. Kwak, K. A., Cho, H. J., Yang, J. Y., & Park, Y. S. (2018). Current perspectives regarding stem cell-based therapy for liver cirrhosis. *Canadian Journal of Gastroenterology and Hepatology*, *2018*, 4197857.
44. Agarwal, R., & Wisnu, W. (2022). The effect of statin therapy on mortality in adult patients with liver cirrhosis: An evidence-based case report. *Acta Medica Indonesiana*, *54*, 491–499.

45. Teunis, C., Nieuwdorp, M., & Hanssen, N. (2022). Interactions between tryptophan metabolism, the gut microbiome and the immune system as potential drivers of non-alcoholic fatty liver disease (NAFLD) and metabolic diseases. *Metabolites*, *12*, 514.
46. Gradisteanu Pircalabioru, G., Liaw, J., Gundogdu, O., Corcionivoschi, N., Ilie, I., Oprea, L., et al. (2022). Effects of the lipid profile, type 2 diabetes and medication on the metabolic syndrome-associated gut microbiome. *International Journal of Molecular Sciences*, *23*, 7509.
47. Iversen, K. N., Dicksved, J., Zoki, C., Fristedt, R., Pelve, E. A., Langton, M., et al. (2022). The effects of high fiber rye, compared to refined wheat, on gut microbiota composition, plasma short chain fatty acids, and implications for weight loss and metabolic risk factors (the RyeWeight Study). *Nutrients*, *14*, 1669.
48. Vallianou, N., Christodoulatos, G. S., Karampela, I., Tsilingiris, D., Magkos, F., Stratigou, T., et al. (2021). Understanding the role of the gut microbiome and microbial metabolites in non-alcoholic fatty liver disease: current evidence and perspectives. *Biomolecules*, *12*, 1–56.
49. Pohl, K., Moodley, P., & Dhanda, A. (2022). The effect of increasing intestinal short-chain fatty acid concentration on gut permeability and liver injury in the context of liver disease: A systematic review. *Journal of Gastroenterology and Hepatology*, *37*, 1498–1506.
50. Singh, V., Yeoh, B. S., Chassaing, B., Xiao, X., Saha, P., Aguilera Olvera, R., et al. (2018). Dysregulated microbial fermentation of soluble fiber induces cholestatic liver cancer. *Cell*, *175*, 679–694. e622.
51. Hussain, S. K., Dong, T. S., Agopian, V., Pisegna, J. R., Durazo, F. A., Enayati, P., et al. (2020). Dietary protein, fiber and coffee are associated with small intestine microbiome composition and diversity in patients with liver cirrhosis. *Nutrients*, *12*, 1395.
52. He, Y., Cheng, B., Guo, B. J., Huang, Z., Qin, J. H., Wang, Q. Y., et al. (2023). Metabonomics and 16S rRNA gene sequencing to study the therapeutic mechanism of Danggui Sini decoction on collagen-induced rheumatoid arthritis rats with Cold Bi syndrome. *Journal of Pharmaceutical and Biomedical Analysis*, *222*, 115109.
53. Nascimento, W. M., Machiavelli, A., Ferreira, L. G. E., Cruz Silveira, L., de Azevedo, S. S. D., Bello, G., et al. (2021). Gut microbiome profiles and associated metabolic pathways in HIV-infected treatment-naïve patients. *Cells*, *10*, 385.
54. Adekunle, R. O., DeSilva, K., & Cartwright, E. J. (2020). Hepatitis C care continuum in a human immunodeficiency virus (HIV) positive cohort: data from the HIV Atlanta Veterans Affairs Cohort Study. *Open Forum Infectious Diseases*, *7*, ofaa085.
55. Liu, X., Mao, B., Gu, J., Wu, J., Cui, S., Wang, G., et al. (2021). Blautia—a new functional genus with potential probiotic properties? *Gut Microbes*, *13*, 1–21.
56. Ryvchin, R., Dubinsky, V., Rabinowitz, K., Wasserberg, N., Dotan, I., & Gophna, U. (2021). Alteration in urease-producing bacteria in the gut microbiomes of patients with inflammatory bowel diseases. *Journal of Crohn's and Colitis*, *15*, 2066–2077.
