



ELSEVIER

Contents lists available at ScienceDirect

# Engineering

journal homepage: [www.elsevier.com/locate/eng](http://www.elsevier.com/locate/eng)



Research  
Medical Engineering—Article

## 一种通过检测特定肠道菌群来评估人体肠道微生态平衡的方法

吴仲文<sup>a,\*</sup>, 潘厦厦<sup>a,#</sup>, 袁音<sup>a,b,c,#</sup>, 楼鹏程<sup>a,b,c</sup>, Lorina Gordejewa<sup>d</sup>, 倪硕<sup>e</sup>, 朱晓飞<sup>f</sup>, 刘博文<sup>a</sup>, 吴凌云<sup>g</sup>,  
李兰娟<sup>a,b,c,\*</sup>, 李博<sup>a,b,c,\*</sup>

<sup>a</sup> 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

<sup>b</sup> Jinan Microecological Biomedicine Shandong Laboratory, Jinan 250000, China

<sup>c</sup> Research Units of Infectious Disease and Microecology, Chinese Academy of Medical Sciences, Beijing 100730, China

<sup>d</sup> School of Medicine, Zhejiang University, Hangzhou 310000, China

<sup>e</sup> 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

<sup>f</sup> Department of Infectious Diseases, Hangzhou Ninth People's Hospital, Hangzhou 310003, China

<sup>g</sup> 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

#### 关键词

肠道微生态

机器学习

微生态失衡

定量聚合酶链反应

中国队列

### 摘要

近几十年来,微生态学的研究涉及领域广泛,人们能更好地了解肠道菌群及其在人类健康和疾病中的作用。常规获取人体肠道菌群状态具有重要意义;然而,目前还没有通过少量粪便微生物来评估肠道微生物组的方法。在本研究中,我们从中国一个大样本人群队列中发现了能够表征整个肠道菌群的10种优势菌,并构建了一套可操作性强、效率高、定量特征好的检测方法,用来分析这10种优势菌。我们建立了10种肠道优势菌的参考范围,发现10种肠道优势菌的丰度和配对比值随年龄变化而变化,提示存在肠道菌群失调。通过比较肝硬化患者与健康对照者的检测结果,分析其差异,并通过机器学习构建两组的分类模型。在建立的6个分类模型中,采用随机森林算法建立的肝硬化预测模型有着最大的曲线下面积(AUC)值和最高的灵敏度。本研究能够简单、快速、稳定、可靠地检测和评估人体肠道菌群的平衡状态,有助于临床工作。

© 2023 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. 引言

人体肠道是一个复杂而精细的微生态系统,能介导宿主与环境之间的相互作用。人类肠道包含数以万亿计的微生物,如细菌、真菌、病毒和其他生命形式,其中大多数存在于结肠中。这些微生物与肠道环境共同构成肠道微生态系统,其多样性是肠道菌群与宿主共同进化的结果[1–

3]。肠道微生态的组成容易受到许多因素的影响,如饮食[4–5]、年龄[6–7]、性别[7]、遗传[8]和药物等[1]。

肠道菌群的组成在调节人体健康和疾病中起着至关重要的作用[9]。肠道微生物群除了参与人体的消化功能外,还能影响人体的发育、生长和生理,包括器官发育、形态发生和代谢[10–11]。肠道菌群在人体免疫系统的发育和诱导中也发挥着不可或缺的作用,能通过调节免疫细胞的

\* Corresponding authors.

E-mail addresses: wuzhongwen@zju.edu.cn (Z. Wu), lji@zju.edu.cn (L. Li), suon@zju.edu.cn (B. Li).

# These authors contributed equally to this work.

分化和免疫介质的产生来维持宿主与肠道菌群之间的相互作用[12–14]。正常肠道菌群的破坏会增加感染、有害病原体过度增殖以及炎症性疾病发生的风险[12]。Qin等[15]对肝硬化（LC）患者肠道菌群的改变进行了研究，发现在属水平上，拟杆菌属在两组中均为优势菌，但其丰度在LC组中显著降低。因此，对健康宿主来说，维持微生态平衡非常重要。事实上，肠道微生态失衡可能会导致多种疾病的发生，包括肝脏疾病、胃肠道疾病、代谢性疾病、心血管疾病等[16–19]。

因此，定期及时获得人体肠道微生物群落的状态对于辅助评估整体健康状况和预测疾病非常重要。过去，基于传统方法如细菌培养或显微检查，常用球菌/杆菌（C/B）比例来反映肠道细菌平衡[20–21]。近年来，随着聚合酶链反应（PCR）技术以及由此发展起来的微生物测序技术的成熟，可以检测到更多的细菌种类，并选择一些关键细菌种类的数量、细菌比例等指标作为肠道微生态平衡指标。例如，研究发现双歧杆菌丰度/肠杆菌丰度（B/E）比值可以判断肝脏疾病进展过程中肠道微生态失调的程度[22–23]。Low等[24]发现早期婴儿中克雷伯菌属/双歧杆菌属丰度的比值（K/B）增加是过敏性疾病风险增加的潜在指标[24]。此外，Ley等[25]首次提出肠道中较高的厚壁菌门/拟杆菌门丰度（F/B）比值可能导致肥胖。微生物测序技术因其高通量的特点，现已成为研究肠道菌群的主流方法。目前关于肠道微生物群的研究多基于16S rRNA基因测序[26–28]和宏基因组测序[29–30]。然而，这两种方法既昂贵又低效，而且步骤通常非常烦琐。因此，微生物测序技术不适合在大规模人群中进行肠道菌群的常规检测。定量PCR（quantitative PCR, qPCR）具有快速、简便、定量的功能[31–32]，是进行常规肠道细菌检测的一种很好的替代方法。

