Engineering 29 (2023) 110–119

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Medical Engineering—Article

An Evaluation Method of Human Gut Microbial Homeostasis by Testing Specific Fecal Microbiota

Zhongwen Wu [a,*,#], Xiaxia Pan [a,#], Yin Yuan [a,b,c,#], Pengcheng Lou [a,b,c], Lorina Gordejeva [d], Shuo Ni [e], Xiaofei Zhu [f], Bowen Liu [a], Lingyun Wu [g], Lanjuan Li [a,b,c,*], Bo Li [a,b,c,*]

[a] State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China
[b] Jinan Microecological Biomedicine Shandong Laboratory, Jinan 250000, China
[c] Research Units of Infectious Disease and Microecology, Chinese Academy of Medical Sciences, Beijing 100730, China
[d] School of Medicine, Zhejiang University, Hangzhou 310000, China
[e] Department of Orthopedic Surgery and Shanghai Institute of Microsurgery on Extremities, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Shanghai 200233, China
[f] Department of Infectious Diseases, Hangzhou Ninth People's Hospital, Hangzhou 310003, China
[g] Department of Radiation Oncology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China

ARTICLE INFO

Article history:
Received 4 January 2023
Revised 22 February 2023
Accepted 28 March 2023
Available online 20 April 2023

Keywords:
Gut microbiota
Machine learning
Microbial dysbiosis
Quantitative polymerase chain reaction
Chinese cohort

ABSTRACT

Research on microecology has been carried out with broad perspectives in recent decades, which has enabled a better understanding of the gut microbiota and its roles in human health and disease. It is of great significance to routinely acquire the status of the human gut microbiota; however, there is no method to evaluate the gut microbiome through small amounts of fecal microbes. In this study, we found ten predominant groups of gut bacteria that characterized the whole microbiome in the human gut from a large-sample Chinese cohort, constructed a real-time quantitative polymerase chain reaction (qPCR) method and developed a set of analytical approaches to detect these ten groups of predominant gut bacterial species with great maneuverability, efficiency, and quantitative features. Reference ranges for the ten predominant gut bacterial groups were established, and we found that the concentration and pairwise ratios of the ten predominant gut bacterial groups varied with age, indicating gut microbial dysbiosis. By comparing the detection results of liver cirrhosis (LC) patients with those of healthy control subjects, differences were then analyzed, and a classification model for the two groups was built by machine learning. Among the six established classification models, the model established by using the random forest algorithm achieved the highest area under the curve (AUC) value and sensitivity for predicting LC. This research enables easy, rapid, stable, and reliable testing and evaluation of the balance of the gut microbiota in the human body, which may contribute to clinical work.

1. Introduction

The human gut is a complex and intricate mini-ecosystem that mediates interactions between the host and the environment. The human gut contains trillions of microorganisms, such as bacteria, fungi, viruses, and other life forms, most of which exist in the colon. These microorganisms and the intestinal environment together constitute the intestinal microecology, and its diversity is the result of the coevolution of the intestinal microbiota and host [1–3]. The composition of the intestinal microecology is easily affected by many factors, such as diet [4,5], age [6,7], sex [7], genetics [8], and medicine [1].

The composition of the intestinal microbiota plays a fundamental role in regulating human health and diseases [9]. In addition to participating in human digestive function, the intestinal microbiota can affect human development, growth, and physiology, including organ development and morphogenesis and metabolism [10,11]. The intestinal microbiota also plays an indispensable role in the development and induction of the human immune system;

* Corresponding authors.
E-mail addresses: wuzhongwen@zju.edu.cn (Z. Wu), ljli@zju.edu.cn (L. Li), suon@zju.edu.cn (B. Li).
# These authors contributed equally to this work.

https://doi.org/10.1016/j.eng.2023.03.007

it regulates the differentiation of immune cells and the production of immune mediators to maintain interaction between the host and the intestinal microflora [12–14]. The destruction of the normal intestinal microbiota will increase the risk of infection and excessive proliferation of harmful pathogens and the occurrence of inflammatory diseases [12]. Qin et al. [15] studied the changes in the gut microbiota of patients with liver cirrhosis (LC) and found that at the genus level, *Bacteroides* was the dominant phylotype in both groups, but its abundance was significantly decreased in the LC group [15]. Therefore, it is very important for a healthy host to maintain microecological homeostasis. Indeed, an unbalanced intestinal microecology leads to the occurrence of a variety of diseases, including liver diseases, gastrointestinal diseases, metabolic diseases, and cardiovascular diseases [16–19].

Therefore, it is of great significance to routinely acquire the status of the human gut microbiota in a timely manner for the auxiliary evaluation of overall health and disease prediction. In the past, the coccus/bacillus (C/B) ratio was commonly used to reflect gut bacterial homeostasis [20,21] based on traditional methods, such as bacterial culture or microscopic examination. More recently, with the maturity of polymerase chain reaction (PCR) technology and the microbial sequencing technologies developed from it, more bacterial species can be detected, and the amount of some key bacterial species, bacterial ratios, or other indices have been selected to serve as balance indicators for the intestinal microecology. For example, it was found that the concentration of *Bifidobacterium*/concentration of Enterobacteriaceae (B/E) ratio can be used to judge the extent of intestinal microecology dysbiosis in the progression of liver diseases [22,23]. Low et al. [24] found that an increase in the abundance of *Klebsiella*/abundance of *Bifidobacterium* (K/B) ratio in early infants is a potential indicator of an increased risk of allergic disease. In addition, Ley et al. [25] first proposed that a higher abundance of Firmicutes/abundance of Bacteroidetes (F/B) ratio in the gut was likely to result in obesity. Microbial sequencing technologies have now become the mainstream method for studying the gut microbiota because the advantage of being high throughput screening. Microbial studies were mostly based on 16S ribosomal RNA (rRNA) gene sequencing [26–28] and metagenomic sequencing [29,30]. However, these two kinds of approaches are expensive and inefficient and are generally very cumbersome. As a result, microbial sequencing is not appropriate for routine detection of the gut microbiota in a large population. For rapid testing and simplicity as well as the absolute quantification function [31,32], a good alternative approach for carrying out routine gut bacterial detection is quantitative PCR (qPCR).

Based on previous studies [15,32], this study attempted to establish an effective and routine gut bacterial detection procedure; ten predominant bacterial groups in the human intestinal tract were detected via qPCR, including probiotics (*Lactobacillus* and *Bifidobacterium*), opportunistic pathogens (Enterobacteriaceae, *Enterococcus*, *Bacteroides*, and *Atopobium*), and other health-promoting symbiotic bacteria (*Faecalibacterium prausnitzii* (*F. prausnitzii*), *Clostridium butyricum* (*C. butyricum*), *Clostridium leptum* (*C. leptum*), and *Eubacterium rectale* (*E. rectale*)). These ten representative bacteria were selected from a large sample cohort of healthy people, which were considered potential indices for the evaluation of the whole human gut microbiome. Furthermore, we sought to obtain reference ranges of a healthy population and the changing patterns of the ten bacterial groups and their pairwise ratios with aging to pave the way for subsequent studies on large-sample disease-specific populations. Moreover, to investigate the probable difference in the qPCR detection results between healthy people and people with specific diseases, we evaluated patients with LC in comparison to healthy control (HC) subjects. To further test the capability of qPCR detection to distinguish people with diseases from the general population, we utilized a machine learning algorithm to mine the information of the detection results of the LC population and the HC population and built several classification models to finally select an optimal one.

