scientific reports

Check for updates

OPEN

Microbiome analysis of gut microbiota in patients with colorectal polyps and healthy individuals

Dayi Deng[1,4], Lin Zhao[1,4], Hui Song[1], Houming Wang[1], Hengjie Cao[2], Huimin Cui[3], Yong Zhou[1]✉ & Rong Cui[1]✉

Colorectal polyps serve as the primary precursors for colorectal cancer. A close relationship has been observed between colorectal polyps and gut microbiota. However, the composition and role of the microbiome associated with tubular adenoma are not well understood. In this study, we prospectively evaluated alterations in gut microbiota among patients with colorectal polyps. A total of 60 subjects were enrolled in this study, including 30 patients with colorectal polyps (CP group) and 30 healthy controls (control group). The 16S rRNA sequencing was employed to characterize the gut microbiome in fecal samples. The results revealed that the beta diversity of the gut microbiota in the CP group significantly differs from that of the control group ($p = 0.001$). At the phylum level, the relative abundance of *Bacteroides*, *Fusobacteria*, and *Proteobacteria* was higher in the CP group compared to the control group ($p < 0.05$), whereas the relative abundance of *Actinobacteria* was higher in the control group in comparison to the CP group ($p < 0.05$). At the genus level, the abundance of *Bacteroides* increased in the CP group ($p < 0.05$), while *Bifidobacterium* declined in the CP group ($p < 0.05$). At the species level, the abundance of *Clostridium perfringens*, *unidentified_Bacteroides*, *unidentified_Dorea*, *Escherichia coli*, *Clostridium ramosum*, and *Ruminococcus gnavus* was higher ($p < 0.05$), whereas the abundance of *Bifidobacterium adolescentis*, *unclassified_Bifidobacterium*, *Bifidobacterium longum*, *Faecalibacterium prausnitzii*, and *unidentified_Bifidobacterium* is lower in CP group compared to the control group ($p < 0.05$). There was a structural imbalance in the composition of intestinal colonization flora for CP patients, characterized by a decrease in beneficial bacteria and an increase in harmful bacteria. *Escherichia*, *Shigella*, and *Bacteroides* may serve as promising biomarkers for early detection of colorectal polyps.

Keywords Colorectal polyps, Microbiome analysis, Gut microbiota

Colorectal cancer (CRC) stands as the third highest contributor to cancer-related mortality globally, resulting in approximately 1.9 million new cases and 935,000 deaths each year[1]. The occurrence and development of CRC involve a multifactorial and multi-step process influenced by a combination of environmental, genetic, and lifestyle factors. Colorectal polyps are protrusions on the surface of the colorectum and serve as the primary precursors to CRC[1]. The transformation of polyps into malignant carcinomas makes them closely associated with colorectal cancer development[2]. Tubular adenomatous polyps and serrated polyps are two prevalent types of precancerous lesions that possess a high potential for malignancy[3]. Therefore, early diagnosis through polyp screening is crucial for preventing the occurrence and progression of colorectal cancer.

Previous studies have indicated an association between gut microbiota and intestinal diseases, including irritable bowel syndrome, inflammatory bowel disease, and CRC[4,5]. The intestinal microbiota has been shown to have a significant impact on the progression from colorectal adenomatous polyps to colorectal cancer[6]. A recent study has reported a significant correlation between an increase in *Fusobacterium_mortiferum* and the development of colorectal polyps[7]. Another study revealed the presence of the *Prevotella* enterotype in patients

[1]Department of Surgery, Jiading Hospital of Traditional Chinese Medicine, 222 Bole Road, Jiading District, Shanghai 201800, China. [2]Department of Surgery, Longhua Hospital Shanghai University of Traditional Chinese Medicine, Shanghai 200030, China. [3]Department of Surgery, Jinan Licheng District Hospital of Chinese Medicine, Jinan 250000, China. [4]Dayi Deng and Lin Zhao contributed to the work equally. ✉email: 13916853065@163.com; 13585551944@163.com

with colorectal adenoma, suggesting it is a distinctive enterotype associated with this condition[8]. Furthermore, gut microbiome differences have been found among healthy individuals and patients with serrated polyps[9,10]. Therefore, the distinctive microbiome alterations associated with premalignant colorectal polyps could function as biomarkers for early cancer detection. Nonetheless, the composition and role of the microbiome associated with tubular adenoma are not well understood.

The characteristics of fecal microbiota in patients with colorectal polyps, primarily tubular adenoma, and healthy individuals were analyzed using 16S rRNA gene sequencing. The objective was to determine the importance of fecal microbial communities in predicting the presence of colorectal polyps, ultimately preventing the occurrence and progression of CRC.

Materials and methods

Subjects and sample collection

A total of 60 subjects, 30 with colorectal polyps and 30 with healthy colorectums, were prospectively included from Shanghai Jiading District Hospital of Traditional Chinese Medicine from March 2024 to August 2024. Written consent forms were obtained from all participants, and the present study was approved by the Ethics Committee of the Hospital. Patients undergoing colonoscopy and diagnosed with colorectal polyps and healthy individuals without colorectal polyps were included. All participants should be older than 18 years old. Exclusion criteria for health control were as follows: (1) with any of the underlying diseases (autoimmune diseases such as ankylosing spondylitis and systemic lupus erythematosus; cachexia; organ failure; infectious diseases; respiratory and cardiovascular diseases); (2) history of surgery and history of chronic drug use; (3) use of antibiotics or probiotics within two months.

Fecal samples were self-collected by participants after enrollment. The samples were immediately collected and placed on dry ice and then stored at − 80 °C before use. Then samples were sent to Bioprofile Co. Ltd (Shanghai, China) for DNA extraction, sequencing, and related analysis.