在前期研究的基础上[15,32]，本研究试图建立一种高效且常规的肠道细菌检测流程，通过qPCR检测到人类肠道中的10种优势菌，包括益生菌（乳杆菌和双歧杆菌）、条件致病菌（肠杆菌、肠球菌、拟杆菌和奇异菌）和其他促进健康的共生菌（普拉梭菌、丁酸梭菌、柔嫩梭菌和直肠真杆菌）。这10种代表性的细菌是从一个中国大样本健康人群中筛选出来的，被认为是评估整个人体肠道菌群平衡状况的潜在指标。此外，我们还尝试获得健康人群肠道中10种菌的正常参考值范围及其两菌配对比值随年龄增长的变化模式，这有助于后续大样本特定疾病人群的研究。为了研究qPCR检测结果在健康人群和特定疾病人群之间可能存在的差异，本研究评估了肝硬化患者和健康对照组之间10种肠道优势菌的差异。为了进一步验证qPCR

检测结果区分疾病人群和普通人群的能力，我们利用机器学习算法深入挖掘肝硬化人群和健康对照人群的检测信息，并构建多个分类模型，最终通过比较筛选出一个最优的分类模型。

## 2. 材料和方法

### 2.1. 受试者招募和样本采集

共纳入510名健康受试者和248例肝硬化患者。纳入和排除标准见附录A中的S1部分。将清晨第一次排的新鲜粪便收集在一次性塑料碗中，然后塞进一次性干净塑料袋中，将塑料袋紧紧固定，整个过程中注意避免尿液污染。为避免食物残渣等干扰因素，选择新鲜粪便中较软的部分，在半小时内装入加有DNA稳定剂（Invitex, 德国）的冻存管，对每个冻存管进行编号后放入-80℃冰箱保存。采集样本的当天上午采集患者血液标本，进行血常规、血生化、C-反应蛋白（CRP）等指标的检测。后续登记所有志愿者的基本信息（表1），包括年龄、性别、BMI等。最终，排除粪便标本不合格、基本信息不完整的受试者后，共纳入500名健康体检者和244例肝硬化患者。所有研究工作均按照浙江大学医学院附属第一医院研究伦理委员会批准的指南（编号:2019-1026和2022-874）进行。

### 2.2. qPCR评估肠道中主要细菌

使用核酸提取纯化试剂（博日公司，浙江杭州）提取744个志愿者粪便中的微生物DNA，具体步骤详见附录A中的S2部分。之后使用Nanodrop one（Thermo Fisher Scientific, USA）检测样本DNA洗脱液中的粪便微生物总DNA浓度。使用实时荧光定量PCR法检测肠道内10种优势菌群，所使用的引物见附录A中的表S1，均由金斯瑞公司（江苏南京）合成。采用VIA TM PCR系统（Applied Biosystems, Inc., Carlsbad, CA, USA）对744份粪便菌群DNA样本进行qPCR，反应体积为20 μL，包括10 μL SYBR Green PCR Master Mix（中诺基因公司，浙江杭州）、8 μL引物对（0.2~0.6 μmol·L<sup>-1</sup>）、2 μL模板DNA或2 μL蒸馏水（阴性对照）。反应条件列于附录A中的表S2中，每个反应重复3次，重复之间的循环阈值（ΔCT）要求小于0.5。将含有各引物组相应扩增片段的质粒DNA标准品按多个梯度比例稀释后，质粒DNA标准品和阴性对照品在同一PCR板上用细菌DNA模板进行扩增。通过与相应细菌DNA标准品扩增得到的标准曲线比较，确定DNA模板中目标细菌的拷贝数。用DNA模板中目标细菌的浓度除以每个样本DNA洗脱液中粪便微生物

表1 744名受试者特征

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 <sup>-2</sup> )	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 (×10 <sup>12</sup> L <sup>-1</sup> )	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 <sup>-1</sup> )	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 (×10 <sup>9</sup> L <sup>-1</sup> )	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 (×10 <sup>9</sup> L <sup>-1</sup> )	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 (×10 <sup>9</sup> L <sup>-1</sup> )	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 (×10 <sup>9</sup> L <sup>-1</sup> )	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 (×10 <sup>9</sup> L <sup>-1</sup> )	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 <sup>-1</sup> )	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 <sup>-1</sup> )	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 <sup>-1</sup> )	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 <sup>-1</sup> )	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 <sup>-1</sup> )	70.65 ± 6.78	73.11 ± 4.17	65.65 ± 8.17	65.00–85.00	< 0.0001
ALB (g·L <sup>-1</sup> )	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 (μmol·L <sup>-1</sup> )	12.80 (8.80–18.20)	10.55 (7.90–14.30)	23.05 (13.73–40.20)	0–21.00	< 0.0001
DBIL (μmol·L <sup>-1</sup> )	4.30 (2.90–6.80)	3.50 (2.50–4.40)	10.15 (5.83–20.58)	0–8.00	< 0.0001
IBIL (μmol·L <sup>-1</sup> )	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 (μmol·L <sup>-1</sup> )	5.90 (3.70–13.68)	4.20 (3.60–6.80)	29.60 (9.83–72.60)	0–10.00	< 0.0001
Cr (μmol·L <sup>-1</sup> )	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 (μmol·L <sup>-1</sup> )	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 <sup>-1</sup> )	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 <sup>-1</sup> )	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 <sup>-1</sup> )	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 <sup>-1</sup> )	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 <sup>+</sup> (mmol·L <sup>-1</sup> )	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 <sup>+</sup> (mmol·L <sup>-1</sup> )	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 <sup>-</sup> (mmol·L <sup>-1</sup> )	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 <sup>2+</sup> (mmol·L <sup>-1</sup> )	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 <sup>-1</sup> )	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).