2. Material and methods

2.1. Volunteer recruitment and sample collection

A total of 510 healthy subjects and 248 patients with LC were recruited; the inclusion and exclusion criteria are listed in Section S1 in Appendix A. Feces of the first defecation in the morning were collected in a clean plastic bag tucked into a disposable plastic bowl, and after defecation, the plastic bag was then fastened tightly to avoid urine pollution. To avoid interference factors, such as food residues, the softer part of the fresh feces was selected and loaded into cryopreservation tubes within half an hour. A DNA stabilizer (Invitek, Germany) was added, and the cryopreservation tubes were numbered before storage at −80 °C for preservation. On the morning of sample collection, a blood sample was collected from the patients and was subjected to routine blood tests, blood biochemistry tests, C-reactive protein (CRP) tests, and tests for other indicators. The basic information of all volunteers (Table 1), including age, sex, and body mass index (BMI), was registered. Volunteers with unqualified fecal samples and incomplete basic information were excluded; 500 healthy people and 244 LC patients were included. All of the work was performed according to guidelines approved by the Research Ethics Committee of the First Affiliated Hospital, College of Medicine, Zhejiang University (No. 2019-1026 and 2022-874).

2.2. qPCR assessment for major gut bacterial species

Microbial DNA was extracted from the feces of the 744 volunteers with a MegaBio soil/fecal genomic DNA purification kit (Bioer, Inc., China). The specific steps are described in Section S2 in Appendix A. The concentration of total DNA of fecal microorganisms (C_t) in each sample DNA eluate was detected using Nanodrop One (Thermo Fisher Scientific, USA). Ten predominant bacterial populations in the intestine were detected by real-time qPCR. Primers were synthesized by GenScript (China), and the primer information is listed in Table S1 in Appendix A. A ViiA™ PCR system (Applied Biosystems, Inc., USA) was used to conduct qPCR on 744 fecal bacterial DNA samples with a reaction volume of 20 μL, including 10 μL of SYBR Green PCR Master Mix (Zhongnuo Gene, Inc., China), 8 μL of primer pairs (0.2–0.6 $\mu mol \cdot L^{-1}$), and 2 μL of template DNA or 2 μL of distilled water (negative control). The reaction conditions are listed in Table S2 in Appendix A. Each reaction was performed in triplicate, and the cycle threshold (ΔCT) between repeats was required to be less than 0.5. Plasmid DNA standards containing the corresponding amplification fragment of each primer group were diluted in multiple gradient ratios and amplified with the bacterial DNA templates in the same PCR plate. The copies of the target bacteria in the DNA template were determined by comparison with the standard curve obtained from amplification of the corresponding bacterial DNA standards. The final concentration of the target bacteria was obtained by dividing the concentration of the target bacteria in the DNA template (N) by the total DNA concentration of fecal microorganisms in each sample DNA eluate and the volume of the template (V). The unit of the concentration is copies per nanogram total DNA of fecal microorganisms, hereinafter referred to as $copies \cdot ng^{-1}$. The formula is listed in Section S4 in Appendix A.

Table 1
Characteristics of the 744 volunteers.

Variable	Total (n = 744)	Healthy (n = 500)	Cirrhosis (n = 244)	Reference range	P value
Age (year)	55.00 (44.00–63.00)	54.00 (40.00–63.75)	56.50 (47.00–63.00)	—	0.0085
Sex (female/male)	268/476	225/275	43/201	—	< 0.0001
BMI ($kg·m^{-2}$)	23.24 (21.41–24.99)	22.99 (21.19–24.67)	23.32 (21.47–25.60)	18.50–24.00	0.0463
RBC ($\times10^{12}$ L^{-1})	4.510 (4.110–4.890)	4.660 (4.360–4.980)	3.885 (3.143–4.518)	3.680–5.130	< 0.0001
Hb ($g·L^{-1}$)	138.0 (122.0–149.0)	142.0 (131.0–151.0)	121.0 (98.3–141.0)	113.0–151.0	< 0.0001
WBC ($\times10^{9}$ L^{-1})	5.60 (4.60–6.90)	6.10 (5.10–7.10)	3.88 (2.62–5.50)	4.00–10.00	< 0.0001
Neu ($\times10^{9}$ L^{-1})	3.20 (2.50–4.02)	3.50 (2.90–4.20)	2.24 (1.47–3.38)	50.00–70.00	< 0.0001
Lym ($\times10^{9}$ L^{-1})	1.68 (1.25–2.11)	1.90 (1.57–2.29)	0.96 (0.61–1.45)	20.00–40.00	< 0.0001
Mono ($\times10^{9}$ L^{-1})	0.37 (0.27–0.48)	0.38 (0.29–0.48)	0.35 (0.24–0.48)	3.00–10.00	0.0450
Plt ($\times10^{9}$ L^{-1})	198.0 (113.0–242.0)	228.0 (193.0–262.0)	77.5 (49.0–119.0)	101.0–320.0	< 0.0001
ALT ($U·L^{-1}$)	18.00 (14.00–26.00)	16.00 (13.00–20.00)	24.00 (17.00–37.25)	7.00–40.00	< 0.0001
AST ($U·L^{-1}$)	22.00 (17.00–29.00)	20.00 (16.00–24.00)	31.00 (23.00–47.00)	13.00–35.00	< 0.0001
ALP ($U·L^{-1}$)	83.00 (66.00–108.00)	77.00 (61.00–99.00)	96.50 (77.25–135.0)	50.00–135.00	< 0.0001
GGT ($U·L^{-1}$)	23.00 (16.00–39.00)	20.00 (15.00–28.00)	37.00 (22.00–74.75)	7.00–45.00	< 0.0001
TP ($g·L^{-1}$)	70.65 ± 6.78	73.11 ± 4.17	65.65 ± 8.17	65.00–85.00	< 0.0001
ALB ($g·L^{-1}$)	44.30 (38.53–47.50)	45.85 (43.00–48.30)	35.10 (29.43–41.48)	40.00–55.00	< 0.0001
TBIL ($\mu mol·L^{-1}$)	12.80 (8.80–18.20)	10.55 (7.90–14.30)	23.05 (13.73–40.20)	0–21.00	< 0.0001
DBIL ($\mu mol·L^{-1}$)	4.30 (2.90–6.80)	3.50 (2.50–4.40)	10.15 (5.83–20.58)	0–8.00	< 0.0001
IBIL ($\mu mol·L^{-1}$)	7.90 (5.50–12.15)	7.10 (5.00–10.10)	11.10 (6.93–17.10)	3.00–14.00	< 0.0001
TBA ($\mu mol·L^{-1}$)	5.90 (3.70–13.68)	4.20 (3.60–6.80)	29.60 (9.83–72.60)	0–10.00	< 0.0001
Cr ($\mu mol·L^{-1}$)	71.00 (61.00–81.00)	72.00 (61.00–81.00)	71.00 (62.25–82.75)	41.00–73.00	0.3355
UA ($\mu mol·L^{-1}$)	301.0 (242.0–356.0)	310.0 (253.0–363.5)	276.0 (222.5–348.0)	155.0–357.0	0.0001
BUN ($mmol·L^{-1}$)	5.020 (4.020–6.020)	5.010 (4.068–5.900)	5.060 (3.833–6.403)	2.600–7.500	0.4746
TG ($mmol·L^{-1}$)	1.015 (0.740–1.390)	1.125 (0.830–1.528)	0.830 (0.610–1.120)	0.300–1.700	< 0.0001
TC ($mmol·L^{-1}$)	4.085 (3.460–4.710)	4.295 (3.843–4.935)	3.350 (2.660–4.040)	3.140–5.860	< 0.0001
Glu ($mmol·L^{-1}$)	4.800 (4.410–5.220)	4.750 (4.410–5.080)	4.925 (4.413–5.988)	3.900–6.100	< 0.0001
K^{+} ($mmol·L^{-1}$)	4.01 (3.72–4.32)	4.08 (3.80–4.39)	3.88 (3.57–4.14)	3.50–5.30	< 0.0001
Na^{+} ($mmol·L^{-1}$)	141.0 (139.0–143.0)	141.0 (139.0–143.0)	141.0 (139.0–142.0)	137.0–147.0	0.0520
Cl^{-} ($mmol·L^{-1}$)	104.0 (101.0–106.0)	103.0 (101.0–105.0)	105.0 (103.0–107.0)	99.0–110.0	< 0.0001
Ca^{2+} ($mmol·L^{-1}$)	2.23 (2.13–2.33)	2.27 (2.19–2.35)	2.13 (2.02–2.27)	2.25–2.75	< 0.0001
CRP ($mmol·L^{-1}$)	0.710 (0.310–2.280)	0.400 (0.190–0.620)	3.190 (1.290–5.485)	0–8.000	< 0.0001

RBC: red blood cell; Hb: hemoglobin; WBC: white blood cell; Neu: neutrophil; Lym: lymphocyte; Mono: monocyte; Plt: platelet; ALT: alanine aminotransferase; AST: aspartate aminotransferase; ALP: alkaline phosphatase; GGT: gamma-glutamyl transferase; TP: total protein; ALB: albumin; TBIL: total bilirubin; DBIL: direct bilirubin; IBIL: indirect bilirubin; TBA: total bile acid; Cr: creatinine; UA: uric acid; BUN: blood urea nitrogen; TG: triglyceride; TC: cholesterol; Glu: blood glucose. Statistical data of TP are presented as the mean ± standard error of mean. Other data in the table are presented as medians with interquartile ranges (IQRs).