57. Rodriguez-Diaz, C., Taminiau, B., García-García, A., Cueto, A., Robles-Díaz, M., Ortega-Alonso, A., et al. (2022). Microbiota diversity in nonalcoholic fatty liver disease and in drug-induced liver injury. *Pharmacological Research*, *182*, 106348.
58. Liang, X. Q., Mai, P. Y., Qin, H., Li, S., Ou, W. J., Liang, J., et al. (2022). Integrated 16S rRNA sequencing and metabolomics analysis to investigate the antidepressant role of Yang-Xin-Jie-Yu decoction on microbe-gut-metabolite in chronic unpredictable mild stress-induced depression rat model. *Frontiers in Pharmacology*, *13*, 972351.
59. Li, R., Yi, X., Yang, J., Zhu, Z., Wang, Y., Liu, X., et al. (2022). Gut microbiome signatures in the progression of hepatitis B virus-induced liver disease. *Frontiers in Microbiology*, *13*, 916061.
60. Dang, J. T., Mocanu, V., Park, H., Laffin, M., Hotte, N., Karmali, S., et al. (2022). Roux-en-Y gastric bypass and sleeve gastrectomy induce substantial and persistent changes in microbial communities and metabolic pathways. *Gut Microbes*, *14*, 2050636.
61. Ganesan, R., Jeong, J. J., Kim, D. J., & Suk, K. T. (2022). Recent trends of microbiota-based microbial metabolites metabolism in liver disease. *Frontiers in Medicine (Lausanne)*, *9*, 841281.
62. Wei, X., Jiang, S., Zhao, X., Li, H., Lin, W., Li, B., et al. (2016). Community-metabolome correlations of gut microbiota from child-turcotte-pugh of A and B patients. *Frontiers in Microbiology*, *7*, 1856.
63. Huang, Z., Xie, N., Illes, P., Di Virgilio, F., Ulrich, H., Semyanov, A., et al. (2021). From purines to purinergic signalling: Molecular functions and human diseases. *Signal Transduction and Targeted Therapy*, *6*, 162.
64. Yin, H., Liu, N., & Chen, J. (2022). The role of the intestine in the development of hyperuricemia. *Frontiers in Immunology*, *13*, 845684.
65. Charni-Natan, M., Aloni-Grinstein, R., Osher, E., & Rotter, V. (2019). Liver and steroid hormones—can a touch of p53 make a difference? *Front Endocrinol (Lausanne)*, *10*, 374.
66. Robeva, R., Mladenović, D., Vesković, M., Hrnčić, D., Bjekić-Macut, J., Stanojlović, O., et al. (2021). The interplay between metabolic dysregulations and non-alcoholic fatty liver disease in women after menopause. *Maturitas*, *151*, 22–30.
67. Moctezuma-Velázquez, C., Low, G., Mourtzakis, M., Ma, M., Burak, K. W., Tandon, P., et al. (2018). Association between low testosterone levels and sarcopenia in cirrhosis: A cross-sectional study. *Annals of Hepatology*, *17*, 615–623.
68. Chu, X., Jin, Q., Chen, H., Wood, G. C., Petrick, A., Strodel, W., et al. (2018). CCL20 is up-regulated in non-alcoholic fatty liver disease fibrosis and is produced by hepatic stellate cells in response to fatty acid loading. *Journal of Translational Medicine*, *16*, 108.
69. Hliwa, A., Ramos-Molina, B., Laski, D., Mika, A., & Sledzinski, T. (2021). The role of fatty acids in non-alcoholic fatty liver disease progression: An update. *International Journal of Molecular Sciences*, *22*, 6900.
70. Zhang, J., Chen, Z., Yu, H., Lu, Y., Yu, W., Miao, M., et al. (2021). Anti-aging effects of a functional food via the action of gut microbiota and metabolites in aging mice. *Aging (Albany NY)*, *13*, 17880–17900.

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