总DNA浓度和模板体积，得到目标细菌的最终浓度。浓度的单位为拷贝数/纳克 (copies/ng) 粪便微生物总DNA，以下简称copies/ng。计算公式详见附录A中的S4部分。

### 2.3. 数据分析

使用R脚本 (4.1.2版)、Rstudio软件 (2022.07.1+554版本) 和Origin 2021软件 (9.95版本) 进行统计学分析和绘图。Ggplot2包 (3.3.5版本) 用于图形绘制。相关性分析的热图由Corplot包 (0.92版本) 绘制。10种肠道特异菌群丰度比较的热图由Pheatmap包 (1.0.12版本) 绘制而

成。数据预处理由Dplyr包 (1.0.8版本) 完成。数据的所有离群值处理采用箱线图法，小于Q1-1.5\*四分位距 (Interquartile range, IQR) 或大于Q3+1.5\* IQR的离群值将会从相应的数据组中删除。参考范围 (双侧) 由健康人群数据剔除离群值后取95%置信区间计算得到。正态分布数据采用正态分布法，上、下限为 $\bar{X} \pm 1.96 \times SD$ 。非正态分布数据采用百分位数法，上、下限分别为P<sub>2.5</sub>和P<sub>97.5</sub>。采用Shapiro-Wilk检验和Kolmogorov-Smirnov检验来检测10种肠道特异菌群的正常分布。两组间差异分析采用Wilcoxon秩和检验。多组间的非正态分布采用Kruskal-

Wallis 检验分析。肝硬化预后相关指标与 10 种肠道菌群的相关性采用 Spearman 秩相关分析。

## 2.4. 机器学习

我们利用机器学习算法建立了用于区分肝硬化和非肝硬化样本的模型。来自同一样本采集区域的 744 例临床样本（肝硬化样本 244 例，健康对照样本 500 例）作为构建模型的数据集。采用随机抽样法，按照 75:25 的比例将数据分为训练集和测试集。其中有 408 个训练数据（183 个肝硬化数据点和 225 个非肝硬化数据点）和 136 个测试数据（61 个肝硬化数据点和 75 个非肝硬化数据点）。

通过预试验筛选，最终使用随机森林（random forest, RF）[33]、梯度提升（gradient boosting machine, GBM）[34]、自适应提升算法（adaptive boosting, AdaBoost）[35]、极限梯度提升算法（eXtreme gradient boosting, XGBoost）[36]、支持向量机（supported vector machine, SVM）\_poly 和 SVM\_Gauss [37] 6 种机器学习方法对数据进行分类和建模，前 4 种方法属于集成学习[38]。

我们选择临床样本中 10 种肠道特异菌群的含量和样本信息中的性别作为训练模型的特征[7]。使用十折交叉验证的随机森林模型来探究这些特征的重要性，使用基尼指数的平均下降幅度作为衡量指标，来确定这些特征的重要性以及在模型中的贡献度。

采用重复十折交叉验证（10 次）构建和验证模型，并采用超参数调优对 6 个模型进行调优。通过超参数调优，得到 6 个模型各超参数值的最优集合，并根据这些值的集合建立各模型的最优模型，SVM\_poly 和其他模型的超参数优化见附录 A。最后，使用曲线下面积（area under the curve, AUC）评估训练模型的质量，选择平均 AUC 值最高的模型为最优模型。当存在过拟合情况时，选择次优模型作为最优模型。各模型相应的 AUC 值、敏感性和特异性由测试数据通过 6 种机器学习算法生成的最终模型验证得出。

机器学习分析过程是在 R 语言 caret 机器学习框架（6.0-93 版本）下实现的。使用 RF（4.7-1.1 版本）进行随机森林分析；使用 GBM 包（2.1.8.1 版本）进行梯度升级；AdaBoost 是使用 Adabag 包（4.2 版本）实现的；XGBoost 是使用 XGBoost 包（1.6.0.1 版本）实现的；并使用 Kernlab 包（0.9-31 版本）实现了两种支持向量机方法。

## 2.5. 伦理

本研究获得浙江大学医学院附属第一医院研究伦理委员会批准（编号：2019-1026 和 2022-874）。患者/参与者

提供了参与本研究的书面知情同意。本研究遵循 1975 年《赫尔辛基宣言》伦理指导原则。

## 3. 结果

### 3.1. 健康人群肠道中主要菌群

本研究共纳入 500 名健康体检者，其中，男性 275 名，女性 225 名。根据严格的排除标准招募健康受试者。符合条件的健康受试者被招募入组，通过 qPCR 检测健康受试者 10 种肠道特异菌群，并对所有数据进行统计处理，建立健康人群中 10 种肠道特异菌群丰度的参考值范围（表 2）。此外，鉴于不同年龄人群的肠道菌群结构和丰度发生改变，我们也分析并获得了 10 种菌在不同年龄分组中的参考值范围（附录 A 中的表 S3）。将不同年龄段的健康人以 20 岁为间隔分为 5 组：0~20 岁、20~40 岁、40~60 岁、60~80 岁、80~100 岁。除直链真杆菌的数量较为均衡，在各年龄分组之间没有显著差异（ $P=0.081$ ）外，其余 9 种菌在 5 个年龄组间均有显著差异。其中，奇异菌的丰度变化最显著（ $P=5.7\times 10^{-8}$ ），其次为肠球菌（ $P=1.9\times 10^{-7}$ ）。

表 2 健康人群中 10 种肠道特异菌的参考值范围

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

### 3.2. 两菌比值是肠道微生态平衡的潜在指标

研究表明，在几乎所有疾病中，厚壁菌门的丰度降低，而拟杆菌门的丰度增加[39]。因此，我们探索了 10 个菌中的两菌比值是否可以作为评估肠道微生态平衡状况的潜在指标。我们将上述 10 种肠道特异菌群丰度对数值进行比较[图 1 (a)]。将任意两菌对比后共得到 45 个两菌比值，并绘制两菌比值随年龄增长的趋势图。有趣的是，我们发现 B/E 值随着年龄的增长呈现出先降低后升高的典型趋势，呈 U 形曲线（ $P=0.021$ ）。然而，仍需要更多的研究来验证。此外，Ec/E 这一指标在危重患者中显著升高