2.3. Statistical analysis

All statistical analyses and plotting/graphical drawings in this study were performed using *R* script (version 4.1.2), Rstudio software (version 2022.07.1+554), and Origin 2021 (version 9.95). ggplot2 pack (version 3.3.5) was used for plotting/graphical drawings. The heatmap for correlation analysis was drawn by the corrplot package (version 0.92). The heatmap for comparing the quantity of the ten bacterial species was drawn by the pheatmap package (version 1.0.12). Data preprocessing was completed by the dplyr package (version 1.0.8). The processing of all outliers of the data adopts the box-plot method, and values less than Q1–1.5* interquartile range (IQR) or greater than Q3+1.5* IQR were determined to be outliers. These outliers were removed from the corresponding data groups. The reference range (bilateral) is calculated by taking a 95% confidence interval (95% CI) after removing outliers from the data of the healthy people. The normal distribution method was adopted for normally distributed data, and the upper and lower limits were $\overline{X}$ ± (1.96 × SD) (SD: standard deviation). The percentile method was adopted for nonnormally distributed data, and the upper and lower limits were $P_{2.5}$ and $P_{97.5}$, respectively. The Shapiro–Wilk test and Kolmogorov–Smirnov test were used to test the normal distribution of ten bacterial species. The Wilcoxon rank sum test was used for the analysis of differences between the two groups. The Kruskal–Wallis test was used to analyze the differences in nonnormal data among multiple groups. Spearman rank correlation analysis was used to analyze the correlation between the liver function indices and the ten associated bacterial species.

2.4. Machine learning

Machine learning algorithms were used to build models for distinguishing between cirrhotic and noncirrhotic samples. A total of 744 clinical samples (244 cirrhosis samples and 500 HC samples) from the same sample collection area were used as the dataset for constructing the model. The random sampling method was utilized to divide the data into training data and test data according to the ratio of 75–25. There were 408 training data points (183 cirrhosis data points and 225 noncirrhosis data points) and 136 test data points (61 cirrhosis data points and 75 noncirrhosis data points).

Through pretest screening, six machine learning methods were finally used to classify and model the data, including RF [33], GBM [34], AdaBoost [35], XGBoost [36], SVM_poly, and SVM_Gauss [37], with the first four methods belonging to ensemble learning [38].

The content of ten predominant bacterial species in the clinical samples and patient sex were selected as characteristics of the training model [7]. A tenfold cross-validation RF model was used to explore the importance of these features, and the mean decrease the Gini index was employed as a metric to determine the importance of these characteristics and their contribution to the model.

Repeated tenfold cross-validation (ten repeats) was applied to build and verify the model, and hyperparameter tuning was used to tune the six models. Through hyperparameter tuning, the optimal set of hyperparameter values of each of the six models was obtained, and the optimal model for each model was established according to the set of these values. The hyperparameter optimization of SVM_poly, and other models are shown in Appendix A. Finally, the area under the curve (AUC) was used to evaluate the quality of the trained model, and the model with the highest

average AUC value was selected as the optimal model. When overfitting occurred, the suboptimal model was chosen as the final result. The corresponding AUC value, sensitivity, and specificity of each model were obtained from the test data to validate the final model generated by the six machine learning algorithms.

The machine learning analysis process was implemented under the caret machine learning framework of R (version 6.0-93). RF was implemented using RF (version 4.7-1.1). Gradient upgrade was implemented using the gbm package (version 2.1.8.1). AdaBoost was implemented using the adabag package (version 4.2). XGBboost was implemented using the XGBoost package (version 1.6.0.1). Two support vector machine methods were implemented using the kernlab package (version 0.9-31).

2.5. Ethics

This study was approved by the Research Ethics Committee of the First Affiliated Hospital, College of Medicine, Zhejiang University (No. 2019-1026 and 2022-874). The patients/participants provided written informed consent to participate in this study. The research protocol complies with the ethical guidelines of the *1975 Declaration of Helsinki*.

3. Results

3.1. The predominant gut microbiota in healthy humans

A total of 500 healthy humans (275 males and 225 females) were included in this study. Healthy subjects were recruited based on strict exclusion criteria. After quality control, ten predominant gut bacterial species were detected by qPCR, and all data were statistically analyzed and processed to find the reference range for healthy people (Table 2). Furthermore, given that the structure and abundance of the gut microbiota are altered in people of different ages, we also analyzed and found the reference range for people of different ages (Table S3 in Appendix A). Healthy people of different ages were divided into five groups at intervals of 20 years: 0–20, 20–40, 40–60, 60–80, and 80–100 years. Except for *E. rectale*, the species showed significant differences among the five age groups. Among those species with significant differences, the abundance of *Atopobium* had the most significant change ($P = 5.7 \times 10^{-8}$), followed by *Enterococcus* ($P = 1.9 \times 10^{-7}$), and only *E. rectale* did not show significant differences with aging ($P = 0.081$).

3.2. Pairwise ratios are potential indicators of gut microbial homeostasis

Studies have reported that the abundance of Firmicutes decreases, and that of Bacteroidetes increases in almost all disease situations [39]. Therefore, we hypothesized that pairwise ratios are potential indicators for evaluating the balance of the gut microbiota. We compared the concentrations of the above ten bacterial species as logarithmic values (Fig. 1(a)). We used the pairwise ratio, and plotted the trends for a total of 45 ratios. Interestingly, we did find that the B/E ratio showed a typical trend of first decreasing and then increasing with age, displaying a U-shaped curve ($P = 0.021$). However, more studies are still needed. Moreover, the *Enterococcus*/Enterobacteriaceae (Ec/E) ratio, an indicator that significantly increases in critical patients [32], showed an increasing trend with age ($P = 0.00025$). The ratio of the two bacterial species with the largest difference between groups was *C. leptum/Bacteroides* ($P = 2.3 \times 10^{-7}$), and its increasing trend with age was also an obvious U-shaped curve that decreased first and then increased (Figs. 1(b) and (c)).

Table 2
Reference range of ten predominant gut bacterial species in overall healthy people.