DNA extraction

DNA of stool samples was extracted using the OMEGA Soil DNA Kit (M5635-02) (Omega Bio-Tek, Norcross, GA, USA), following the manufacturer's instructions, and stored at -20 °C before further analysis. The quantity and quality of extracted DNAs were measured using a NanoDrop NC2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and agarose gel electrophoresis, respectively.

16S rRNA sequencing and analysis

Pfu high-fidelity DNA polymerase was used for Polymerase chain reaction (PCR) amplification. PCR amplification of the bacterial 16S rRNA genes V3—V4 region was performed using the forward primer 338F (5'-ACTCCTACGGGAGGCAGCA-3') and the reverse primer 806R (5'-GGACTACHVGGGTWTCTAAT-3'). PCR products were purified using Vazyme VAHTSTM DNA Clean Beads (Vazyme, Nanjing, China). The 16S rRNA amplicon libraries were prepared by TruSeq Nano DNA LT Library Prep Kit (Illumina, San Diego, CA, USA). The final 16S rRNA amplicon libraries were sequenced on a MiSeq sequencing platform (Illumina, San Diego, CA, USA) using a MiSeq Reagent Kit V3 (600 cycles) for paired-end sequencing.

Microbiome bioinformatics were performed with QIIME2 2019.4 with slight modification according to the official tutorials (https://docs.qiime2.org/2019.4/tutorials/). Briefly, raw sequence data were demultiplexed using the demux plugin following by primers cutting with cutadapt plugin[11]. Sequences were then quality filtered, denoised, merged and chimera removed using the DADA2 plugin[12]. Non-singleton amplicon sequence variants (ASVs) were aligned with mafft[13] and used to construct a phylogeny with fasttree2[14]. Taxonomy was assigned to ASVs using the classify-sklearn naive Bayes taxonomy classifier in feature-classifier plugin[15] against the Greengenes database (Release 13.8). Greengenes database (Release 13.8)[16] was selected for the 16S rRNA gene analysis of bacteria or archaea. We used local nt database for annotation for functional genes or others.

Statistical analysis

ASV-level alpha diversity indices, such as Chao1 richness estimator, Shannon diversity index, Simpson index were calculated using the ASV table in QIIME2, and visualized as box plots. ASV-level ranked abundance curves were generated to compare the richness and evenness of ASVs among samples. To evaluate the diversity for species complexity across samples, beta diversity calculations were analyzed using Principal co-ordinates analysis (PCoA) analysis. The statistical significance of the differences in beta diversity between groups was evaluated by permutational ANOVA (PERMANOVA) and Permdisp using QIIME2. Bacterial abundance and diversity were compared using the Wilcoxon rank-sum test and Welch's t-test. Taxa abundances at the ASV levels were statistically compared among samples or groups by MetagenomeSeq, and visualized as Manhattan plots[17]. Linear discriminant analysis effect size (LEfSe) was performed to detect differentially abundant taxa across groups using the default parameters[18]. Random forest analysis was employed with QIIME2's default settings to differentiate samples from different groups. It involved employing the unrarefied ASV/0TU table and the abundance table at taxonomic levels, including phylum, class, order, family, and genus, for conducting nested cross-validation analysis. Nested stratified tenfold cross validation was used for automated hyperparameter optimization and sample prediction. The area under the receiver operating characteristic (ROC) curve (AUC) was employed to assess the diagnostic value of potential bacterial markers in discriminating between the CP and control group. This evaluation was performed by a diagnostic model that is based on only one species.

Statistical analysis was conducted utilizing R software. The results of continuous variables were presented as mean ± standard deviation (mean ± SD), while categorical variables were presented as frequency and percentage. T-test or Wilcoxon tests were employed to analyze differences in the abundance between two groups, for normally or not normally distributed data, respectively. A p value of less than 0.05 was considered statistically significant.

Results

Clinical characteristics study population

A total of 60 subjects were enrolled in this study: 30 healthy individuals in the control group and 30 individuals in the colorectal polyp (CP) group. The histological classification demonstrated that tubular adenomas constituted 86.7% of patients within the CP group. No significant differences were observed between the CP and control groups in terms of age, gender, height, weight, or BMI. The detailed demographic data of the participants is summarized in Table 1.

Alpha diversity and beta diversity analysis

To assess the differences in microbial communities and microbiota distribution within fecal microbiota between the CP and control groups, sequence alignment was performed to estimate alpha and beta diversity. There were no significant differences in alpha diversity indices including Chao1 index ($p = 0.48$) (Fig. 1a), Simpson index ($p = 0.78$) (Fig. 1b), and Shannon index ($p = 0.42$) (Fig. 1c) between the CP and control groups. This finding showed that microbial communities in feces are similar between the CP and control groups. This is in line with previous research conducted on patients diagnosed with tubular adenomas[19]. Nonetheless, both unweighted and weighted PCoA plots showed a significant difference in the microbiota distribution of the two groups ($p = 0.001$; Fig. 1d, e). Moreover, the Permdisp analysis revealed a homogeneous dispersion of data between the CP and control groups (Table 2).

Altered fecal microbiota in the CP and control groups

Community analysis was performed utilizing a Venn diagram, in which 17,779 Operational Taxonomic Units (OTUs) were unique to the CP group, 12,743 OTUs were unique to the control group, and 998 OTUs overlapped in the CP and control groups (Fig. 2a). Analysis of the relative abundance of each group in the Venn diagram at the phylum level revealed that *Firmicutes* was the most predominant in the CP and control groups (Fig. 2b). In addition, the relative abundance of *Bacteroides, Fusobacteria,* and *Proteobacteria* were higher in the CP group compared to the control group, whereas the relative abundance of *Actinobacteria* was higher in the control group in comparison to CP group. At the genus level, the abundance of *Bacteroides* was increased in the CP group, while *Bifidobacterium* declined in the CP group (Fig. 2c).