[32], 本研究中发现 Ec/E 随年龄增长呈递增趋势 ( $P=0.00025$ )。组间差异最大的两种细菌比值为柔嫩梭菌/拟杆菌 ( $P=2.3\times 10^{-7}$ ), 其随年龄增长的变化趋势也是呈明显的先下降后上升的 U 形曲线[图 1 (b)、(c)]。

### 3.3. 肝硬化患者肠道菌群存在失衡

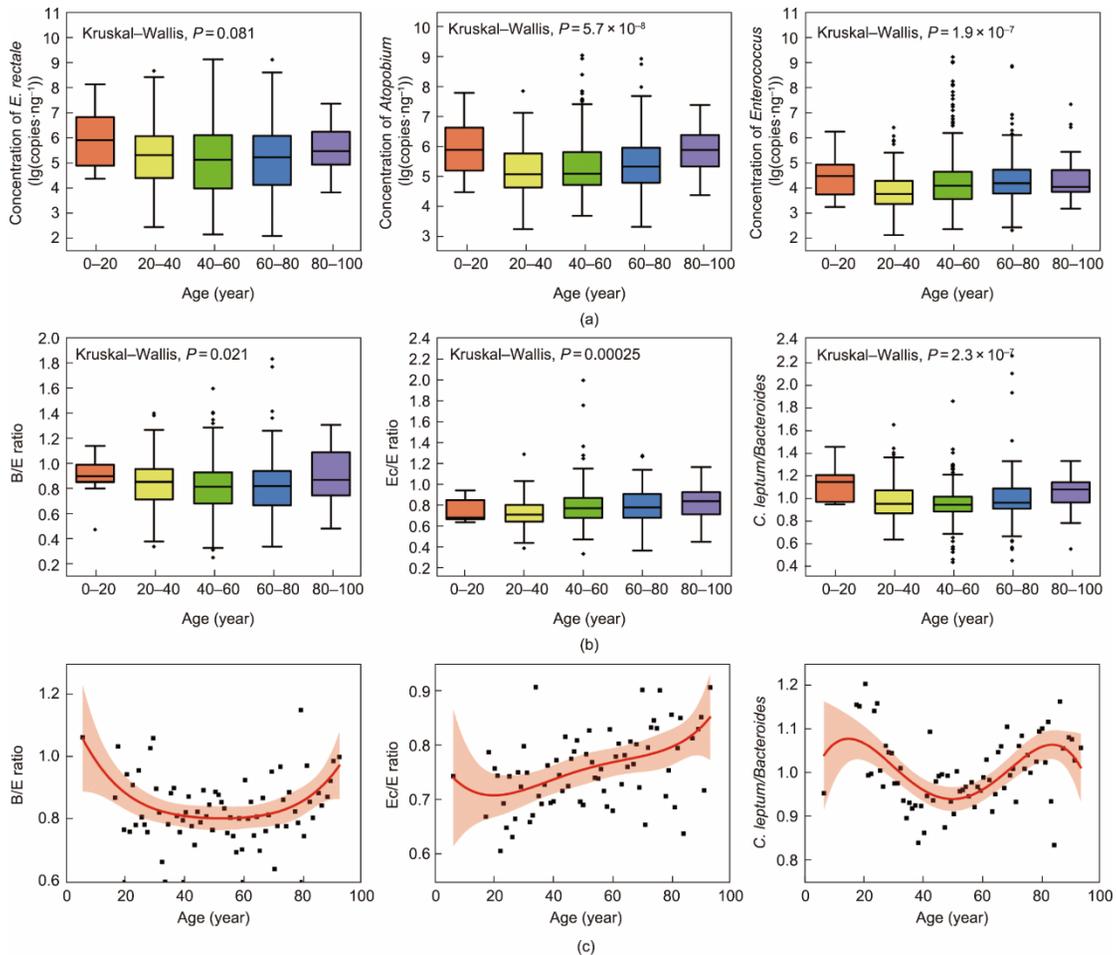
为了进一步验证上述 10 种肠道特异菌群是否可以作为评估肠道微生态平衡状态的指标, 我们收集并采用 qPCR 方法测定了 244 例肝硬化患者的粪便样本。通过比较健康人群和肝硬化患者肠道菌群的 qPCR 检测结果发现, 有 7 种菌群在两组间存在显著差异[图 2 (a)]; 而普拉梭菌、柔嫩梭菌和奇异菌在两组间无显著差异。在 7 种菌群中, 肠球菌、直肠真杆菌和拟杆菌差异最大 ( $P=2.50\times 10^{-12}$ 、 $3.47\times 10^{-10}$ 、 $6.57\times 10^{-10}$ ), 其次是丁酸梭菌 ( $P=1.76\times 10^{-5}$ )、乳杆菌 ( $P=1.13\times 10^{-3}$ )、肠杆菌 ( $P=5.59\times 10^{-3}$ ) 和

双歧杆菌 ( $P=9.71\times 10^{-3}$ ), 这与热图中显示的结果一致[图 2 (b)]。此外, 双歧杆菌、肠杆菌和丁酸梭菌的丰度在健康人和肝硬化患者中相对稳定, 离群值很少。肝硬化患者和健康人群 4 种细菌的离群值分别为 (1/244, 1/500)、(0/244, 2/500)、(4/244, 1/500) 和 (2/244, 3/500)。

