Cancer Biol Med 2026. doi: 10.20892/j.issn.2095-3941.2025.0846

ORIGINAL ARTICLE

Comprehensive microbiome profiling reveals mucosal microbiome heterogeneity in patients with left- and right-sided colorectal neoplasia

Sunny H. Wong[1,2], Thomas N.Y. Kwong[3], Ru Zhang[3,4], Thomas Y.T. Lam[5], Andrew M.Y. Ho[3], Wing Wah Leung[6], Xing Kang[1,3], Harry C.H. Lau[3], Eagle S.H. Chu[3], Marc T.L. Wong[3], Louis H.S. Lau[3], Rashid N.S. Lui[3], Raymond S.Y. Tang[3], James Y.W. Lau[6], Simon S.M. Ng[6], Jun Yu[3,7], Joseph J.Y. Sung[1,2]

[1]Centre for Microbiome Medicine, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore 308232, Singapore; [2]Department of Gastroenterology and Hepatology, Tan Tock Seng Hospital, National Healthcare Group, Singapore 308433, Singapore; [3]Institute of Digestive Disease, State Key Laboratory of Digestive Disease, Li Ka Shing Institute of Health Sciences, Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China; [4]Shenzhen People's Hospital, Shenzhen 518020, China; [5]The Nethersole School of Nursing, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China; [6]Department of Surgery, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China; [7]CUHK-Shenzhen Research Institute, Shenzhen 518057, China

ABSTRACT

Objective: Left- and right-sided colorectal cancer (CRC) exhibit distinct molecular and clinicopathologic features. However, little is known about the spatial heterogeneity of microbial signatures. In this study the profiles and ecologic patterns of disease-associated intestinal microbiome were investigated in patients with an adenoma(s) or CRC at different anatomic locations.

Methods: A total of 690 stool, colonic aspirate, and mucosal biopsy samples were prospectively collected from 32 healthy, 30 adenoma, and 31 CRC patients.

Results: CRC was associated with alterations in fecal and mucosal microbiomes. Furthermore, the overall composition of the mucosal microbiome, stratified by metacommunities, differed between the patients with left- and right-sided neoplastic lesions. Patients with right-sided CRC had an elevated inter-phylum ecologic network, while patients with left-sided CRC had an enriched abundance of *Fusobacterium*. Interestingly, rectal neoplasia harbored a tumor microbiome that was distinctly different from the tumor microbiome at other anatomic sites.

Conclusion: The mucosal microbiome of right-sided CRC was distinctly different from the mucosal microbiome of left-sided CRC patients, suggesting distinct microbial ecology and heterogeneous host-microbial ecologic relationships that may contribute to differences in the tumor microenvironment between left- and right-sided CRC.

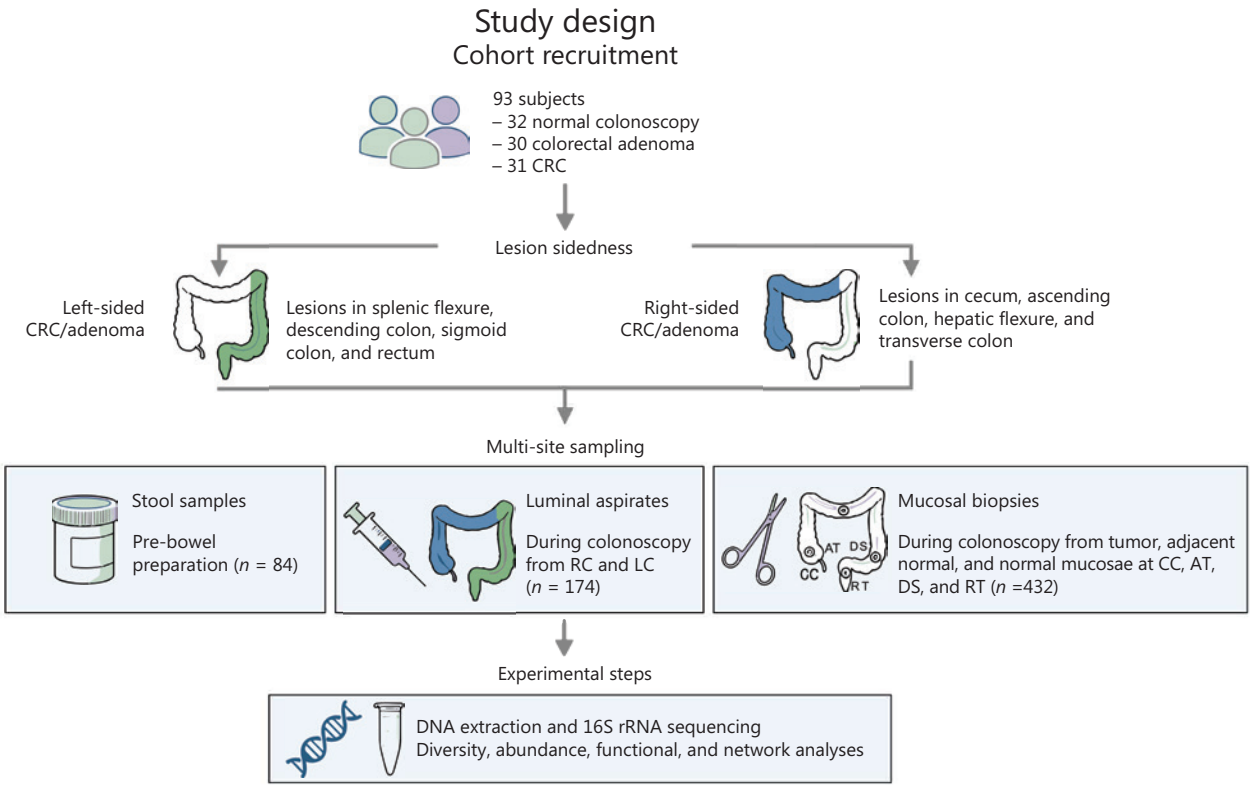
KEYWORDS Colorectal cancer; colorectal adenomas; colon carcinogenesis; colonic microflora; gut microbiome

Introduction

Colorectal cancer (CRC) is a significant global health burden[1]. The development of CRC is a complex process that is influenced by genetic and environmental factors. Recent studies have highlighted the importance of the gut microbiome in the pathogenesis of CRC. The CRC microbiome exhibits a significant compositional shift compared to the microbiome in healthy individuals with certain bacteria, such as *Fusobacterium nucleatum*, *Bacteroides fragilis*, and *Peptostreptococcus anaerobius*, implicated in tumor development and progression.