Identification of biomarkers for colorectal polyps

To identify dominant fecal microbiota as biomarkers, LEfSe analysis was conducted for all samples. A total of 48 OTUs exhibiting statistically significant differences in abundance were identified. Among these, 34 OTUs showed significantly higher abundance in the CP group compared to the control group, while 14 OTUs exhibited significantly higher abundance in the control group in comparison to the CP group ($p < 0.05$; Fig. 3a, b and Supplementary Table 1).

	Colorectal polyps (N = 30)	Control (N = 30)	*p* value
Age, year			0.954
Mean (SD)	59.6 (6.97)	59.5 (6.49)	
Median [Min, Max]	59.0 [45.0, 73.0]	60.0 [45.0, 69.0]	
Gender			0.785
Male, n (%)	21 (70.0%)	19 (63.3%)	
Female, n (%)	9 (30.0%)	11 (36.7%)	
Height, cm			0.474
Mean (SD)	166 (8.10)	164 (8.06)	
Median [Min, Max]	167 [150, 178]	163 [150, 178]	
Weight, kg			0.949
Mean (SD)	65.4 (8.04)	65.3 (7.99)	
Median [Min, Max]	66.5 [46.0, 80.0]	65.5 [46.0, 80.0]	
BMI, kg/m^2			0.596
Mean (SD)	23.8 (2.59)	24.1 (2.30)	
Median [Min, Max]	23.6 [18.7, 30.1]	24.2 [20.4, 30.1]	
Histology			
Tubular adenoma, n (%)	19 (63.3%)	/	
Tubular adenoma/ serrated polyps, n (%)	1 (3.3%)	/	
Tubular adenoma/ hyperplastic polyp, n (%)	5 (16.7%)	/	
Tubular adenoma/ hyperplastic polyp/ serrated polyps, n (%)	1 (3.3%)	/	
Serrated adenoma, n (%)	1 (3.3%)	/	
Hyperplastic polyp, n (%)	3 (10.0%)	/	

Table 1. Demographic and disease characteristics of included individuals. BMI: body mass index; Max: maximum; Min: minimum; SD: standard deviation.