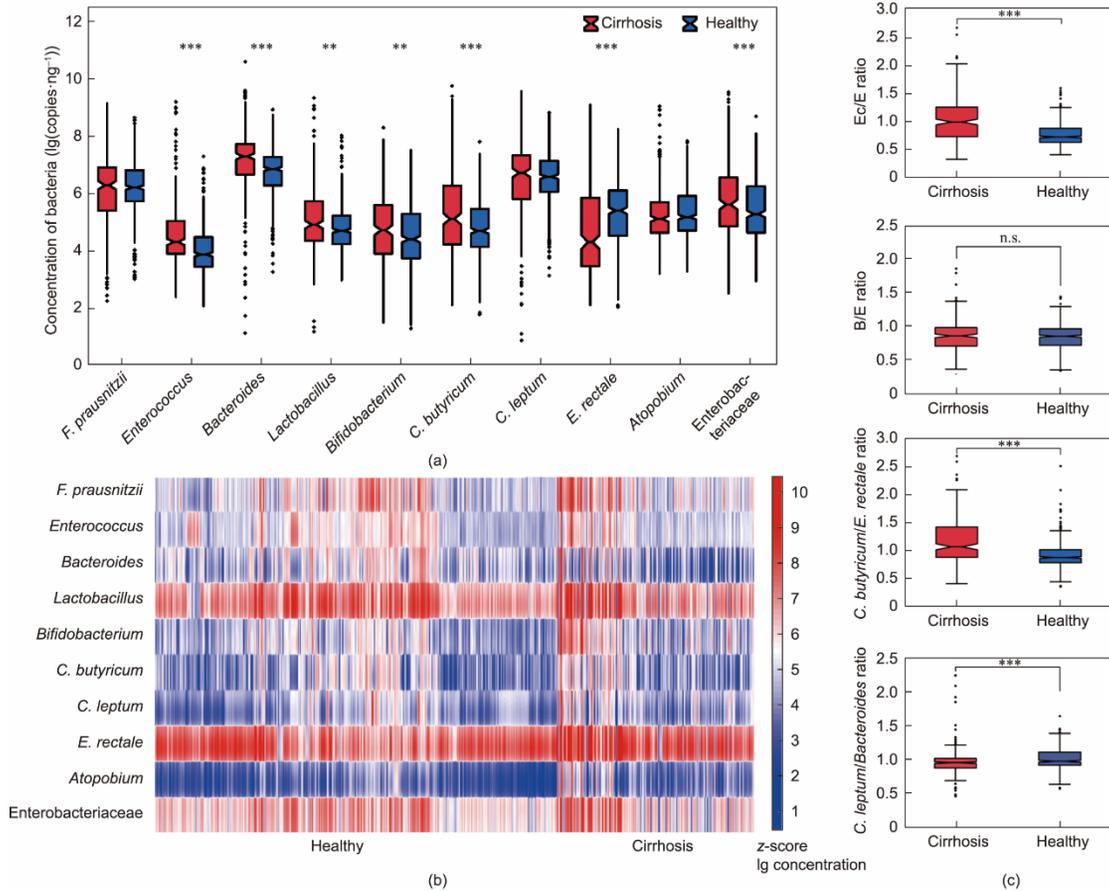
此外, 与健康人群相比, 肝硬化患者肠道中 Ec/E 比值 ( $P=3.08\times 10^{-24}$ ) 和柔嫩梭菌/拟杆菌比值 ( $P=2.14\times 10^{-18}$ ) 存在显著差异, 这也表明肝硬化患者存在肠道菌群失衡[图 2 (c)]。这些结果与我们之前的研究结果一致[15]。然而, 在这项研究中, 两组之间 B/E 值几乎保持不变。

### 3.4. 肠道菌群与肝硬化严重程度相关

肠道菌群与肝功能指标的相关性分析显示, 血清 ALB、DBIL、TG、TBA 水平与拟杆菌的丰度呈正相关,



**图 1.** 健康人群中的肠道菌群。(a) 不同年龄组中 10 种肠道菌群的差异。从左到右依次显示了直肠真杆菌、奇异菌和肠球菌在不同年龄组之间的差异。(b) 不同年龄组中 10 种肠道菌群的成对比值差异, 纵坐标表示 10 种肠道菌群的成对比值。从左到右依次为 B/E 值差异、肠球菌/肠杆菌 (Ec/E) 比值差异和柔嫩梭菌/拟杆菌比值差异与年龄的关系。(c) 10 种肠道菌群的成对比值与年龄的关系。纵坐标表示 10 种肠道菌群的成对比值。黑点表示每个年龄段两种肠道菌群比值的平均值。红色曲线是通过多项式拟合实现的, 浅红色区域是拟合曲线的 95% 置信区间。箱线图用四分位数表示, 箱体上、下限表示从第一到第三四分位数的范围, 箱内的线表示中位数。箱体外顶端和底端的横线表示第一和第三四分位数之外 1.5 倍四分位距的最小和最大值。黑点表示离群值。



**图2.** 肝硬化患者与正常对照组的肠道菌群。(a) 箱型图表示的肝硬化组和正常对照组中10种优势肠道菌种的丰度。纵轴表示10种菌丰度的对数值。箱体的上、下限表示第一四分位和第三四分位外1.5倍IQR内的最小和最大值。黑点表示在箱体之外的异常点。切口表示中位数的95%置信区间。(b) 肝硬化组与正常对照组间10种优势菌种含量差异的热力图。水平轴表示肝硬化和正常对照组中的各样本。肝硬化组有244个样本，正常对照组有500个样本。垂直轴表示样本中的10种优势细菌。不同的颜色表示每个样本中10种细菌的丰度的对数值。暗红色表示含量较高，暗蓝色表示含量较低。(c) 肝硬化组与正常对照组之间肠道细菌的两两比值。\*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ 。n.s.: 无显著差异。

与其他9种菌群的丰度呈负相关。血清ALP水平与拟杆菌的丰度呈负相关，与其他9种菌群的丰度呈正相关。肠杆菌的丰度与B/E值呈负相关（图3）。

### 3.5. 多种机器学习模型来区分和预测健康人群和肝硬化患者

此外，我们构建了多种机器学习模型来分析上述肠道菌群的检测结果，并试图用机器学习模型区分健康对照者和肝硬化患者[图4(a)]。对6种模型进行超参数调优，得到超参数的最优组合。基于多项式核的SVM\_poly模型的超参数与模型性能的关系如图4(b)所示。由图可知，超参数的最优组合为多项式阶数=2，尺度=0.100， $C=0.75$ 。其他5种模型的性能估计值与调优参数的关系见附录A中的图S1至图S3。