Target bacteria	Reference range (copies·ng^{-1})	Reference range (lg)
F. prausnitzii	4.14×10^4–1.15×10^8	4.616530–8.059953
Enterococcus	4.96×10^2–2.47×10^5	2.695877–5.392757
Bacteroides	2.09×10^5–9.66×10^7	5.319932–7.984785
Lactobacillus	3.63×10^3–1.99×10^6	3.560451–6.297778
Bifidobacterium	1.63×10^2–6.36×10^6	2.212161–6.803233
C. butyricum	8.60×10^2–5.30×10^6	2.934515–6.724199
C. leptum	1.06×10^5–2.55×10^8	5.023961–8.407158
E. rectale	2.52×10^3–2.07×10^7	3.401409–7.316846
Atopobium	1.15×10^4–9.49×10^6	4.060623–6.977340
Enterobacteriaceae	4.35×10^3–5.93×10^7	3.638726–7.772846

3.3. An imbalanced gut microbiota in LC patients

To further verify whether the above ten predominant bacterial species may be used as indicators for microbial homeostasis, we collected and measured 244 fecal samples from LC patients. By comparing the results of healthy people with those of LC patients, it was found that there were significant differences between seven bacterial species (Fig. 2(a)); in contrast, *F. prausnitzii*, *C. leptum*, and *Atopobium* showed no differences between the two populations. Among the seven bacterial species, *Enterococcus*, *E. rectale*, and *Bacteroides* exhibited the largest differences ($P = 2.50 \times 10^{-12}$, 3.47×10^{-10}, and 6.57×10^{-10}, respectively). This was followed by *C. butyricum* ($P = 1.76 \times 10^{-5}$), *Lactobacillus* ($P = 1.13 \times 10^{-3}$), Enterobacteriaceae ($P = 5.59 \times 10^{-3}$), and *Bifidobacterium* ($P = 9.71 \times 10^{-3}$), which is consistent with the results visualized in the heatmap (Fig. 2(b)). In addition, concentrations of *Bifidobacterium*, *E. rectale*, Enterobacteriaceae, and *C. butyricum* were relatively stable in both healthy individuals and cirrhosis patients, with few outliers. The number of outliers of the four bacterial species in cirrhosis patients and healthy people was (1/244, 1/500), (0/244, 2/500), (4/244, 1/500), and (2/244, 3/500), respectively.

Furthermore, compared to healthy individuals, the ratios of Ec/*E. rectale* ($P = 3.08 \times 10^{-24}$), *C. leptum/Bacteroides* ($P = 2.14 \times 10^{-18}$), and *C. butyricum/E. rectale* ($P = 2.54 \times 10^{-18}$) were significantly different in cirrhosis patients, indicating that the gut microbiota was imbalanced in cirrhosis patients (Fig. 2(c)). These results are consistent with our previous work [15]. However, the B/E ratio remained almost unchanged in this work.

3.4. The gut microbiota is associated with the severity of LC

Correlation analysis of the gut microbiota and liver function indicators revealed that the serum levels of alternate (ALT), albumin (ALB), direct bilirubin (DBIL), triglyceride (TG), and total bile acid (TBA) positively correlated with the abundance of *Bacteroides* but negatively correlated with that of the other nine bacterial species. The serum level of alkaline phosphatase (ALP) negatively correlated with the abundance of *Bacteroides* but positively correlated with that of the other nine bacterial species. Negative correlations between the abundance of Enterobacteriaceae and the B/E ratio were also found (Fig. 3).

3.5. Multiple machine learning models to distinguish and predict healthy people and patients with LC

Furthermore, we constructed multiple machine learning models to analyze the above results of the gut microbiota and to distinguish HC subjects from LC patients (Fig. 4(a)). By tuning the hyperparameters of the six models, the optimal combination of hyperparameters was obtained. The relationship between the hyperparameters of the SVM_poly model based on the polynomial kernel and the model performance is shown in Fig. 4(b). From the

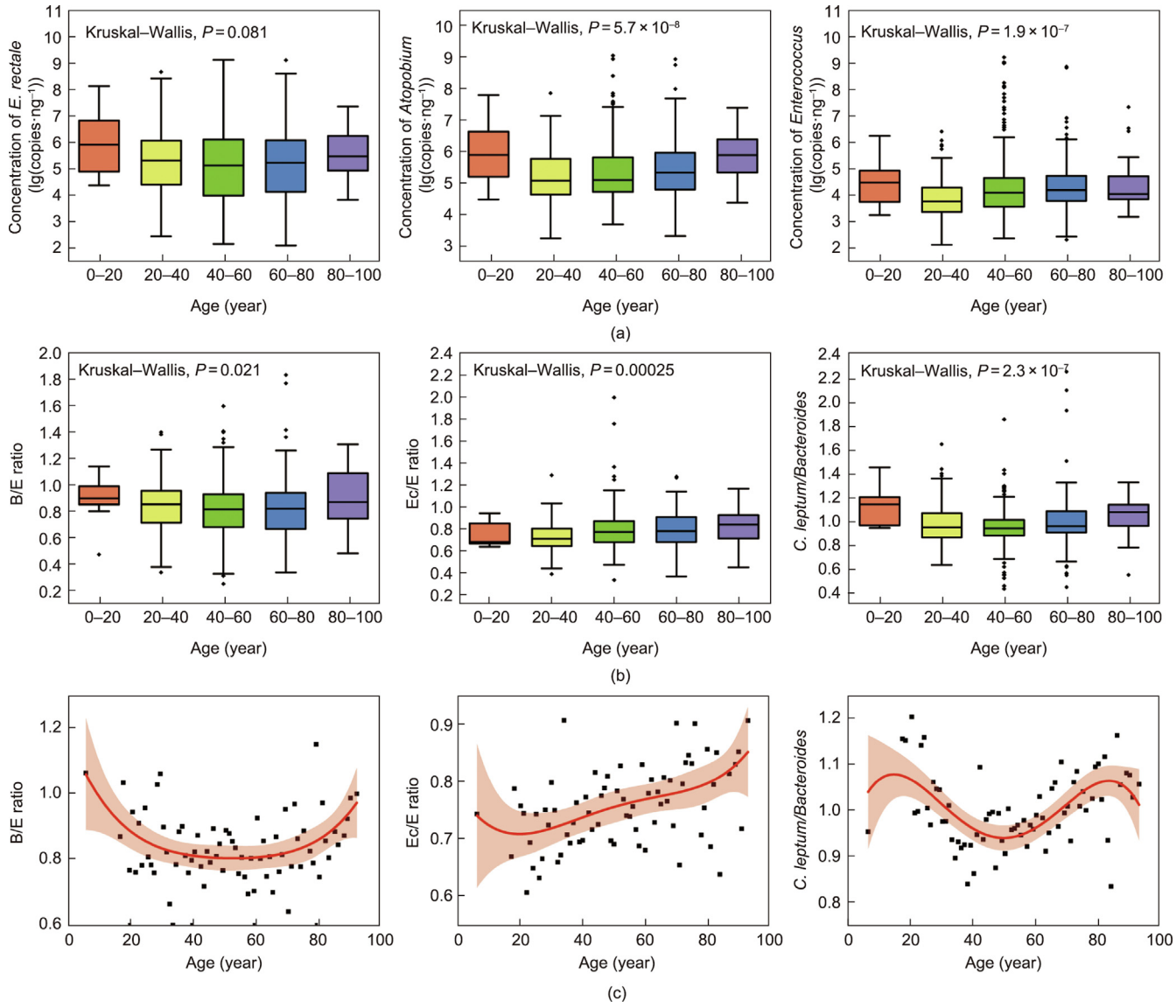

Fig. 1. The gut microbiota in healthy people. (a) The differences in ten bacterial species in different age groups. From left to right, the differences in *E. rectale*, *Atopobium*, and *Enterococcus* among different age groups are shown. (b) The difference in the pairwise ratios of the ten bacterial species in different age groups, and the ordinate represents the pairwise ratios of the ten bacterial species. From left to right is the difference in the B/E ratio, *Enterococcus*/Enterobacteriaceae (Ec/E) ratio, and *C. leptum/Bacteroides* ratio with age. (c) The relationship between the pairwise ratio of ten bacterial species and age. The ordinate represents the pairwise ratios of the ten bacterial species. The black dots represent the mean of the ratios of the two bacterial species at each age. The red curve is the fitted curve of these points, implemented using polynomial fitting, and the light red area is the 95% CI of the fitted curve. From left to right, the relationships between the B/E ratio and age, the Ec/E ratio and age, and the *C. leptum/Bacteroides* ratio and age are shown. The boxes represent the IQR, from the first to the third quartile, and the lines within the boxes represent the median. Whiskers show the minimum and maximum values within 1.5 times IQR outside the first and third quartiles. Black dots represent outliers outside the whisker.

figure, the optimal combination of the hyperparameters is polynomial degree = 2, scale = 0.100, and C = 0.75. Relationship diagrams between the estimates of performance and the tuning parameters of the other five models are shown in Figs. S1–S3 in Appendix A.