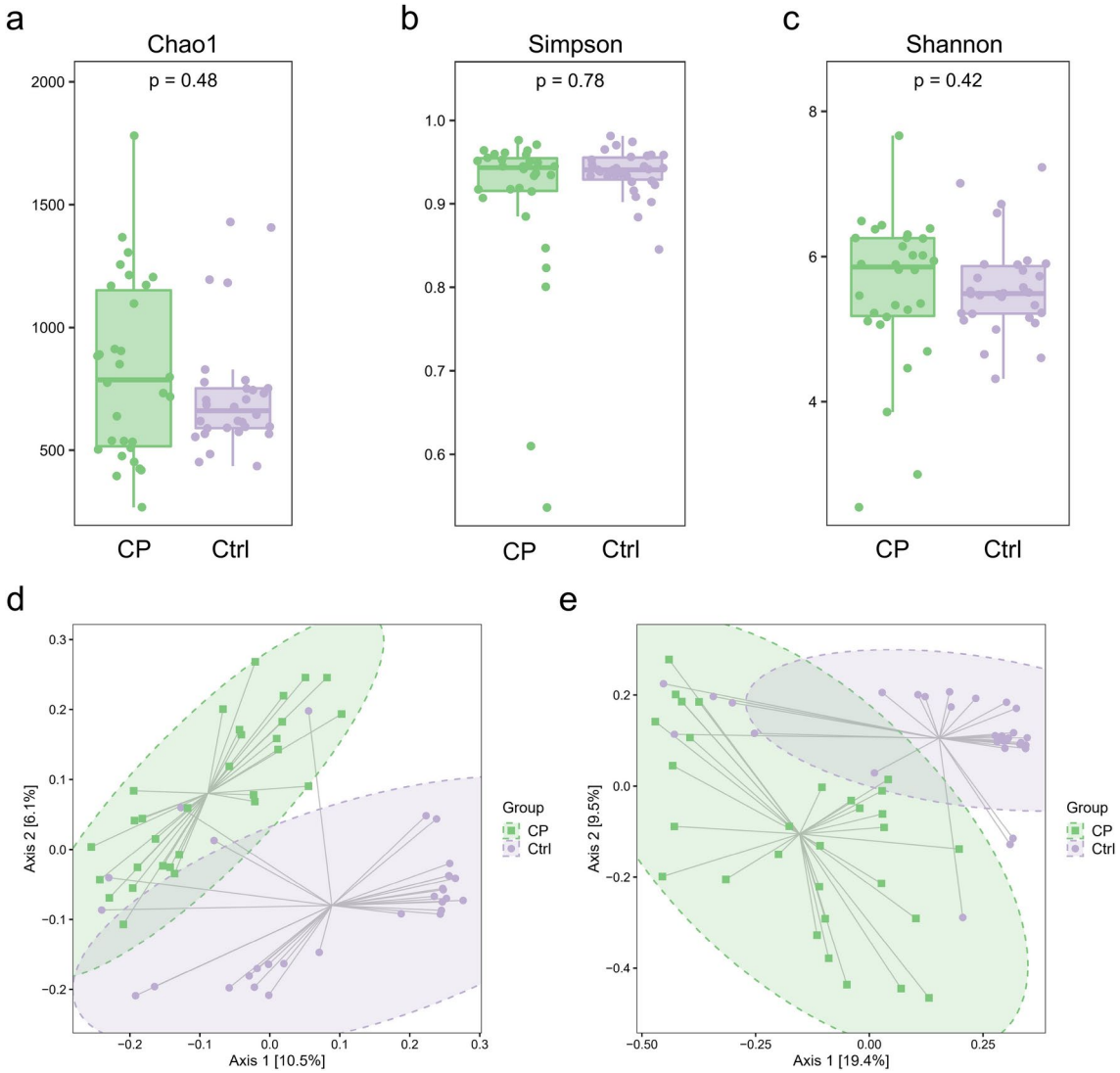

Fig. 1. Microbial diversity analyses between the CP and control groups. (**a**) Chao1 index, (**b**) Simpson index, (**c**) Shannon index, (**d**) unweighted PCoA analysis, (**e**) weighted PCoA analysis. CP: colorectal polyps. PCoA: Principal co-ordinates analysis.

Sample size	Permutations	*p* value
60	999	0.001

Table 2. Permdisp results between the colorectal polyp and control groups.

Furthermore, using LEfSe and metagenomeSeq analysis, we observed that microbiota of the CP group showed a significantly increased abundance of *Bacteroidetes*, *Verrucomicrobia*, and *Proteobacteria* at the phylum level ($p < 0.05$; Fig. 3 and Supplementary Table 1); *Bacteroides*, *Ruminococcus,* and *Akkermansia* at the genus level ($p < 0.05$; Fig. 3 and Supplementary Table 1); *Clostridium perfringens*, *unidentified_Bacteroides*, *unidentified_Dorea*, *Escherichia coli*, *Clostridium ramosum*, *Ruminococcus gnavus* at the species level (logFC > 2, $p < 0.05$) compared with the controls (Supplementary Table 2). Additionally, compared with the CP group, the microbiota of the control group showed an increased abundance of *Actinobacteria* at the phylum level ($p < 0.05$; Fig. 3 and Supplementary Table 1); and *Bifidobacterium*, *Faecalibacterium,* and *Gemmiger* at the genus level ($p < 0.05$; Fig. 3 and Supplementary Table 1); *Bifidobacterium adolescentis*, *unclassified_Bifidobacterium*, *Bifidobacterium longum*, *Faecalibacterium prausnitzii*, *unidentified_Bifidobacterium* at the species level (logFC > 4, $p < 0.05$; Supplementary Table 2). The analysis of the random forest indicates that our results exhibit a significantly high level of accuracy (0.88; Table 3).

Subsequently, we evaluated the predictive capability of specific bacterial species in relation to colorectal polyps using ROC curves. *Escherichia*, *Shigella*, and *Bacteroides* were employed to identify biomarkers indicative

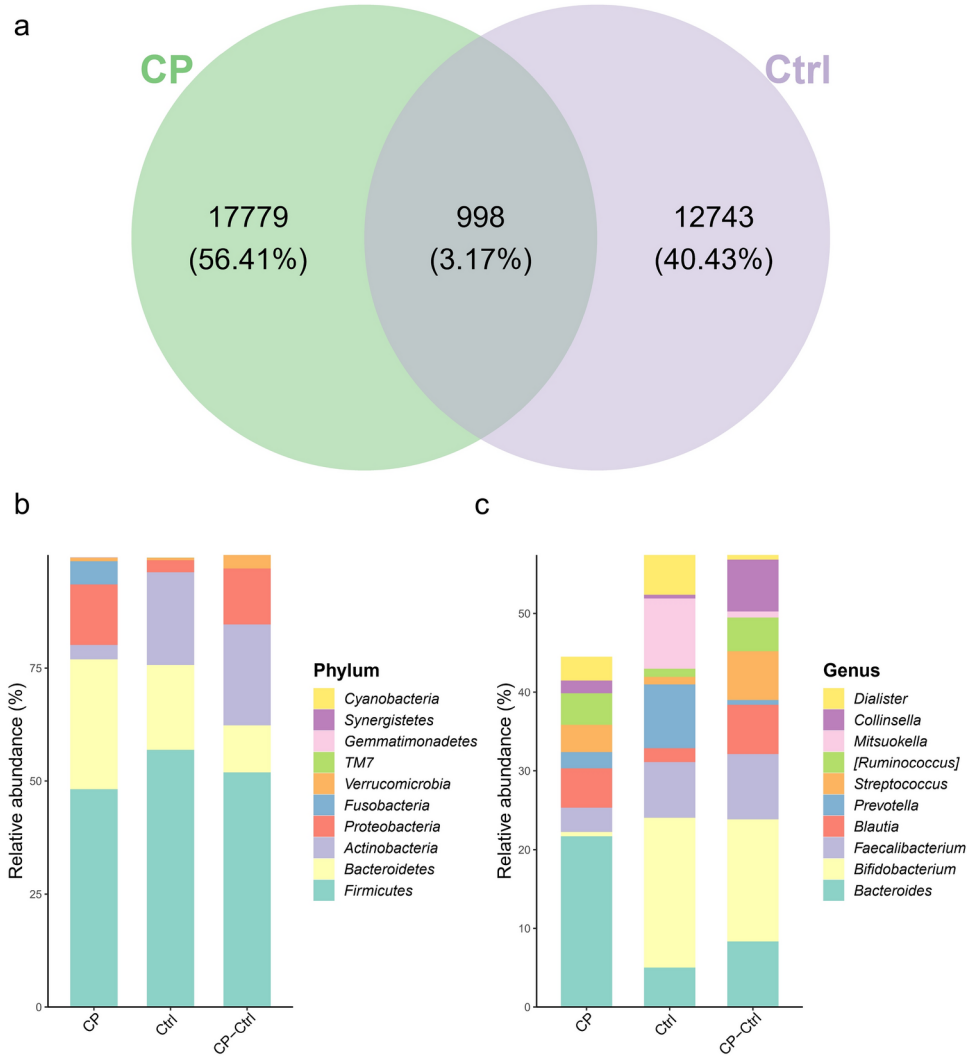

Fig. 2. Microbial abundance profiling between the CP and control groups. (**a**) Venn diagram showing the number of the overlapping OTUs between the CP group and control group (**b**) Taxonomic composition type at the phylum level, (**c**) Taxonomic composition at the genus level. CP: colorectal polyps. OTUs: Operational Taxonomic Units. CP-Ctrl: CP versus control.