6种模型的训练结果显示，XGBoost模型的训练效果最好，平均AUC值为0.9376（95% CI, 0.9158~0.9595），其次为SVM\_Gauss和SVM\_Poly，平均AUC值分别为0.9050（95% CI, 0.8754~0.9346）和0.9040（95% CI,

0.87449~0.9331）。RF模型的训练结果最差，但平均AUC也达到了0.8746（95% CI, 0.8414~0.9078）[图4(c)]。

然后使用机器学习模型对50例真实临床样本进行分析。检验结果（表3）显示，6种模型均取得了较好的预测结果，其中，RF的AUC值最好，达到了0.8776（95% CI, 0.8159~0.9393），其次是XGBoost和Adaboost，AUC值分别达到了0.8726（95% CI, 0.8097~0.9355）和0.8630（95% CI, 0.7969~0.9290）。敏感度最高的3个模型是RF、XGBoost和Adaboost，分别为0.8361、0.7869和0.7869。特异度最高的3个模型是XGBoost、AdaBoost和GBM，分别为0.8667、0.8533和0.8400[图4(c)和附录A中的表S4]。

## 4. 讨论

肠道菌群的组成是动态的，肠道菌群失衡是指肠道内正常菌群数量、比例和种类发生改变，影响人体健康，导

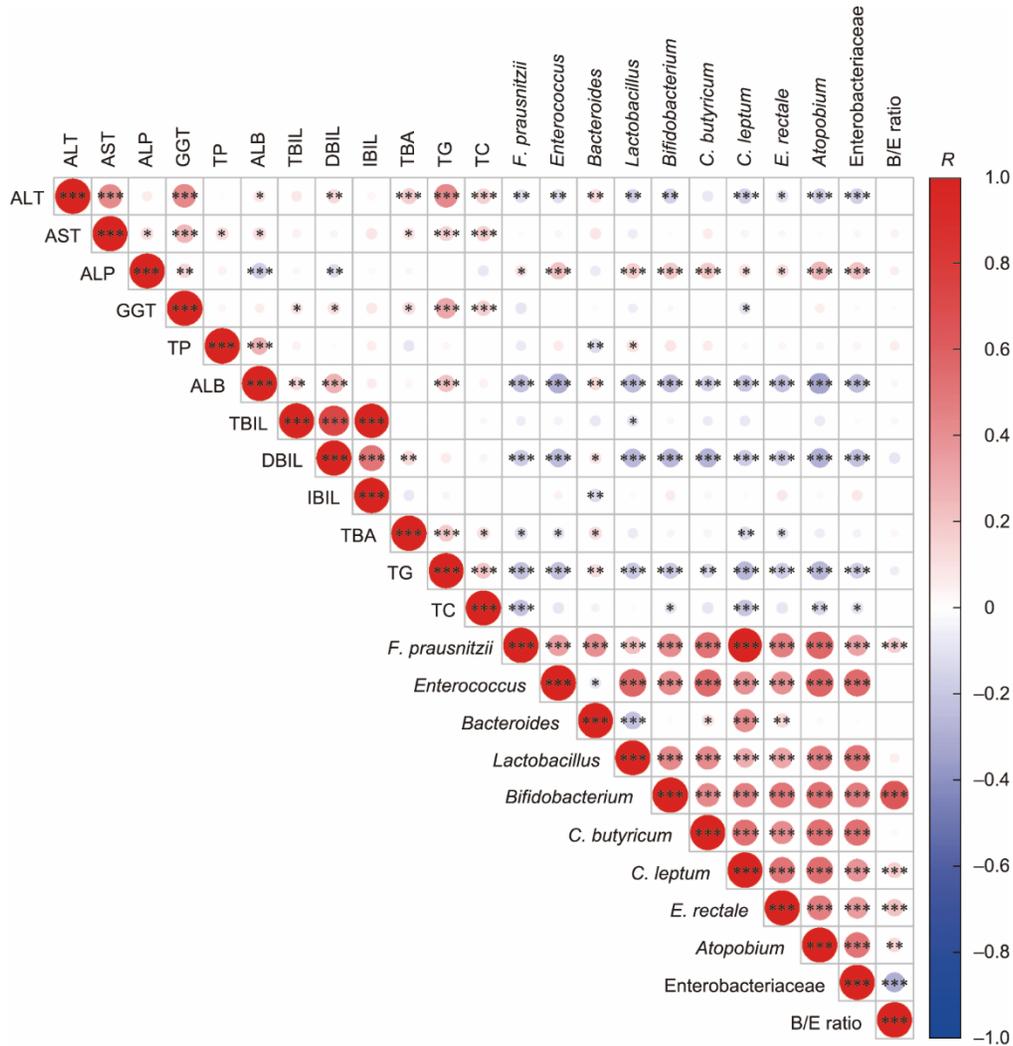


图3. 肝功能指标与肠道菌群之间的相关性。点的颜色深浅和大小显示了相关性的强度 ( $R$  值)。红色越深, 正相关性越强。蓝色越深, 负相关性越强。白色表示无相关性。点越大, 相关性越强。\*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$ 。