It is well-established that different subtypes of CRC preferentially arise at different locations within the colorectum. CRC that arises from the proximal (right-sided) colon relative to the splenic flexure differs from CRC that arises from the distal (left-sided) colorectum, including age- and gender-specific incidence rates, clinical characteristics, and pathologic features. In fact, patients with right-sided tumors are more likely to have the following characteristics: female gender; older age; exhibit advanced, mucinous, and signet-ring histology; respond differently to treatment; and have poorer disease

Correspondence to: Joseph J.Y. Sung
E-mail: josephsung@ntu.edu.sg
ORCID ID: https://orcid.org/0000-0003-3125-5199
Received January 8, 2026; accepted March 25, 2026;
published online May 26, 2026.
Available at www.cancerbiomed.org
©2026 The Authors. Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)



Study flowchart Study design and sample collection workflow. This schematic overview of the study design illustrates cohort recruitment, lesion classification, multi-site sampling, and downstream analyses. A total of 93 patients were recruited, including patients with normal colonoscopy findings (*n* = 32), colorectal adenomas (*n* = 30), and CRC (*n* = 31). Lesions were classified as left- (splenic flexure, descending colon, sigmoid colon, and rectum) or right-sided (cecum, ascending colon, hepatic flexure, and transverse colon). Multi-site sampling included pre-bowel preparation stool samples (*n* = 84), luminal aspirates collected during colonoscopy from the right and left colon (*n* = 174), and mucosal biopsies obtained from tumor, adjacent normal, and anatomically defined normal mucosae across the colorectum (cecum, ascending/transverse colon, descending/sigmoid colon, and rectum; *n* = 432). Microbial DNA was extracted from all samples and subjected to 16S rRNA gene sequencing, followed by diversity, differential abundance, functional inference, and microbial network analyses.

prognosis[2]. These differences are rooted in embryology with the proximal colon and distal colorectum developing from the midgut and hindgut, respectively[3]. In addition, right-sided tumors exhibit a distinct molecular profile that is characterized by oncogenic *BRAF* mutations, a positive CpG island methylator phenotype (CIMP), and aberrations in mismatch repair signaling. In contrast, left-sided tumors are characterized by more frequent loss of heterozygosity and *TP53* and *APC* mutations. These molecular differences were highlighted in a recent single-cell transcriptomic study that revealed a distinct genetic dichotomy with a predominance of the intrinsic Consensus Molecular Subtype (iCMS)3 subtype in right-sided tumors and the iCMS2 subtype in left-sided tumors[4].

Despite the well-established differences in molecular signatures and clinical features of CRC at different locations along the colorectum, there is still limited knowledge regarding the spatial heterogeneity of microbial signatures. A few studies have compared proximal and distal neoplasia in the colorectum and reported that *F. nucleatum* is more frequently detected in proximal[5-8], microsatellite instability (MSI)-high[9,10], and CIMP-high neoplastic lesions[11]. Conversely, few studies have noted *Fusobacterium* in patients with left-sided CRC[12-14], which is consistent with the mechanistic role of FadA adhesin in activating β-catenin/Wnt signaling[15,16]. Because studies have produced inconsistent results[17-20], the microbiome pattern of left- *vs.* right-sided CRC remains unclear.

In this study patients with colorectal tumors at different sites were recruited prospectively and the fecal, luminal, and tissue microbiomes on- and off-tumor locations were comprehensively profiled.

Methods

Patient recruitment

We recruited a Chinese cohort consisting of 32 patients with normal colonoscopy findings, 30 patients with

colorectal adenomas, and 31 patients with CRC. Participants were recruited prior to colonoscopy at the Shaw Endoscopy Centre in the Prince of Wales Hospital at the Chinese University of Hong Kong. Participants who had received systemic antibiotics within 8 weeks prior to colonoscopy or who had a history of inflammatory bowel disease, prior colorectal surgery, or hereditary colorectal cancer syndromes were excluded from the study. Baseline clinical variables, including smoking status, alcohol consumption, and comorbid conditions, were recorded and compared across study groups. None of these variables differed significantly between groups (all $P > 0.05$; **Table S1**) and were therefore not included as additional covariates to avoid model overfitting. Age and gender were included as covariates in all differential abundance and diversity analyses.

Fecal samples were collected before bowel preparation for colonoscopy, whereas aspirate and mucosal biopsy samples were collected during endoscopy. Samples were stored at −20°C within 4 h of collection and subsequently at −80°C for 24 h for long-term storage. Informed consent was obtained from all participants and the clinical study protocol was approved by the Joint Chinese University of Hong Kong and the New Territories East Cluster Clinical Research Ethics Committee (CREC 2018.251). Neither the patients nor the public were involved in the design, conduct, reporting, or dissemination of the research.