of colorectal polyps. The ROC analysis results indicated a good predictive capability of these bacterial species for colorectal polyps, with corresponding AUC values of 0.723, 0.712, and 0.705 (Fig. 4).

Discussion

In this study, we performed 16S rRNA sequencing on stool samples obtained from CP and control groups. The objective of this study was to identify gut microbiota related to polyps, primarily tubular adenomas, and investigate potential biomarkers linked to them. Here, we found that alpha diversity (including Chao1 index, Shannon index, and Simpson index) does not exhibit any significant differences between the CP and control groups. However, the beta diversity of the gut microbiota in the CP group significantly differs from that of the control group. Previous studies on colorectal polyps have yielded inconsistent findings in terms of community diversity[20–22]. Some studies reported no significant differences in microbial diversity, while other studies found higher microbial diversity in patients with polyps. These discrepancies could potentially be influenced by various factors, including sample size, statistical power, and the presence of microbial drivers specific to certain populations. Furthermore, using LEfSe analysis, we observed significant differences in the microbial composition of gut microbiota between the two groups.

The majority of colorectal cancer cases arise from the presence of colorectal polyps. The intestinal microbiota has been identified as an important factor in promoting the progression from colorectal polyps to colorectal cancer[6]. In this study, we found that the predominant components of the intestinal microbiome in both groups are composed mainly of *Firmicutes*, *Bacteroidetes*, *Actinobacteria*, and *Proteobacteria*. *Firmicutes* and *Bacteroidetes* are two types of obligate anaerobic bacteria found in the intestinal tract, which have important functions in maintaining human health. Imbalance in these bacteria leads to inflammatory bowel disease,

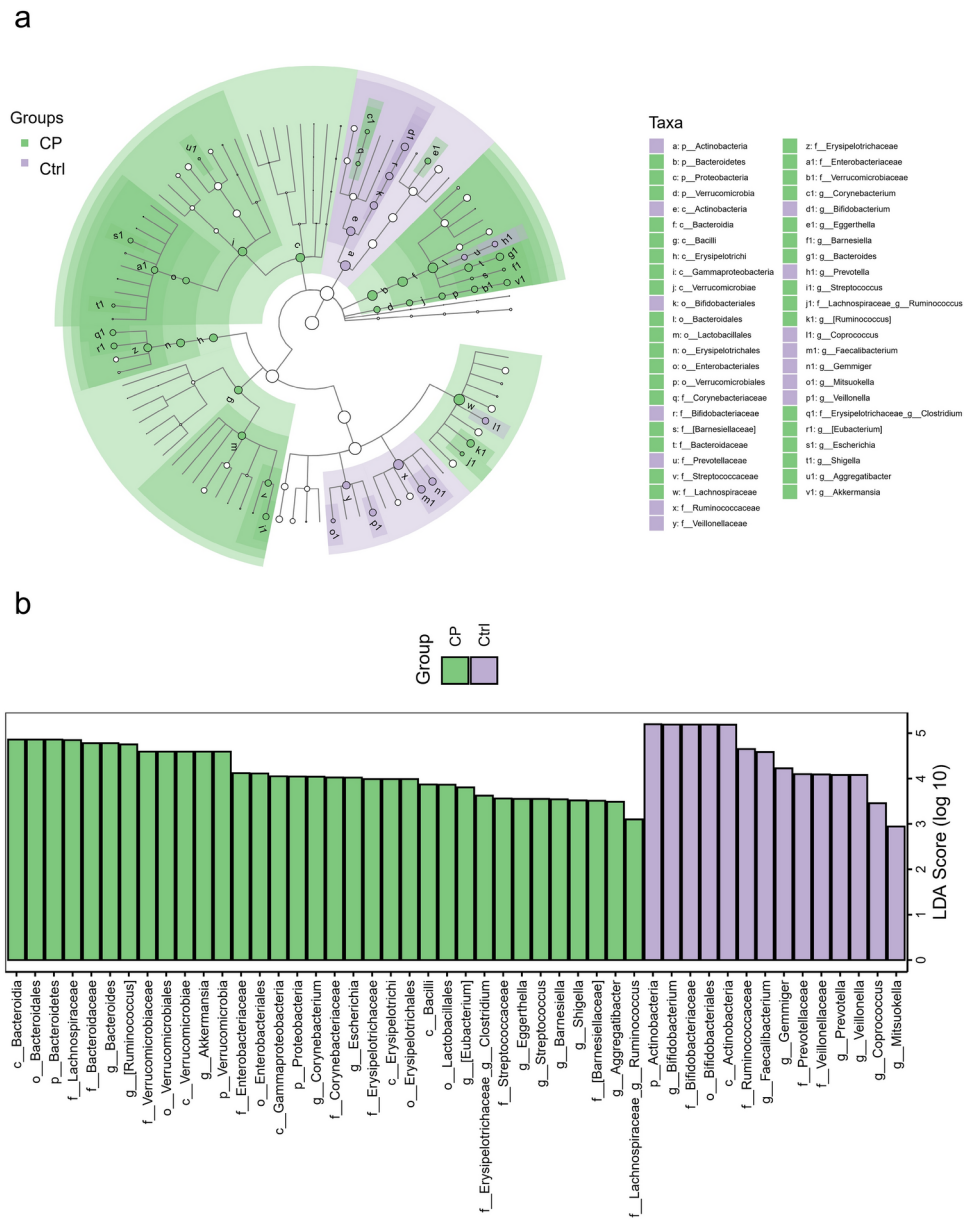

Fig. 3. Linear discriminant analysis effect size analysis of the CP and control groups. (**a**) Cladogram indicating the phylogenetic distribution of microbiota correlated with the CP or control groups. (**b**) The differences in abundance between the CP and control groups. CP: colorectal polyps.