致一系列异常的生理和病理现象[40–41]。当肠道微生物群失衡时, 肠道微生物群和宿主免疫系统之间的平衡会受到影响[42]。肠道菌群在人体增龄过程中发生了显著变化[43]。Zhang 等[7]发现随着人类增龄, 不同队列人群的肠道菌群发生一致变化, 均表现为 $\alpha$ 多样性增加。口腔菌群中多个成员, 如产短链脂肪酸的肠杆菌和梭状芽孢杆菌丰度随年龄增长而增加。他们还发现, 数种双歧杆菌(短双歧杆菌、两歧双歧杆菌、长双歧杆菌和青春双歧杆菌)与年龄呈负相关。Biagi 等[44]研究发现核心共生菌群(主要属于瘤胃球菌科、毛螺菌科和拟杆菌科)的累积丰度随着年龄的增长而下降。然而, 与健康相关的菌群阿克曼菌属、双歧杆菌属和克里斯滕森菌科在超高龄人群中富集, 尤其是在 105~109 岁的百岁老人中。另一项研究还发现, 健康老年人的肠道菌群核心属丰度下降, 尤其是拟杆菌, 这与较长的预期寿命相关[45]。虽然不同的研究均发现了

人类增龄过程中肠道菌群的变化, 但关于增龄过程中肠道菌群的变化模式尚未达成共识。因此, 在本研究中, 我们的目的是基于大型队列的研究结果, 建立适用于所有年龄以及不同年龄段健康人群中 10 种肠道特异菌群的参考值范围, 并确定健康老龄化的潜在特征。

在我们前期工作的基础上, 我们通过 qPCR 方法定量检测健康人群中 10 种肠道特异菌群。本研究选择的 10 种肠道特异菌群在维持肠道微生态平衡中发挥重要作用。其中, 乳杆菌和双歧杆菌对调节免疫和维持肠道屏障功能及肠道微生物群多样性和稳定性具有重要作用[46–48], 而普拉梭菌、丁酸梭菌、柔嫩梭菌和直肠真杆菌均可产生短链脂肪酸, 短链脂肪酸在抑制条件致病菌过度生长、维持肠上皮屏障完整性、增强免疫力等方面具有重要作用[49–51]。我们的结果表明, 10 种肠道特异菌群的丰度随年龄的增长发生变化。以 20 年为间隔将健康人群按年龄分为

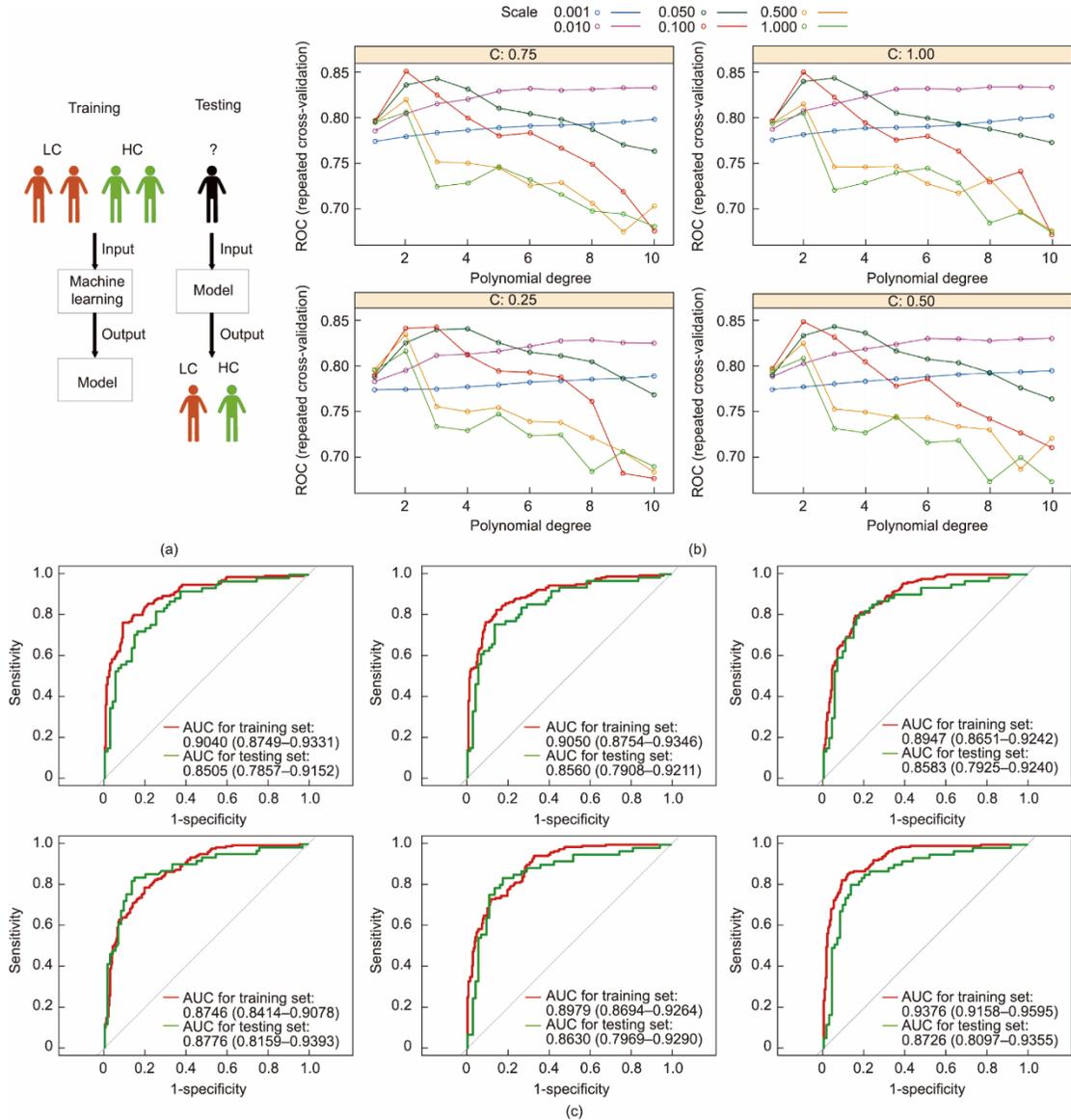


图4. 多种机器学习模型。(a) 机器学习分析过程分为两个步骤：训练和测试。训练通过机器学习方法学习输入的分类的肝硬化和健康对照数据，最终得到优化的分类模型。测试使用训练好的模型对未分类的数据进行分类，并通过相关指标评估模型的质量。(b) 基于多项式核的支持向量机分类模型的超参数调整，通过调整SVM中的C值和scale和多项式阶数，找到平均AUC值最高的超参数组合。最终组合为多项式阶数=2，scale=0.100，C=0.75。(c) 六种机器学习算法对训练集和测试集的AUC值。左上方是SVM\_poly算法的AUC值；左中方是SVM\_Gauss算法的AUC值；右上方是GBM算法的AUC值；左下方是RF算法的AUC值；下中方是AdaBoost算法的AUC值；右下方是XGBoost算法的AUC值。ROC：受试者工作特征曲线。