Sample collection, DNA extraction, and 16S rRNA gene sequencing

DNA extraction and purification for 16S rRNA gene amplicon sequencing were performed on a total of 690 samples, including 432 tissue biopsies, 174 aspirates, and 84 stool samples collected from 93 patients (**Table S1**). Although stool collection was attempted for all participants, some samples were unavailable or excluded after quality control, resulting in the final number as above. The QIAamp PowerFecal Pro Kit (Qiagen, Hilden, Germany) was used to extract microbial DNA from stool and aspirate samples according to the manufacturer's instructions. The DNA concentration was measured using a Thermo Scientific NanoDrop (Thermo Fisher Scientific, Wilmington, DE, USA). For tissue biopsies, samples were first digested using lysozyme and mutanolysin enzymes (Sigma-Aldrich, Hong Kong, China), followed by bead-beating and DNA purification using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). The Illumina Primer pair, 515f (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806r (5′-GGACTACHVGGGTWTCTAAT-3′), targeting the V4 hypervariable regions of the 16S rRNA gene was used for sequencing on the Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA). The Invitrogen SequelPrep Normalization Plate Kit (Thermo Fisher Scientific, Waltham, MA, USA) was used for library cleanup and normalization.

Data processing and taxonomic assignment

Filtering, merging, and taxonomic assignment of amplicon sequencing libraries were performed using QIIME 2 2021.02[21]. Briefly, raw paired-end sequencing reads were denoised and quality-filtered using the QIIME2. The filtered reads were trimmed to 220 bases, merged, and de-replicated into non-redundant amplicon sequence variants (ASVs) using the DADA2 plugin[22]. Taxonomic assignment to ASVs was performed using the feature classifier plugin with a pre-trained naive Bayes classifier against the GreenGene 13_8 99% 16S rRNA gene reference sequence database[23]. Rooted phylogenetic trees were constructed based on the alignment of representative sequences using mafft and FastTree. The feature table and phylogenetic tree were imported into R using the phyloseq package[24].

Raw ASV inference generated a feature table containing 98,632 ASVs prior to trimming. After quality trimming and denoising using the DADA2 pipeline, the resulting feature table contained 47,453 ASVs. A multi-step filtering procedure was applied to remove contaminating taxa and reduce sparsity in the dataset. First, potential contaminant ASVs were identified using the *decontam* R package based on the relationship between ASV abundance and extracted DNA concentration in each sample (frequency method)[25]. Second, the filtered feature table was further processed by removing ASVs assigned to taxa commonly associated with reagent contamination or laboratory environments, as well as sequences derived from mitochondria or chloroplasts[26]. Third, ASVs with a prevalence <1% across samples (presence defined as ≥5 reads) or with ambiguous taxonomic assignment (e.g., taxa resolved only to the kingdom level) were removed from downstream analyses. The dataset contained 5,127 taxa after these filtering steps. Taxa were subsequently collapsed to the species level and rarefied for downstream analyses, resulting in 399 species in the final analyses. These steps reduced sparsity and mitigated the impact of excessive zero counts typical of microbiome datasets. Six negative control samples were included, consisting of two extraction blanks for each sample type (stool, aspirate, and biopsy), to monitor potential contamination during DNA extraction and sequencing, and processed alongside biological samples.

Spatial heterogeneity of microbial composition

Microbial biomass differed among sample types (e.g., stool *vs.* biopsy). Therefore, rarefaction was performed in a sample type-specific manner. Stool, aspirate, and biopsy samples were rarefied to the lowest total reads in the respective sample types (47,000, 19,000, and 2,500 reads, respectively). To assess the diversity and composition of the microbiome, ASVs were collapsed into taxonomic aggregates at higher taxonomic ranks, such as the family or phylum level. Alpha diversity was

measured by calculating evenness, the Shannon-Weaver index, and phylogenetic diversity. Beta-diversity, which evaluates microbial composition among samples, was assessed based on the Bray-Curtis dissimilarity or unweighted UniFrac phylogenetic diversity. The R package vegan was used to compute the diversity metrics[27]. The microbial composition diversity was ordinated by principal coordinate analysis (PCoA) using the pcoa() function from the ape package[28] and visualized with a ggplot. Pairwise permutation multivariate analysis of variance (PERMANOVA) was performed based on the unweighted UniFrac metric with 999 iterations using the adonis2 function of the vegan package to evaluate the microbial composition among different sampling locations or fecal samples from patients with left- or right-sided lesions. Pairwise microbial composition with a PERMANOVA false discovery rate (FDR) adjusted $P < 0.05$ was considered statistically significant. Interaction terms were not included in the adjustment due to the limiting subgroup sample sizes.

Differential abundance taxa across groups

The compositional structure of each patient's fecal microbiome was evaluated and stratified into different metacommunities using an unsupervised learning approach prior to performing a differential abundance analysis. First, ASV counts were consolidated at the family level and transformed into relative abundance. Core taxa with a minimum prevalence of 0.1 and a minimum abundance of 0.05 were then selected. The Jensen-Shannon Divergence (JSD) was computed among core taxa and the clustering process was iteratively performed for each clustering number (max_k from 1 to 8) using the partition around medoids (PAM) algorithm to determine the optimal number of clusters. The silhouette information was calculated for each max_k value to assess the clustering performance and the optimal number of clusters was determined using the max_k value with the highest silhouette information value. The PAM algorithm and estimation of silhouette information were performed using the R package cluster[29]. Core taxa were also fitted to a Dirichlet-multinomial model (DMM) for clustering using the R package Dirichlet multinomial[30]. The optimal number (k) of clusters was determined using the minimized Laplace criterion to assess the goodness of fit. Dunn's index, which estimates the ratio of intra-cluster to inter-cluster UniFrac distances using the R package clValid[31], was used to determine the clustering performance. A regression analysis was performed for metacommunity with other potential variables (e.g., gender and age) to assess whether metacommunity could be a confounding factor. Differential abundance analysis was performed using ANCOM-BC with age, gender, and metacommunity included as covariates and modelled as fixed effects within the linear regression framework[32]. Samples from patients with bilateral neoplasia ($n = 4$) were excluded from the analyses. Taxa exhibiting a strong effect on typing metacommunity were selected using a random forest-based feature selection algorithm with the R package, Boruta[33]. *P*-values in the differential abundance analysis were adjusted for multiple comparison corrections using the Bonferroni method. Given the subgroup stratification, analyses involving anatomic sites along the colorectum and by microbiome metacommunity classification were considered exploratory and performed to generate hypotheses regarding spatial heterogeneity of the colorectal microbiome.