The training results of the six models showed the XGBoost model to have the best training results, with an average AUC value reaching 0.9376 (95% CI, 0.9158–0.9595). This was followed by SVM_Gauss and SVM_poly, with average AUCs reaching 0.9050 (95% CI, 0.8754–0.9346) and 0.9040 (95% CI, 0.8749–0.9331), respectively. The worst training result was observed for the RF model, but the average AUC also reached 0.8746 (95% CI, 0.8414–0.9078) (Fig. 4(c)).

Then, the machine learning model was also used to analyze 50 real clinical samples. The test results (Table 3) show that the six models achieved good prediction results, among which RF achieved the best AUC value, reaching 0.8776 (95% CI, 0.8159–0.9393). This was followed by XGBoost and Adaboost, with AUC values reaching 0.8726 (95% CI, 0.8097–0.9355) and 0.8630 (95% CI, 0.7969–0.9290), respectively. The three models with the highest sensitivity were RF, XGBoost, and Adaboost, with values of 0.8361, 0.7869, and 0.7869, respectively. The three models with the highest specificity were XGBoost, Adaboost, and GBM, with values of 0.8667, 0.8533, and 0.8400, respectively (Fig. 4(c) and Table S4 in Appendix A).

4. Discussion

Gut microbiota dysbiosis is an abnormal change in the number, proportion, and species of the normal microbiota in the gut that

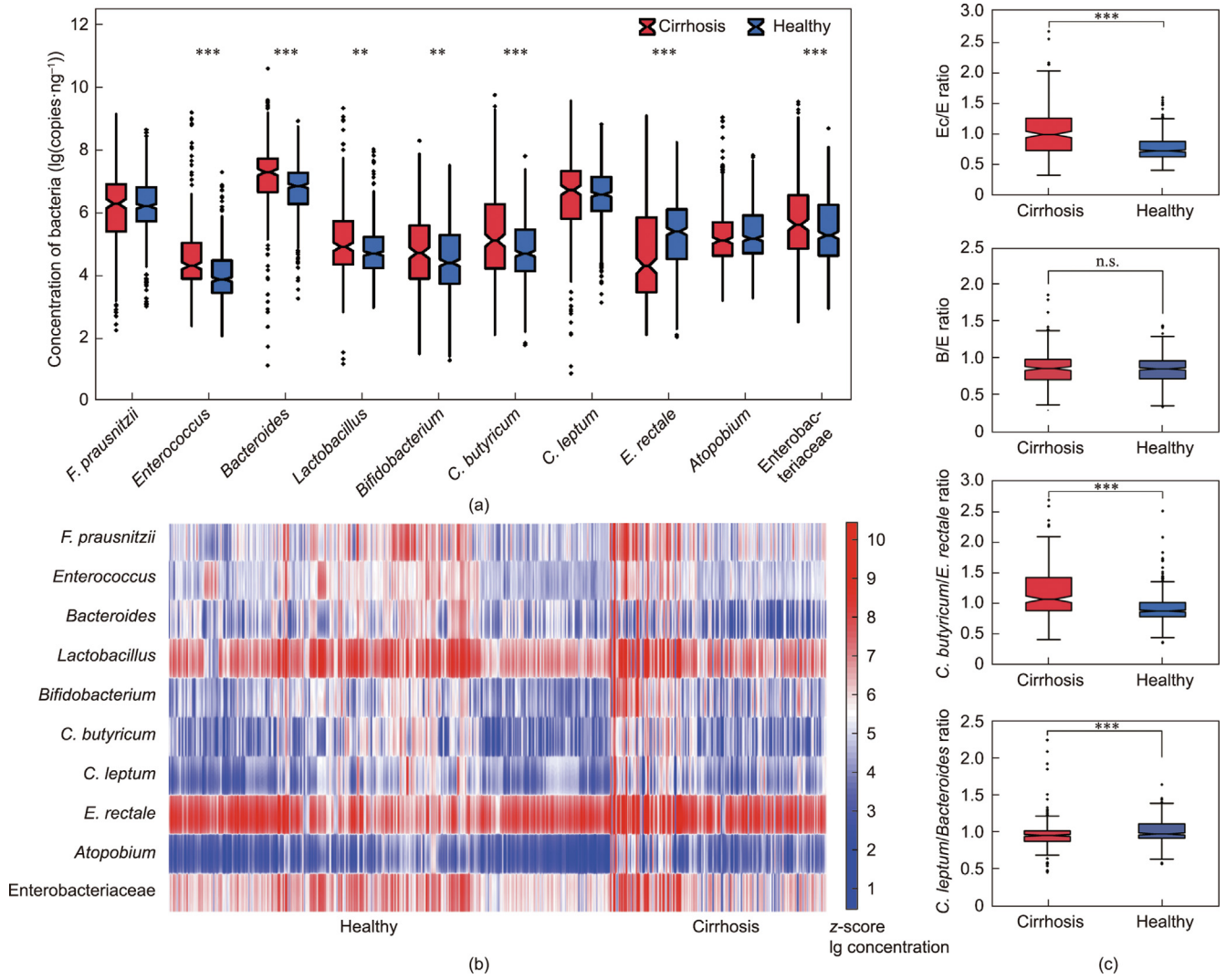

Fig. 2. Gut microbiota in LC patients and HC subjects. (a) The abundances of the ten predominant gut bacterial species in cirrhosis and HC subjects. The ordinate represents the lg value of the concentrations of the bacterial species. The boxes represent the IQR, from the first to the third quartiles, and the lines within the boxes represent the median. Whiskers show the minimum and maximum values within 1.5 times IQR outside the first and third quartiles. Black dots represent outliers outside the whisker. The cut represents the 95% CI of the median. (b) Heatmap of the difference in the content of ten predominant bacterial species between the LC group and the HC group. The horizontal axis represents the samples, which were divided into two groups according to LC and HC. There were 244 samples in the LC group and 500 samples in the HC group. The vertical axis represents the ten predominant bacteria in the sample. The different colors represent the lg values of the concentrations of the ten bacterial species in each sample. Darker red represents higher content, and darker blue represents lower content. (c) The pairwise ratios of gut bacteria between the cirrhosis and HC groups. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. n.s.: no significance.

affects human health, leading to a series of abnormal physiological and pathological phenomena [40,41]. Homeostasis between the gut microbiota and the host immune system is compromised when the former is imbalanced [42]. The gut microbiota changes significantly during human aging [43]. Zhang et al. [7] found consistent changes in gut microbiota during aging in humans, as characterized by increased α diversity. The abundance of multiple members of the oral microbiota, Enterobacteriaceae and *Clostridia,* which are short-chain fatty acid (SCFA) producers, increased with age. They also found that several *Bifidobacterium* species (*B. breve*, *B. bifidum*, *B. longum*, and *B. adolescentis*) negatively correlated with age. Biagi et al. [44] revealed that the cumulative abundance of symbiotic bacterial taxa (mostly belonging to the dominant Ruminococcaceae, Lachnospiraceae, and Bacteroidaceae families) decreased with age. Nevertheless, health-related *Akkermansia, Bifidobacterium*, and Christensenellaceae were enriched in elderly individuals, especially semisupercentenarians (105–109 years old). Another study also found a decrease in the abundance of core genera in the gut microbiota of healthy aging individuals, particularly *Bacteroides*, which is associated with longer life expectancy [45]. While different studies have identified gut microbiota changes during human aging, there is no consensus on the changing patterns of the gut microbiome during aging. Therefore, in this study, we aimed to establish the reference range for all ages as well as for different age groups of healthy people based on the results of a large cohort and to determine underlying characteristics during healthy aging.