表3 6种机器学习算法的测试结果

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

5组（0~20、20~40、40~60、60~80、80~100岁），10种肠道特异菌群中有9种在5个年龄组中均有不同程度的差异。奇异菌的差异最显著（ $P=5.7 \times 10^{-8}$ ），其次为肠球菌

（ $P=1.9 \times 10^{-7}$ ）。此外，除拟杆菌随年龄增长呈先升高后降低外，其余菌属随年龄增长均呈先降低后升高的总体趋势。先前大多数大样本研究都是基于16S rRNA或宏基因组测序，但测序的主要局限性之一是，分类群只能根据细菌基因组中单个区域的序列来划分，而且只能得到相对丰度的结果，这就很难产生一个稳定可靠的参考值[31,52]。相比之下，qPCR可以通过校准已知丰度的标准物质来定量检测细菌丰度，具有重复性好、结果快速和操作简单的特点。及时检测肠道微生物失调对临床医生来说有助于实现早期诊断和治疗，并改善预后[32,53]。

我们前期通过宏基因组学研究分析了肝硬化患者肠道

菌群的变化,在属水平上,拟杆菌属在两组肠道中均为主要菌群,但拟杆菌属在肝硬化组中明显减少,在其余菌属中,韦荣球菌属、链球菌属、梭菌属和普雷沃菌属在肝硬化组中富集,而健康对照组中以真杆菌属和 *Alistipes* 为主。在肝硬化组中丰度下降最多的物种中,12种是拟杆菌门,7种是厚壁菌门,特别是 *Clostridiales* [15]。然而,宏基因组测序成本高,数据量大,需要大量的资源来进行分析。因此,在本研究中,我们旨在使用较少数量的肠道菌群来表征肝硬化患者肠道微生态的变化。我们构建了多种机器学习模型来进一步分析健康人群和肝硬化患者的肠道菌群检测结果。有趣的是,我们发现在集成学习框架下4种算法构建的分类模型的预测结果都优于SVM和其他一些机器学习算法。其中,RF分类模型的AUC值和灵敏度最高,XGBoost分类模型的AUC值次之,特异度最高,相对优于其他4种模型。比较这两种分类模型,RF模型具有更高的AUC值和灵敏度,能更好地识别肝硬化或疑似肝硬化患者,但假阳性更多。XGBoost模型特异度较高,识别健康人群的能力较好,但假阴性较多,可能导致未被识别的肝硬化患者未能及时就诊,延误诊断和治疗。因此,灵敏度最高的RF模型为最佳选择。虽然RF模型可能会导致一些健康人被误诊,但能及时发现更多真正的肝硬化患者,利大于弊。然而,地理位置对人体肠道菌群有较大影响,且地理差异限制了健康人群和疾病模型中肠道特异菌群的参考值范围的应用。为了验证肠道微生物群在健康或疾病状态之间的差异,需要标准化的实验方案、区域研究设计和广泛的抽样,并且需要将地理作为特征添加到模型中。针对不同的地理区域需要训练不同的预测模型[54–55]。考虑到地域对人体肠道微生物群的影响,未来将从不同地区采集样本,重新建立分类模型。我们将采取两种策略:①在模型中加入地域作为附加特征;②根据不同区域分别独立地建立分类模型,并根据结果选择一个更好的分类模型。

荧光定量PCR与机器学习的结合使检测和分析人体肠道菌群变得简单、快速、稳定、可靠。随着新样本的纳入,我们的健康人群队列也将随之扩大,健康人群中肠道特异菌群检测结果的参考值范围将随着样本量的扩大而更新,相关机器学习模型预测疾病的能力将通过不断的学习和预测而进一步提高,对未来相关临床工作的开展产生积极的辅助作用。

## 5. 结论

我们对中国人群大样本队列的粪便研究发现,10种

肠道特异细菌能够表征整个肠道菌群的平衡状态。通过qPCR方法定量测定健康人群中10种肠道特异细菌的丰度,建立了10种肠道特异细菌的正常参考值范围,并发现了这10种细菌随年龄增长和疾病状态的变化特点。此外,我们利用机器学习算法从检测结果中深度提取差异信息,构建了可靠的肝硬化预测模型。本研究表明,用少量的肠道微生物信息可描述和预测中国健康人群肠道微生态的平衡状态。通过建立健康人群10种肠道特异细菌的参考值范围,可用于描述和预测肠道微生态失衡。在今后的工作中,仍需要进一步拓展并完善健康人群中10种肠道特异细菌的参考值范围及理论模型。

## 致谢

本研究得到以下项目资助:国家重点研发计划项目(2018YFC2000500)、中央高校科研基金项目(2022ZJH003)、传染病诊疗国家重点实验室独立课题(2022zz22)、国家自然科学基金项目(81703430、32170058、82200994)、中国医学科学院医学科学创新基金项目(2019-I2M-5-045)、济南微生物医学山东实验室科研项目(JNL-2022051B)。

特别感谢KW Liu的支持,感谢中诺基因公司提供的细菌检测和模型建立。

## 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. Longterm 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. Zhenganxifeng 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 JL. 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;45:53–61.
- [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, Elbedwi M, Nambiar RB, Zhou X, et al. *Lactocaseibacillus 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.