Microbial network analysis and functional profiling

The microbial correlation network was estimated using Fastspar, a C implementation of the SparCC algorithm that can infer correlation coefficients from compositional datasets[34]. Differentially abundant taxa and metacommunity markers were selected to construct a correlation network. The exact *P*-values and correlation coefficients were calculated based on 1,000 iterations and corrected for multiple comparisons using the FDR. The correlation matrix generated by Fastspar was analyzed in R using the igraph package[35] and network analysis and visualization were conducted using Cytoscape (version 3.9.1)[36]. Correlations with a $|\rho| > 0.3$ and FDR-adjusted $P < 0.05$ were retained for network visualization. The $|\rho| > 0.3$ threshold was selected to focus on moderate-to-strong correlations while excluding weaker associations that are more likely to reflect noise. Functional abundance of the microbiome was inferred using PICRUSt2 (version 2.4.1) based on quality-filtered representative sequences[37], although microbial gene or metabolic activity was not directly measured. ASVs were searched against the expanded Human Oral Microbiome database (eHOMD) 16S rRNA RefSeq database (version 15.22) using HOMD's online BLASTn program[38] to identify oral-associated taxa. ASVs with a BLAST hit and a minimum percent identity of 97% were considered oral-associated. The functional pathway comparisons were adjusted using the FDR where applicable.

Results

Patient samples and characteristics

In this study the microbiome profiles of different CRC phenotypes were investigated by performing 16S rRNA gene sequencing on fecal, luminal, and mucosal biopsy samples collected from patients with a normal colorectum ($n = 32$), histologically proven adenomas ($n = 30$), and invasive adenocarcinoma at different sites of the colon and rectum ($n = 31$). Patients were recruited from the Prince of Wales Hospital of the Chinese University of Hong Kong. The baseline characteristics of the patients are summarized in **Table S1**.

A stool sample was collected from each participant before bowel preparation for colonoscopy. During colonoscopy two separate colonic aspirate samples were taken from the distal colorectum (either at the rectum or sigmoid colon) and proximal colon (either from the ascending colon or cecum). Mucosal samples were taken from tumor tissues (cancers or polyps >1 cm) and normal mucosae at 2 sites in the right colon (ascending or transverse colon and cecum) and 2 sites in the left colorectum (descending or sigmoid colon and rectum; **Figure 1A**). This procedure resulted in 84 stool samples, 174 colonic aspirate samples, 432 tumors, and colorectal biopsy samples.

Altered composition of the mucosal microbiome in the CRC

The microbial diversity was first examined in all patient samples without stratification into left- or right-sided groups. Stool samples exhibited the highest microbial richness and diversity regardless of the disease phenotype, followed by biopsy and aspirate samples before and after rarefaction ($P < 0.05$; Chao1 and Shannon's diversity; **Figures 1B and S1**). Moreover, the microbial compositions across sample types were significantly different, as shown by the distinct clusters in the PCoA plot (PERMANOVA, $P < 0.05$; **Figure 1C**). Generally, biopsy samples were dominated by taxa from Firmicutes and Bacteroidetes, whereas fecal samples contained a notable proportion of Bacteroidetes (**Figure 1D**). Fusobacteria were prevalent in all sample types across disease phenotypes. Mucosal-associated *Akkermansia* spp. from the phylum Verrucomicrobia were specifically associated with the biopsy samples (**Figure 1D**). When comparing taxa among sample types across disease phenotypes some taxa were more prevalent in specific sample types and disease phenotypes. For example, *Leptotrichia* spp. were almost exclusively observed in biopsy samples of CRC patients but were nearly absent in stool samples. Similarly, *Helicobacter* spp. were present in one-half of the aspirate samples but were nearly absent in stool and biopsy samples from patients with CRC. This finding suggests that stool and aspirate samples may only partially reflect the microbiome present in the tumor and adjacent mucosa.

Community typing across biogeographic niches

Partitioning around medoids (PAM) and Dirichlet multinomial mixture (DMM) models were used to partition microbiome communities into clusters. The performance of the clustering algorithms was evaluated using Dunn's index. The analysis showed that clustering fecal and mucosal samples using the PAM algorithm resulted in a better Dunn index. The fecal microbiome was partitioned into three metacommunities (fecal MC1-3; **Figures 2A and B**), whereas five mucosal metacommunities were identified (mucosal metacommunities A–E; **Figures 2C and D**). Fecal MC1 was characterized by phylotypes from the Bacteroidaceae family, including representative members of *Bacteroides caccae*, *B. uniformis*, and *B. plebeius*. MC2 was distinguished by an expansion of Ruminococcaceae, whereas MC3 was dominated by *Prevotellaceae*. The combination of metacommunities varied across the disease phenotypes (CRC *vs.* normal: $P = 0.05$; CRC *vs*. adenoma: $P = 0.04$; **Figure 2B**). A high proportion of MC2 was noted in CRC but was low in adenomas (61% *vs.* 26%, $P = 0.05$). Five metacommunities were identified for the mucosal microbiome based on core taxa with a mean relative abundance ≥0.1% (**Figure 2C**). These metacommunities were designated as A–E, which was consistent with our previous study[39] and defined based on the same set of biomarkers. Briefly, Bacteroidaceae was enriched in metacommunities A and E. Members of Lachnospiraceae and Ruminococcaceae predominated metacommunity B, whereas metacommunity D harbored a higher abundance of taxa from Lactobacillaceae and Verrucomicrobiaceae (primarily *Akkermansia muciniphila*). An elevated abundance of CRC-associated Fusobacteria was observed in metacommunities C, D, and E.