	Control	colorectal polyps	Overall Accuracy
Control	0.833333	0.166667	
Colorectal polyps	0.066667	0.933333	
Overall Accuracy			0.883333
Baseline Accuracy			0.5
Accuracy Ratio			1.766667

Table 3. Accuracy of feces microbiota as a potential tool for detecting colorectal polyps.

formation of adenomas, and development of CRC[23,24]. Remarkably, the relative abundance of *Fusobacteria* in the CP group is higher than control group. *Fusobacteria*, a group of Gram-negative bacteria commonly present in the gastrointestinal tract, have been linked to a range of diseases[25,26]. A high abundance of *Fusobacterium* has been observed in CRC, yet its role in disease development remains elusive[26,27]. The abnormal increase of

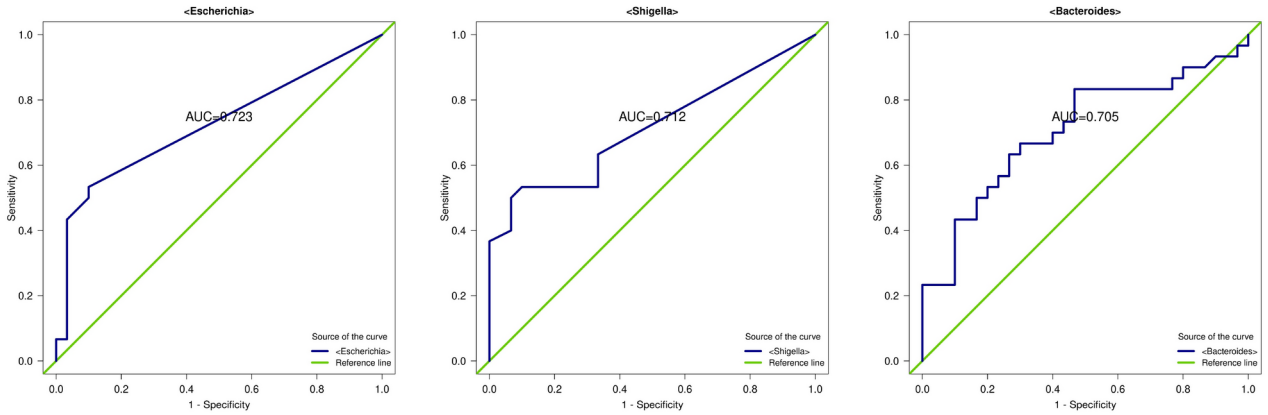

Fig. 4. ROC analysis of the association of *Escherichia*, *Shigella*, and *Bacteroides* with colorectal polyps. ROC: receiver operating characteristic.

Fusobacteria detected in feces may be related to colorectal polyps. However, further studies with larger sample sizes are required to confirm this correlation.

The human microbiome comprises more than 1000 microbial species and the gastrointestinal tract harbors approximately 100 trillion bacteria[28]. Although numerous types of bacteria are present in the gut, the abundance of each species varies widely. The gut microbiome consists of 30–40 bacterial species, accounting for over 99%[29]. These bacteria in the gut can be classified into three main groups based on their distinct physiological functions: pathogenic bacteria, conditional pathogens, and commensal bacteria. Commensal bacteria make up over 99% of the gut microbiota, producing beneficial substances that contribute to human health[30]. In our study, we observed a decreased abundance of commensal bacteria such as *Bifidobacterium*, *Faecalibacterium*, and *Gemmiger* in the microbiota of the CP group compared to the control group. In comparison to commensal bacteria, the presence of conditional pathogens in the gut is less common. However, the proliferation of these pathogens under specific conditions can be harmful to the body. Pathogenic bacteria, including *Salmonella*, *Shigella*, *Escherichia*, and *Klebsiella*, are directly responsible for causing diseases. In this study, using LEfSe analysis, we identified 34 distinct types of gut microbiota that showed significant differences between the CP and control groups. Specifically, *Escherichia* and *Shigella* were abundant in the CP group. *Shigella* is classified as a gram-negative intracellular bacterial pathogen that initiates infection by invading cells, leading to intense inflammation in the colonic and rectal epithelium. Notably, a recent study reported that Shigella is associated with malignancy[31]. In accordance with our findings, Zhong et al. reported that *Shigella* is abundant in the polyp and the mucosal sample[32]. Furthermore, we also found the abundance of *Clostridium perfringens*, *unidentified_Bacteroides*, *unidentified_Dorea*, *Clostridium ramosum*, and *Ruminococcus gnavus*. *Clostridium perfringens* and *Clostridium ramosum* have been reported to be associated with CRC[33,34]. The research conducted on mice revealed a significant increase in the abundance of *Bacteroides* in mice with CRC. *Ruminococcus gnavus* can produce pro-inflammatory polysaccharides, which disrupt the integrity of the intestinal mucosal barrier and promote inflammation[35–37]. These results indicate a structural imbalance in the composition of intestinal colonization flora, characterized by a decrease in beneficial bacteria and an increase in harmful bacteria. Therefore, we hypothesize that the imbalance of intestinal colonization flora could be a contributing factor to colorectal carcinogenesis.