Based on our previous work, we detected ten predominant gut bacterial species in each healthy individual by qPCR. The ten predominant gut bacterial species selected for this study were found to play an important role in maintaining intestinal homeostasis. For instance, *Lactobacillus* and *Bifidobacterium* are important for regulating immunity and maintaining gut barrier function [46–48], and *F. prausnitzii*, *C. butyricum*, *C. leptum*, and *E. rectale* can produce SCFAs, which are important in inhibiting the overgrowth of opportunistic pathogens, maintaining the integrity of

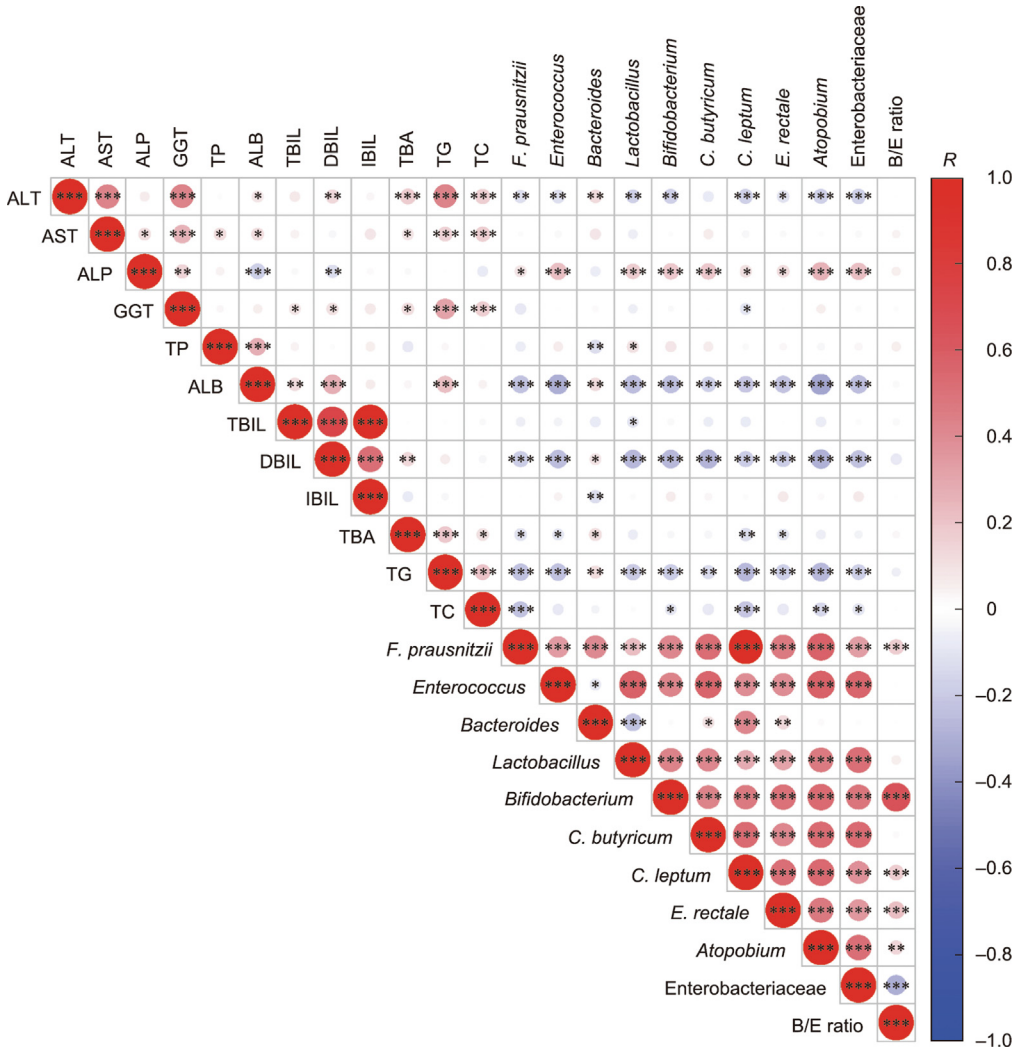

Fig. 3. Correlation between liver function indicators and the gut microbiota. The color depth and size of the dots show the strength of the correlation (*R* value). The darker the red is, the stronger the positive correlation is. The darker the blue is, the stronger the negative correlation is. The white represents no correlation. The larger the dot is, the stronger the correlation is. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

the intestinal epithelial barrier, and enhancing immunity [49–51]. Our results revealed that the concentrations of the ten predominant gut bacterial species changed differently with age. The healthy population was divided into five groups at intervals of 20 years (0–20, 20–40, 40–60, 60–80, and 80–100). Nine of the ten predominant bacterial species showed some variation in all five age groups. *Atopobium* showed the most significant difference ($P = 5.7 \times 10^{-8}$); this was followed by *Enterococcus* ($P = 1.9 \times 10^{-7}$). In addition, except for *Bacteroides*, the concentration of which increased and then decreased with age, the other bacteria showed a general trend of decreasing and then increasing. Most large-sample studies have been based on 16S rRNA or metagenomic sequencing, but one of the main limitations of sequencing is that taxa can only be assigned according to the sequence of a single region in the bacterial genome, and only relative abundance results are obtained, which makes it difficult to generate a stable and reliable reference value [31,52]. In contrast, qPCR can be used to quantitatively detect bacterial concentration by calibration with known concentrations of standard substances, which has the characteristics of good repeatability, rapid results, and simple operation. Timely detection of intestinal microbial dysbiosis is helpful for clinicians to achieve early diagnosis and treatment and to improve prognosis [32,53].

Our previous study analyzed the changes in the gut microbiota of patients with LC. We found that at the genus level, *Bacteroides* was the dominant phylotype in both groups, but its abundance was significantly decreased in the LC group [15]. Of the remaining genera, *Veillonella*, *Streptococcus*, *Clostridium*, and *Prevotella* were enriched in the LC group, while *Eubacterium* and *Alistipes* were dominant in the HC subjects. Of the species that decreased the most in abundance in the LC group, twelve were *Bacteroidetes* and seven were Firmicutes, specifically from the order *Clostridiales*. However, the cost of metagene sequencing is high, and the amount of data is large, which requires a large amount of computing resources to perform analysis. Therefore, in this study, we aimed to characterize the changes in the gut microbiota of cirrhosis patients using a smaller number of bacteria. Moreover, we also constructed multiple machine learning models to further analyze the microecological results. Interestingly, we found that the prediction results of the classification models built by four algorithms under the framework of ensemble learning were better than those of SVM and some other machine learning algorithms. Among them, the classification model of RF achieved the highest AUC value and sensitivity, and the classification model of XGBoost achieved the second highest AUC value and the highest specificity, which were relatively better than those of the other four models. Comparing