The distribution of the mucosal metacommunity was significantly different across disease phenotypes. Specifically, CRC had expansion of metacommunities D and E and shrinkage of metacommunities A and B compared to the adenomatous and normal mucosae (**Figure 2D**; CRC *vs.* normal: $P < 0.01$; CRC *vs*. adenoma: $P < 0.01$). Notably, left- and right-sided tumors encompassed different distributions of mucosal metacommunities (**Figure 2E**; $P = 0.0047$). Greater than 95% of biopsy samples from patients with left-sided CRC belong to metacommunity D, whereas right-sided CRC samples are classified almost dichotomously into metacommunities D and E. The abundance of oral-associated taxa, as defined by eHOMD, differed among metacommunities with samples assigned to mucosal metacommunity D and fecal MC2 having the highest abundance of oral taxa (mean abundances of 20% and 7%, respectively; **Figure 2F**).

Next, microbial composition was assessed using PERMANOVA based on unweighted UniFrac and Bray-Curtis distances. The overall composition of the fecal, luminal, and mucosal microbiomes in CRC patients (both right- and left-sided) dramatically shifted from the overall composition in normal subjects and adenoma patients (**Figures 3 and S2A, B**). Indeed, the mucosal microbiome in CRC patients was significantly different from the mucosal microbiome in normal subjects and adenoma patients, even in the adjacent normal (off-tumor) mucosae of CRC patients (**Figures S2C and D**). Differences in the mucosal microbiome were detected by PERMANOVA between patients with left- and right-sided CRC (**Figure 3**), providing further evidence that the mucosal microbiome of CRC, as well as their adjacent mucosa, were distinctly different between the left and right sides of the colon.