There are limitations in this study. Firstly, since this study was conducted at a single center, it is crucial to validate our findings through multicenter studies involving larger sample sizes. Secondly, in contrast to metagenomic sequencing, 16S rRNA sequencing does not provide sufficient depth for species-level identification of specific bacteria[38]. Conversely, metagenomic sequencing allows for high-throughput sequencing of all microbial genomes in each sample and provides comprehensive species-level annotation[39]. Consequently, the inclusion of metagenomic sequencing data is frequently recommended for additional support. Thirdly, it is essential to acknowledge that fecal samples may not fully represent the entire landscape of the gut microbiota. Therefore, it is advisable to consider additional endoscopic sampling of gut mucosal samples for co-analysis to obtain a more comprehensive dataset of microbial information.

Overall, in this study, the abundance of gut microbiota at different levels was assessed in both CP patients and healthy individuals. We found that there is a structural imbalance in the composition of intestinal colonization flora, characterized by a decrease in beneficial bacteria and an increase in harmful bacteria. The presence of distinct microbial genera in each group suggests the presence of potential biomarkers for the disease. Our findings showed that *Escherichia*, *Shigella*, and *Bacteroides* may serve as promising biomarkers for early detection of colorectal polyps. Nonetheless, a larger sample size is necessary to further validate the candidate biomarkers.

Data availability

The datasets generated and/or analysed during the current study are available in the National Center for Biotechnology Information repository, available at http://www.ncbi.nlm.nih.gov/bioproject/1,180,202. The NCBI Sequence Read Archive (SRA) ID is SRA31178251-310。

Received: 15 October 2024; Accepted: 21 February 2025
Published online: 28 February 2025