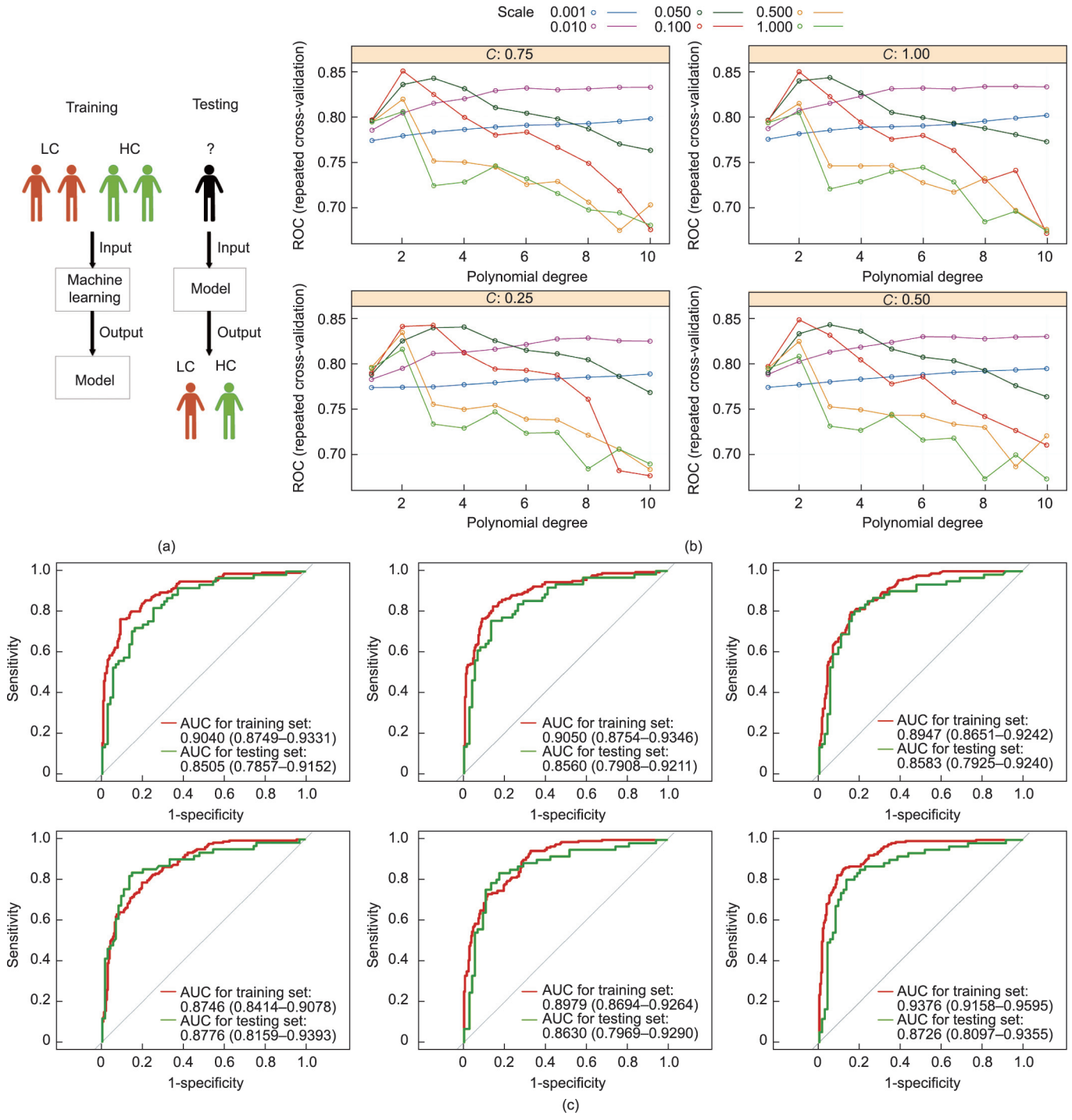

Fig. 4. Multiple machine learning models. (a) The machine learning analysis process is divided into two steps: training and testing. The training is to learn the input classified HC and LC data through machine learning methods and finally obtain the optimized classification model. The test classifies the unclassified data with the trained model and evaluates the quality of the model through relevant indicators. (b) Hyperparameter tuning of the classification model established by support vector machine based on polynomial kernel by adjusting the values of *C* in SVM, scale, and polynomial degree to find the hyperparameter combination with the highest average AUC value. The final combination was polynomial degree = 2, scale = 0.100, *C* = 0.75. (c) AUC values of six machine learning algorithms for the training and test sets. The upper left is the AUC value of the SVM_poly algorithm; the upper middle is the AUC value of the SVM_Gauss algorithm; the upper right is the AUC value of the GBM algorithm; the bottom left is the AUC value of the RF algorithm; the bottom middle is the AUC value of the AdaBoost algorithm; the bottom right is the AUC value of the XGBoost algorithm. ROC: receiver operating characteristic.

the two classification models, the RF model had a higher AUC value and sensitivity and better identified patients with cirrhosis or suspected cirrhosis, but there were more false-positives. The XGBoost model showed higher specificity and was better able to identify healthy people. However, there were more false-negatives, which may result in unrecognized cirrhosis patients failing to visit the hospital in a timely manner, delaying diagnosis and treatment. The RF model with the highest sensitivity was the best choice. Although the RF model may cause some healthy people to be misdiagnosed, the benefits of finding more patients with true cirrhosis

Table 3
Test results of six machine learning algorithms.

Algorithm	AUC	Sensitivity	Specificity
SVM_poly	0.8505	0.7377	0.7867
SVM_Gauss	0.8560	0.7705	0.8000
GBM	0.8583	0.7541	0.8400
RF	0.8776	0.8361	0.8000
AdaBoost	0.8630	0.7869	0.8533
XGBoost	0.8726	0.7869	0.8667

in a timely manner outweigh the risks. However, geography and location have strong influences on the human gut microbiome, and geographical differences limit the application of the reference range of the healthy gut microbiome and disease models. To verify the prevalence of microbiota differences between healthy or disease states, standardized experimental protocols, regional study designs, and extensive sampling are needed, and geography as a feature needs to be added to the model. Different predictive models need to be trained for different geographical regions [54,55]. Considering the influence of region on the human microbiota, samples from different regions will be collected in the future to re-establish the classification model. We will adopt two strategies: ① add region as an additional feature to the model and ② establish classification models according to different regions relatively independently and select a better classification model according to the results.

The combination of qPCR and machine learning makes is an easy, rapid, stable, and reliable method to test and analyze the gut microbiome of the human body. With the inclusion of new samples, our healthy population cohort will be subsequently expanded, the reference range for the healthy population results will be updated as the sample size expands, and the capability of relevant machine learning models to predict illnesses will be improved through continuous learning and prediction, which may contribute to clinical work.

5. Conclusions

Ten kinds of predominant gut bacterial species that characterize the whole microbiome in the human gut were found from a large-sample Chinese cohort. We established the reference ranges of these ten predominant gut bacterial groups by detecting their concentrations by qPCR and discovered the changing patterns of the ten bacterial groups with aging and disease. In addition, we utilized machine learning algorithms to deeply extract differential information from the detection results and built and selected a reliable classification model for predicting LC. This study revealed that it is highly necessary to describe and predict the changes in gut microbiota in a healthy Chinese population with a small amount of information. Based on this healthy range, it can be widely used to predict and describe intestinal microecological dysbiosis in various diseases. However, more new theoretical models and clinical practice are still needed in future work.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFC2000500), the Fundamental Research Funds for the Central Universities (2022ZFJH003), the Independent Task of State Key Laboratory for Diagnosis and Treatment of Infectious Diseases (2022zz22), the National Natural Science Foundation of China (81703430, 32170058, and 82200994), the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2019-I2M-5-045), and the Research Project of Jinan Microecological Biomedicine Shandong Laboratory (JNL-2022051B).

Special thanks to KW Liu for self-confidence and spiritual support and Zhongnuo Gene, Inc., for bacteria detection and model establishment.

Compliance with ethics guidelines

Zhongwen Wu, Xiaxia Pan, Yin Yuan, Pengcheng Lou, Lorina Gordejeva, Shuo Ni, Xiaofei Zhu, Bowen Liu, Lingyun Wu, Lanjuan Li, and Bo Li declare that they have no conflicts of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2023.03.007.