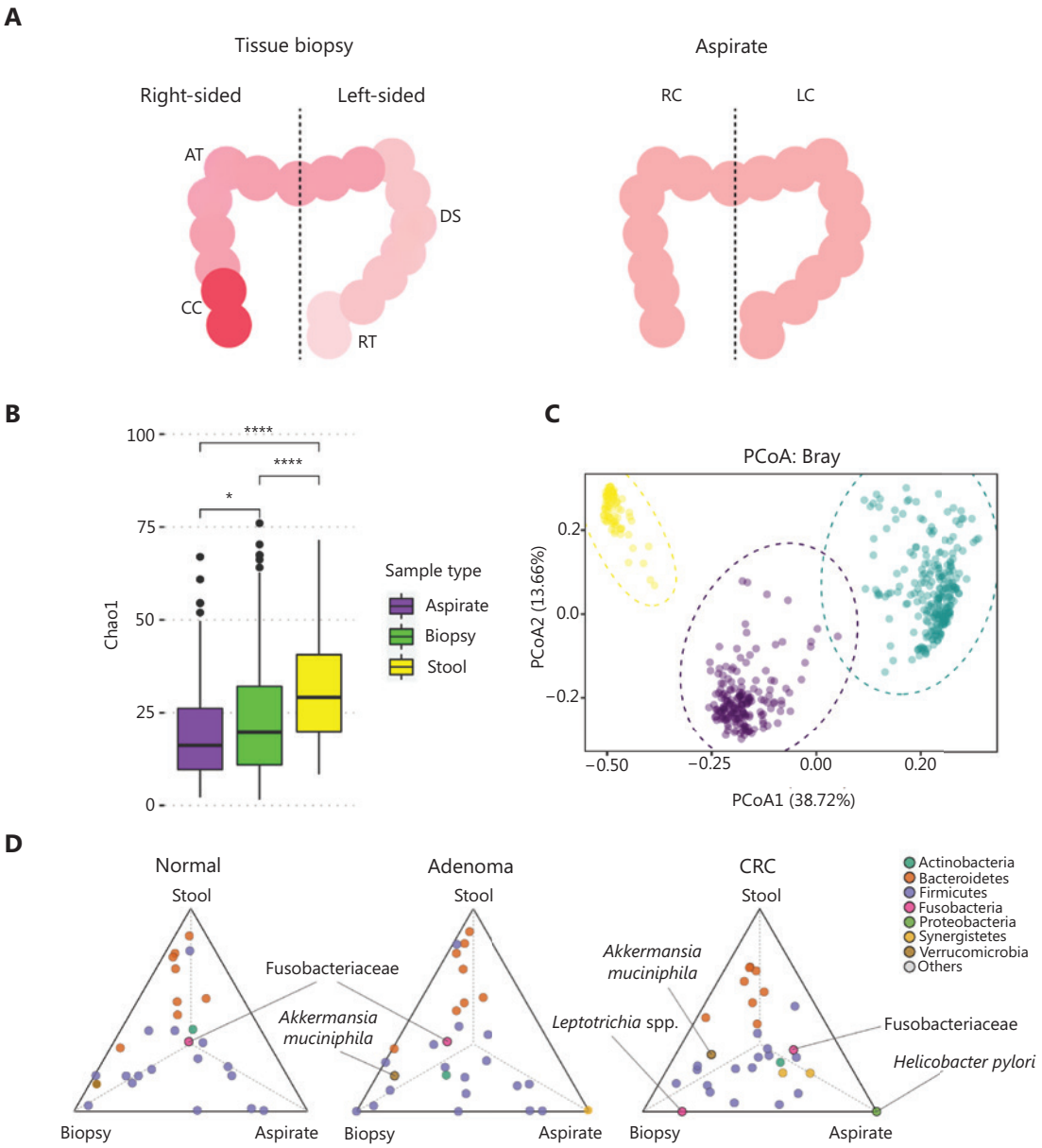


Figure 1 (A) Strategy for sample collection in this study. Patients were grouped into left- or right-sided according to the site of the lesion(s). Patients with lesions from the cecum, ascending colon, hepatic flexure, and transverse colon were assigned to the right-sided (R-) group, while patients with lesions from the splenic flexure, descending colon, sigmoid colon, and rectum to the left-sided (L-) group. Tissue biopsy samples were designated as follows: AT, ascending colon, hepatic flexure, and transverse colon; CC, cecum; DS, splenic flexure, descending colon, and sigmoid colon; RT, rectum. Aspirate samples were grouped into RC or LC. (B) Alpha diversity among stool, aspirates, and biopsies. The Chao1 metric was calculated based on rarefied samples. (C) Principal coordinate analysis of stool, aspirates, and biopsies was based on the Bray–Curtis distance. (D) The ternary plot shows the prevalence of taxonomic families among sample types (aspirates, biopsies, and stool) across disease phenotypes (normal, adenomas, and CRC). Each circle represents a family and is color-coded according to the phylum. The prevalence of each taxonomic family was determined by collapsing ASVs (minimum read count, ≥ 5) from the same family, followed by counting the number of samples with the taxonomic family detected. Taxonomic families with a minimum prevalence of 0.4 are shown.

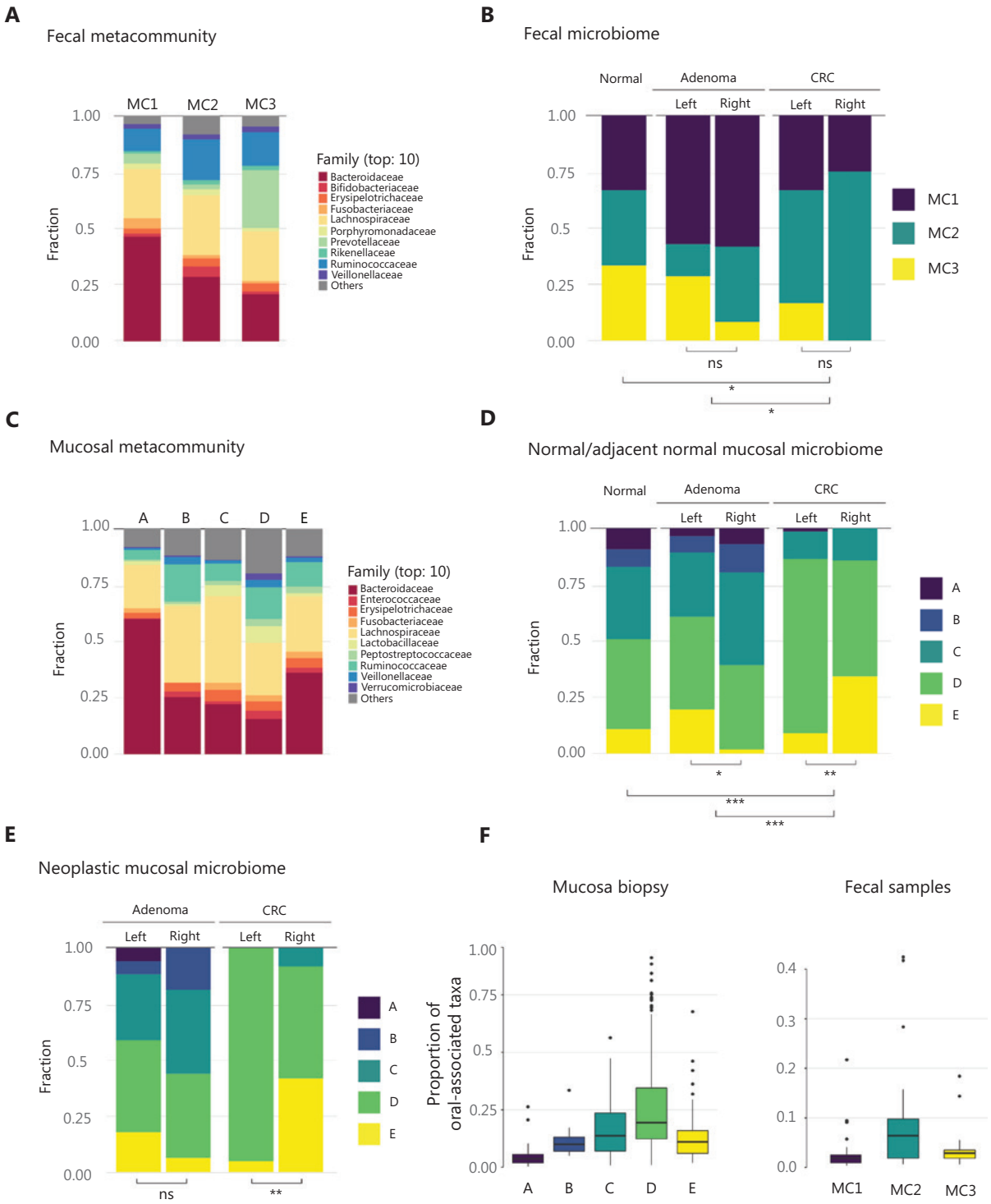


Figure 2 (A) The taxonomic profile for the top 10 abundant families in the fecal metacommunities. The fecal metacommunities (MC1–3) were detected in stool samples based on core taxa with a minimum prevalence of 0.1 using the PAM algorithm. (B) The distribution of the fecal microbiome metacommunities across disease phenotypes. Pairwise comparison on the proportion of the metacommunity among sample groups was performed using Fisher's exact test. CRC *vs.* normal: $P = 0.05$; CRC *vs.* adenoma: $P = 0.04$; L-adenoma *vs.* R-adenoma: $P = 0.41$; L-CRC *vs.* R-CRC: $P = 0.58$. (C) The taxonomic profile for the top 10 abundant families in the mucosal metacommunities. The mucosal metacommunities (A–E) were detected in biopsy samples based on core taxa with a minimum prevalence of 0.1 using the PAM algorithm. (D) The distribution