References

1. Sung, H. et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **71**, 209–249. https://doi.org/10.3322/caac.21660 (2021).
2. Barberis, E. et al. A new method for investigating microbiota-produced small molecules in adenomatous polyps. *Anal. Chim. Acta* **1179**, 338841. https://doi.org/10.1016/j.aca.2021.338841 (2021).
3. He, X. et al. Long-term risk of colorectal cancer after removal of conventional adenomas and serrated polyps. *Gastroenterology* https://doi.org/10.1053/j.gastro.2019.06.039 (2020).
4. Biondi, A., Basile, F. & Vacante, M. Familial adenomatous polyposis and changes in the gut microbiota: New insights into colorectal cancer carcinogenesis. *World J. Gastrointest. Oncol.* **13**, 495–508. https://doi.org/10.4251/wjgo.v13.i6.495 (2021).
5. Cai, J., Sun, L. & Gonzalez, F. J. Gut microbiota-derived bile acids in intestinal immunity, inflammation, and tumorigenesis. *Cell Host Microbe* **30**, 289–300. https://doi.org/10.1016/j.chom.2022.02.004 (2022).
6. Tjalsma, H., Boleij, A., Marchesi, J. R. & Dutilh, B. E. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nat. Rev. Microbiol.* **10**, 575–582. https://doi.org/10.1038/nrmicro2819 (2012).
7. Liang, S. et al. Gut microbiome associated with APC gene mutation in patients with intestinal adenomatous polyps. *Int. J. Biol. Sci.* **16**, 135–146. https://doi.org/10.7150/ijbs.37399 (2020).
8. Lv, M. et al. Analysis of the relationship between the gut microbiota enterotypes and colorectal adenoma. *Front. Microbiol.* **14**, 1097892. https://doi.org/10.3389/fmicb.2023.1097892 (2023).
9. Rezasoltani, S. et al. The association between fecal microbiota and different types of colorectal polyp as precursors of colorectal cancer. *Microb. Pathog.* **124**, 244–249. https://doi.org/10.1016/j.micpath.2018.08.035 (2018).
10. Avelar-Barragan, J. et al. Distinct colon mucosa microbiomes associated with tubular adenomas and serrated polyps. *NPJ. Biofilms Microb.* **8**, 69. https://doi.org/10.1038/s41522-022-00328-6 (2022).
11. Kechin, A., Boyarskikh, U., Kel, A. & Filipenko, M. cutPrimers: A new tool for accurate cutting of primers from reads of targeted next generation sequencing. *J. Comput. Biol.* **24**, 1138–1143. https://doi.org/10.1089/cmb.2017.0096 (2017).
12. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Method.* **13**, 581–583. https://doi.org/10.1038/nmeth.3869 (2016).
13. Katoh, K., Misawa, K., Kuma, K.-I. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
14. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650. https://doi.org/10.1093/molbev/msp077 (2009).
15. Bokulich, N. A. et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**, 90. https://doi.org/10.1186/s40168-018-0470-z (2018).
16. DeSantis, T. Z. et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ. Microbiol.* **72**, 5069–5072. https://doi.org/10.1128/aem.03006-05 (2006).
17. Zgadzaj, R. et al. Root nodule symbiosis in Lotus japonicus drives the establishment of distinctive rhizosphere, root, and nodule bacterial communities. *Proc. Natl. Acad. Sci. U S A* **113**, E7996–E8005 (2016).
18. Segata, N. et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60. https://doi.org/10.1186/gb-2011-12-6-r60 (2011).
19. Intarajak, T. et al. Distinct gut microbiomes in Thai patients with colorectal polyps. *World J. Gastroenterol.* **30**, 3336–3355. https://doi.org/10.3748/wjg.v30.i27.3336 (2024).
20. Zhang, L. et al. Salivary and fecal microbiota: potential new biomarkers for early screening of colorectal polyps. *Front. Microbiol.* **14**, 1182346. https://doi.org/10.3389/fmicb.2023.1182346 (2023).
21. Clavenna, M. G. et al. Distinct Signatures of Tumor-Associated Microbiota and Metabolome in Low-Grade vs High-Grade dysplastic colon polyps: Inference of their role in tumor initiation and progression. *Cancers (Basel)*. https://doi.org/10.3390/cancers15123065 (2023).
22. Peters, B. A. et al. The gut microbiota in conventional and serrated precursors of colorectal cancer. *Microbiome* **4**, 69. https://doi.org/10.1186/s40168-016-0218-6 (2016).
23. Ahlawat, S., Asha, N. & Sharma, K. K. Gut-organ axis: a microbial outreach and networking. *Lett. Appl. Microbiol.* **72**, 636–668. https://doi.org/10.1111/lam.13333 (2021).
24. Quaglio, A. E. V., Grillo, T. G., De Oliveira, E. C. S., Di Stasi, L. C. & Sassaki, L. Y. Gut microbiota, inflammatory bowel disease and colorectal cancer. *World J. Gastroenterol.* **28**, 4053–4060. https://doi.org/10.3748/wjg.v28.i30.4053 (2022).
25. Fan, Z. et al. Fusobacterium nucleatum and its associated systemic diseases: epidemiologic studies and possible mechanisms. *J. Oral Microbiol.* **15**, 2145729. https://doi.org/10.1080/20002297.2022.2145729 (2023).
26. Tahara, T. et al. Fusobacterium in colonic flora and molecular features of colorectal carcinoma. *Cancer Res.* **74**, 1311–1318. https://doi.org/10.1158/0008-5472.CAN-13-1865 (2014).
27. Amitay, E. L. et al. Fusobacterium and colorectal cancer: causal factor or passenger? Results from a large colorectal cancer screening study. *Carcinogenesis* **38**, 781–788. https://doi.org/10.1093/carcin/bgx053 (2017).
28. Ley, R. E., Peterson, D. A. & Gordon, J. I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**, 837–848 (2006).
29. Kimura, K., McCartney, A. L., McConnell, M. A. & Tannock, G. W. Analysis of fecal populations of bifidobacteria and lactobacilli and investigation of the immunological responses of their human hosts to the predominant strains. *Appl. Environ. Microbiol.* **63**, 3394–3398 (1997).
30. Sears, C. L. A dynamic partnership: celebrating our gut flora. *Anaerobe* **11**, 247–251 (2005).
31. Mori, G. et al. Shifts of faecal microbiota during sporadic colorectal carcinogenesis. *Sci. Rep.* **8**, 10329. https://doi.org/10.1038/s41598-018-28671-9 (2018).
32. Zhong, X. et al. Gut microbiota signatures in tissues of the colorectal polyp and normal colorectal mucosa, and faeces. *Front. Cell. Inf. Microbiol.* **12**, 1054808. https://doi.org/10.3389/fcimb.2022.1054808 (2022).
33. Huang, C.-Y. & Wang, M.-C. Clostridium perfringens bacteremia associated with colorectal cancer in an elderly woman. *Turk J Gastroenterol* **31**, 960–961. https://doi.org/10.5152/tjg.2020.19987 (2020).
34. Yang, J. et al. Establishing high-accuracy biomarkers for colorectal cancer by comparing fecal microbiomes in patients with healthy families. *Gut Microbes* **11**, 918–929. https://doi.org/10.1080/19490976.2020.1712986 (2020).
35. Henke, M. T. et al. Ruminococcus gnavus, a member of the human gut microbiome associated with Crohn's disease, produces an inflammatory polysaccharide. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 12672–12677. https://doi.org/10.1073/pnas.1904099116 (2019).
36. Henke, M. T. et al. Capsular polysaccharide correlates with immune response to the human gut microbe Ruminococcus gnavus. *Proc. Natl. Acad. Sci. U.S.A.* https://doi.org/10.1073/pnas.2007595118 (2021).
37. Bell, A. et al. Elucidation of a sialic acid metabolism pathway in mucus-foraging Ruminococcus gnavus unravels mechanisms of bacterial adaptation to the gut. *Nat. Microbiol.* **4**, 2393–2404. https://doi.org/10.1038/s41564-019-0590-7 (2019).
38. Mignard, S. & Flandrois, J. P. 16S rRNA sequencing in routine bacterial identification: a 30-month experiment. *J. Microbiol. Method.* **67**, 574–581 (2006).

39. Franzosa, E. A. et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Method.* **15**, 962–968. https://doi.org/10.1038/s41592-018-0176-y (2018).

Acknowledgements
We thank Bioprofile Co. Ltd (Shanghai, China) for the assistance in deep-sequencing data analysis.

Author contributions
DD, YZ, and RC conceived and designed the study. DD, LZ, HC, and HC interpreted the analysis results and wrote the manuscript. DD, HS, and HW collected the samples and related information. DD and LZ wrote the manuscript. All authors revised the manuscript and approved the final manuscript.

Funding
This work was supported by Natural science research project of Jiading District, Shanghai (JDKW-2022–0030 and JDKW-2023–0051).

Declarations

Competing interests
The authors declare no competing interests.

Ethics approval
All experimental protocols were approved by the Ethics Committee of the Jiading Hospital of Traditional Chinese Medicine. Written informed consent was obtained from all participants. All methods were carried out in accordance with the Declaration of Helsinki and relevant guidelines and regulations.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-91626-4.

Correspondence and requests for materials should be addressed to Y.Z. or R.C.

Reprints and permissions information is available at www.nature.com/reprints.

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