References

[1] Weersma RK, Zhernakova A, Fu J. Interaction between drugs and the gut microbiome. Gut 2020;69(8):1510–9.
[2] Ley RE, Peterson DA, Gordon JI. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell 2006;124(4):837–48.
[3] Brody H. The gut microbiome. Nature 2020;577(7792):S5.
[4] Ghosh TS, Rampelli S, Jeffery IB, Santoro A, Neto M, Capri M, et al. Mediterranean diet intervention alters the gut microbiome in older people reducing frailty and improving health status: the NU-AGE 1-year dietary intervention across five European countries. Gut 2020;69(7):1218–28.
[5] Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S, et al. Gut microbiota composition correlates with diet and health in the elderly. Nature 2012;488(7410):178–84.
[6] Claesson MJ, Cusack S, O'Sullivan O, Greene-Diniz R, de Weerd H, Flannery E, et al. Composition, variability, and temporal stability of the intestinal microbiota of the elderly. Proc Natl Acad Sci USA 2011;108(Suppl 1):4586–91.
[7] Zhang X, Zhong H, Li Y, Shi Z, Ren H, Zhang Z, et al. Sex- and age-related trajectories of the adult human gut microbiota shared across populations of different ethnicities. Nature Aging 2021;1:87–100.
[8] Burberry A, Wells MF, Limone F, Couto A, Smith KS, Keaney J, et al. C9orf72 suppresses systemic and neural inflammation induced by gut bacteria. Nature 2020;582(7810):89–94.
[9] Heintz-Buschart A, Wilmes P. Human gut microbiome: function matters. Trends Microbiol 2018;26(7):563–74.
[10] Karasov WH, Martinez del Rio C, Caviedes-Vidal E. Ecological physiology of diet and digestive systems. Annu Rev Physiol 2011;73:69–93.
[11] Sommer F, Backhed F. The gut microbiota—masters of host development and physiology. Nat Rev Microbiol 2013;11(4):227–38.
[12] Kamada N, Seo SU, Chen GY, Nunez G. Role of the gut microbiota in immunity and inflammatory disease. Nat Rev Immunol 2013;13(5):321–35.
[13] Smith PM, Garrett WS. The gut microbiota and mucosal T cells. Front Microbiol 2011;2:111.
[14] Kawamoto S, Tran TH, Maruya M, Suzuki K, Doi Y, Tsutsui Y, et al. The inhibitory receptor PD-1 regulates IgA selection and bacterial composition in the gut. Science 2012;336(6080):485–9.
[15] Qin N, Yang F, Li A, Prifti E, Chen Y, Shao L, et al. Alterations of the human gut microbiome in liver cirrhosis. Nature 2014;513(7516):59–64.
[16] Frost F, Kacprowski T, Ruhlemann M, Pietzner M, Bang C, Franke A, et al. Long-term instability of the intestinal microbiome is associated with metabolic liver disease, low microbiota diversity, diabetes mellitus and impaired exocrine pancreatic function. Gut 2021;70(3):522–30.
[17] Lynch SV, Pedersen O. The human intestinal microbiome in health and disease. N Engl J Med 2016;375(24):2369–79.
[18] Schnabl B, Brenner DA. Interactions between the intestinal microbiome and liver diseases. Gastroenterology 2014;146(6):1513–24.
[19] Bajaj JS, Heuman DM, Hylemon PB, Sanyal AJ, White MB, Monteith P, et al. Altered profile of human gut microbiome is associated with cirrhosis and its complications. J Hepatol 2014;60(5):940–7.
[20] Yu X, Zhang X, Jin H, Wu Z, Yan C, Liu Z, et al. Zhengganxifeng decoction affects gut microbiota and reduces blood pressure via renin-angiotensin system. Biol Pharm Bull 2019;42(9):1482–90.
[21] Zhang Y, Wang Y, Ke B, Du J. TMAO: how gut microbiota contributes to heart failure. Transl Res 2021;228:109–25.
[22] Lu H, Wu Z, Xu W, Yang J, Chen Y, Li L. Intestinal microbiota was assessed in cirrhotic patients with hepatitis B virus infection. Intestinal microbiota of HBV cirrhotic patients. Microb Ecol 2011;61(3):693–703.
[23] Li L. Infectious microecology: theory and applications. Hangzhou: Zhejiang University Press; 2014.

[24] Low JSY, Soh SE, Lee YK, Kwek KYC, Holbrook JD, van der Beek EM, et al. Ratio of *Klebsiella/Bifidobacterium* in early life correlates with later development of paediatric allergy. Benef Microbes 2017;8(5):681–95.
[25] Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. Nature 2006;444(7122):1022–3.
[26] Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, et al. A core gut microbiome in obese and lean twins. Nature 2009;457(7228):480–4.
[27] Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, et al. The long-term stability of the human gut microbiota. Science 2013;341(6141):1237439.
[28] Bender JM, Li F, Adisetiyo H, Lee D, Zabih S, Hung L, et al. Quantification of variation and the impact of biomass in targeted 16S rRNA gene sequencing studies. Microbiome 2018;6(1):155.
[29] Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature 2012;490(7418):55–60.
[30] Lepage P, Leclerc MC, Joossens M, Mondot S, Blottiere HM, Raes J, et al. A metagenomic insight into our gut's microbiome. Gut 2013;62(1):146–58.
[31] Poretsky R, Rodriguez RL, Luo C, Tsementzi D, Konstantinidis KT. Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. PLoS One 2014;9(4):e93827.
[32] Tang L, Gu S, Gong Y, Li B, Lu H, Li Q, et al. Clinical significance of the correlation between changes in the major intestinal bacteria species and COVID-19 severity. Engineering 2020;6(10):1178–84.
[33] Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
[34] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001;29(5):1189–232.
[35] Yoav Freund RES. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 1997;55(1):119–39.
[36] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. Assoc Comp Machinery 2016:785–94.
[37] Cortes C, Vapnik V. Support vector networks. Mach Learn 1995;20(3):273–97.
[38] Dietterich TG. Ensemble methods in machine learning. Berlin Heidelberg: Springer; 2000. p. 1–15.
[39] Su Q, Liu Q, Lau RI, Zhang J, Xu Z, Yeoh YK, et al. Faecal microbiome-based machine learning for multi-class disease diagnosis. Nat Commun 2022;13(1):6818.
[40] Shin NR, Whon TW, Bae JW. Proteobacteria: microbial signature of dysbiosis in gut microbiota. Trends Biotechnol 2015;33(9):496–503.
[41] Sun MF, Shen YQ. Dysbiosis of gut microbiota and microbial metabolites in Parkinson's disease. Ageing Res Rev 2018:4553–61.
[42] Hooper LV, Littman DR, Macpherson AJ. Interactions between the microbiota and the immune system. Science 2012;336(6086):1268–73.
[43] Badal VD, Vaccariello ED, Murray ER, Yu KE, Knight R, Jeste DV, et al. The gut microbiome, aging, and longevity: a systematic review. Nutrients 2020;12(12).
[44] Biagi E, Franceschi C, Rampelli S, Severgnini M, Ostan R, Turroni S, et al. Gut microbiota and extreme longevity. Curr Biol 2016;26(11):1480–5.
[45] Wilmanski T, Diener C, Rappaport N, Patwardhan S, Wiedrick J, Lapidus J, et al. Gut microbiome pattern reflects healthy ageing and predicts survival in humans. Nat Metab 2021;3(2):274–86.
[46] Yu Q, Yuan L, Deng J, Yang Q. *Lactobacillus* protects the integrity of intestinal epithelial barrier damaged by pathogenic bacteria. Front Cell Infect Microbiol 2015;5:26.
[47] Peng X, Ed-Dra A, Song Y, Elbediwi M, Nambiar RB, Zhou X, et al. *Lacticaseibacillus rhamnosus* alleviates intestinal inflammation and promotes microbiota-mediated protection against Salmonella fatal infections. Front Immunol 2022;13:973224.
[48] Salazar N, Gueimonde M, de Los RG, Ruas-Madiedo P. Exopolysaccharides produced by lactic acid bacteria and bifidobacteria as fermentable substrates by the intestinal microbiota. Crit Rev Food Sci Nutr 2016;56(9):1440–53.
[49] Parada Venegas D, De la Fuente MK, Landskron G, Gonzalez MJ, Quera R, Dijkstra G, et al. Short chain fatty acids (SCFAs)-mediated gut epithelial and immune regulation and its relevance for inflammatory bowel diseases. Front Immunol 2019;10:277.
[50] Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, Bunger MK, et al. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. Cell Metab 2011;13(5):517–26.
[51] Park J, Kim M, Kang SG, Jannasch AH, Cooper B, Patterson J, et al. Short-chain fatty acids induce both effector and regulatory T cells by suppression of histone deacetylases and regulation of the mTOR-S6K pathway. Mucosal Immunol 2015;8(1):80–93.
[52] Laudadio I, Fulci V, Palone F, Stronati L, Cucchiara S, Carissimi C. Quantitative assessment of shotgun metagenomics and 16S rDNA amplicon sequencing in the study of human gut microbiome. OMICS 2018;22(4):248–54.
[53] Bartosch S, Fite A, Macfarlane GT, McMurdo ME. Characterization of bacterial communities in feces from healthy elderly volunteers and hospitalized elderly patients by using real-time PCR and effects of antibiotic treatment on the fecal microbiota. Appl Environ Microbiol 2004;70(6):3575–81.
[54] He Y, Wu W, Zheng HM, Li P, McDonald D, Sheng HF, et al. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. Nat Med 2018;24(10):1532–5.
[55] Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature 2012;486(7402):222–7.