of the mucosal metacommunity in normal/adjacent normal mucosal microbiome across five sample groups (normal, L-adenoma, R-adenoma, L-CRC, and R-CRC). Five metacommunities (A–E) were detected in mucosal samples. Pairwise comparison on the proportion of the metacommunity between sample groups was performed using Fisher's exact test. CRC *vs.* normal: P = 1.2e–7; CRC *vs.* adenoma: P = 1.4e–7; adenoma *vs.* normal: P = 0.8; L-CRC *vs.* R-CRC: P = 0.009; L-adenoma *vs.* R-adenoma: P = 0.01. (E) The distribution of the five mucosal metacommunities in lesion biopsy samples (CRC and adenoma); L-CRC *vs*. R-CRC: P = 0.0047; L-adenoma *vs*. R-adenoma: P = 0.59. (F) The fraction of oral-associated taxa in mucosal and fecal samples among metacommunities. Representative sequences were searched against the eHOMD 16S rRNA Refseq database (version 15.22) using HOMD online BLASTn program. Representative sequences obtained BLAST hits with an identity ≥ 97% were considered as oral-associated taxa. ***$P \leq 0.001$, **$P \leq 0.01$, *$P \leq 0.05$. CRC, colorectal cancer; ns, not significant; PAM, partition around medoids.

Comprehensive tumor microbiome profiling along the colorectum

Individual taxon profiles were examined using a log-linear regression model with adjusted confounding factors to further investigate the altered microbial composition at different anatomic locations. *Fusobacterium* spp. and *Ruminococcus* spp. were enriched in fecal samples from left-sided CRC patients, whereas *B. plebeius* and *Prevotella* spp. were more abundant in right-sided CRC patients (**Figure 4A**). Unique microbial signatures specific to certain locations or shared between different locations along the colorectum were observed. Notably, the rectum of left-sided CRC patients had an enrichment of *Fusobacteria* spp., *Solobacterium moorei*, and *Parvimonas* spp., while the proximal colon of right-sided CRC patients exhibited elevated levels of *Eubacterium* spp. and *Holdenmanella biformis* (**Figure 4B**).

Heterogeneity in metagenome functional profiles and microbial ecology among mucosal-associated microbes

Co-occurrence analysis was performed using the SparCC algorithm to gain a deeper understanding of the ecologic network among taxa in mucosal microbiomes across different disease phenotypes, lesion sidedness (left- or right-sided), and biogeography along the colorectum. We focused on taxa that were identified as differentially abundant and/or metacommunity markers and obtained exact P-values using 1,000 iterative permutations. The microbial co-occurrence network exhibited distinct correlation patterns between left- and right-sided CRC patients (**Figure S3**). An inter-phylum polymicrobial correlation network was observed consisting of members from Firmicutes and *Bacteroides* in right-sided CRC, whereas such a network was nearly absent in left-sided CRC, including tumor tissues and adjacent normal mucosae. In contrast, left-sided CRC tumors manifested an increased abundance of *Fusobacterium* spp., which had a strong negative correlation with members of the polymicrobial network.

The strength and direction (co-occurring or -excluding) remained broadly similar in healthy subjects and within adenoma and CRC subjects when the overall pattern of microbial interactions was considered across the colorectum (average Pearson correlations of ~0.50; **Figure S4A–C**). Nevertheless, a lower degree of coherence was observed in the microbial interactions between normal and adenoma or CRC subjects (average Pearson correlations, ~0.33 and ~0.43, respectively; **Figures S4D and E**). With clear differences in the pattern of the microbial network between patients with left- and right-sided lesions. Coherence in microbial interactions was not detected between left- and right-sided CRC or adenoma patients (**Figure S4F**). The co-occurrence patterns (e.g., cluster blocks) in the cecum of right-sided CRC patients were undetectable in left-sided CRC patients (**Figures S4F and 5**), suggesting a high spatial heterogeneity in the microbial interactome.

Finally, the functional profiles associated with a shift in overall microbial composition were examined. PICRUSt2 was used to predict the metagenomic functional pathways of the mucosal microbiome based on representative 16S rRNA gene sequences. MetaCyc functional pathways were shown to be enriched in left- and right-sided CRC along different regions of the colorectum (**Figure S6**). Amine and polyamine degradation pathways, such as creatinine degradation I (CRNFORCAT-PWY) and glycine betaine degradation I (PWY-3661), were significantly enriched in the cecum of the left-sided CRC mucosal microbiome. The pathway for butanediol biosynthesis (P125-PWY) was enriched in the cecum and ascending colon in the right-sided CRC mucosa. The CO_2 fixation pathway, the reductive acetyl coenzyme A pathway (CODH-PWY), was strongly associated with mucosa along the entire left-sided CRC colorectum. In contrast, the right-sided CRC mucosa was associated with carbohydrate or nucleotide degradation, such as starch (PWY-6731), sucrose (PWY-3801) or lactose and galactose (LACTOSECAT-PWY).

Discussion

Fecal, luminal, and colorectal mucosal biopsies were prospectively collected from normal controls, patients with adenomas, and CRC patients to understand the ecology of the gut microbiome in primary CRC from different anatomic locations and comprehensive profiling of the microbiome was performed with extensive geospatial mapping. A

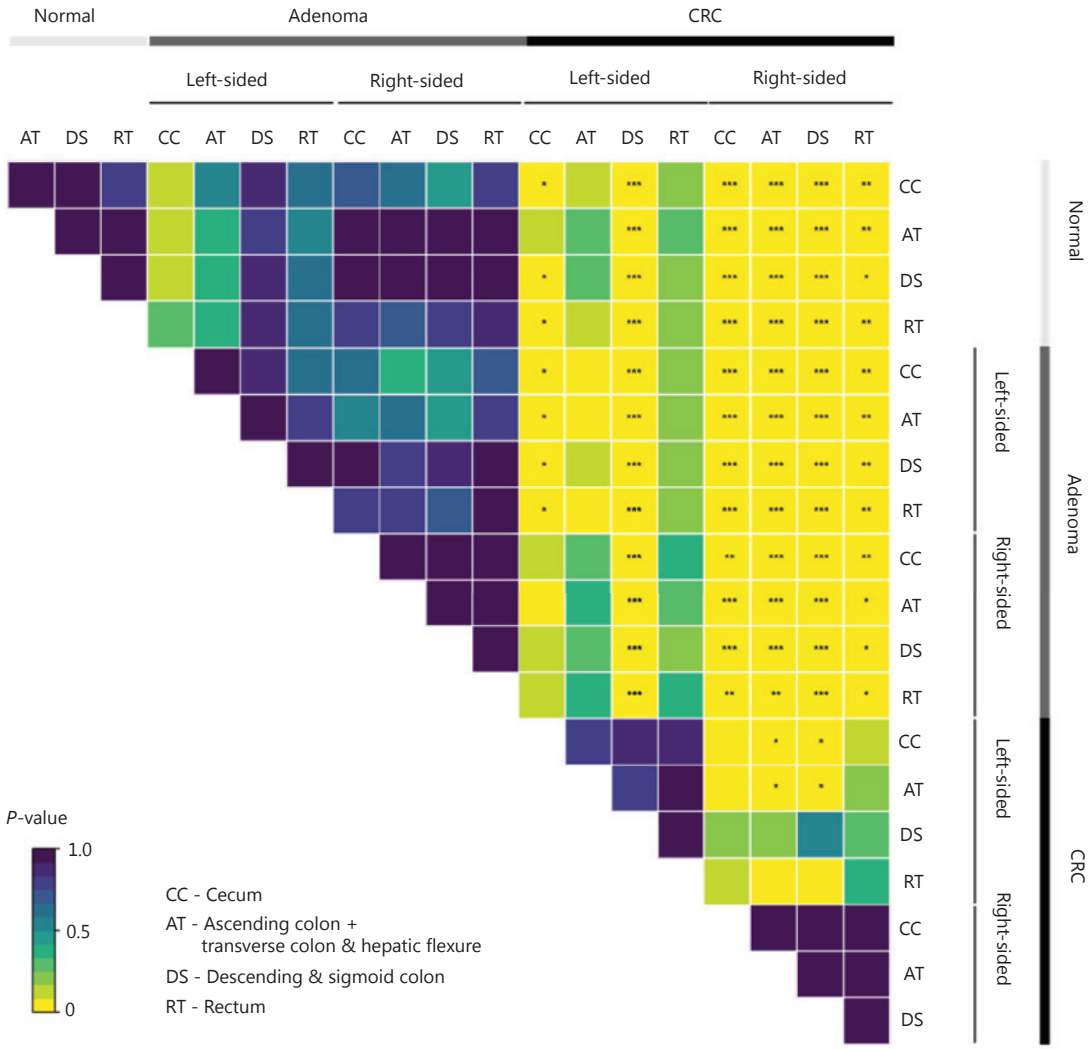


Figure 3 Pairwise difference in the composition of the mucosal microbiome in biopsy samples. Normal (healthy individuals) or adjacent normal mucosal samples (patients with adenomas or CRC) were compared using PERMANOVA on the unweighted UniFrac distance. Each square represents a pairwise compositional difference between sample groups (row *vs.* column) and was filled with a color according to the q-value (FDR-adjusted PERMANOVA *P*-value). Left- and right-sided indicate the location of lesions (adenoma or CRC) in the patients from whom biopsy samples were collected. AT, ascending colon, hepatic flexure, and transverse colon; CC, cecum; CRC, colorectal cancer; DS, splenic flexure, descending colon, and sigmoid colon; ns, not significant; RT, rectum. ***$P \leq 0.001$, **$P \leq 0.01$, *$P \leq 0.05$.

significant alteration of the gut microbiome in CRC samples compared to normal and adenoma samples was observed, which was consistent with previous findings, including our findings[39]. This alteration was evident from the drastic shift in the gut microbiome, as shown in the compositional ternary plots (**Figure 1D**), heatmaps of pairwise comparisons (**Figures 3 and S1A, B**), and metacommunity profiling (**Figures 2A and E**). These findings are consistent with the large-scale literature. Specifically, a pooled analysis of 3,741 stool metagenomes from 18 independent studies confirmed a reproducible CRC-associated fecal microbial signature across populations[18].

Moreover, extensive mucosal sampling across different locations provided additional insight into the geospatial ecology of the colorectum in this study. Notably, the presence of cancerous tumors was noted to extensively alter the mucosal microbiome. This finding was indicated by widespread changes in the microbiome, which extend beyond the tumor site to affect nearly the entire colorectum to mark contrasting differences between normal and adenoma samples (**Figure 3**). Given the established notion of the cancer field effect[40,41], this finding raises the possibility of a metagenomic-equivalent concept in which detectable microbiome changes are evident in non-tumorous

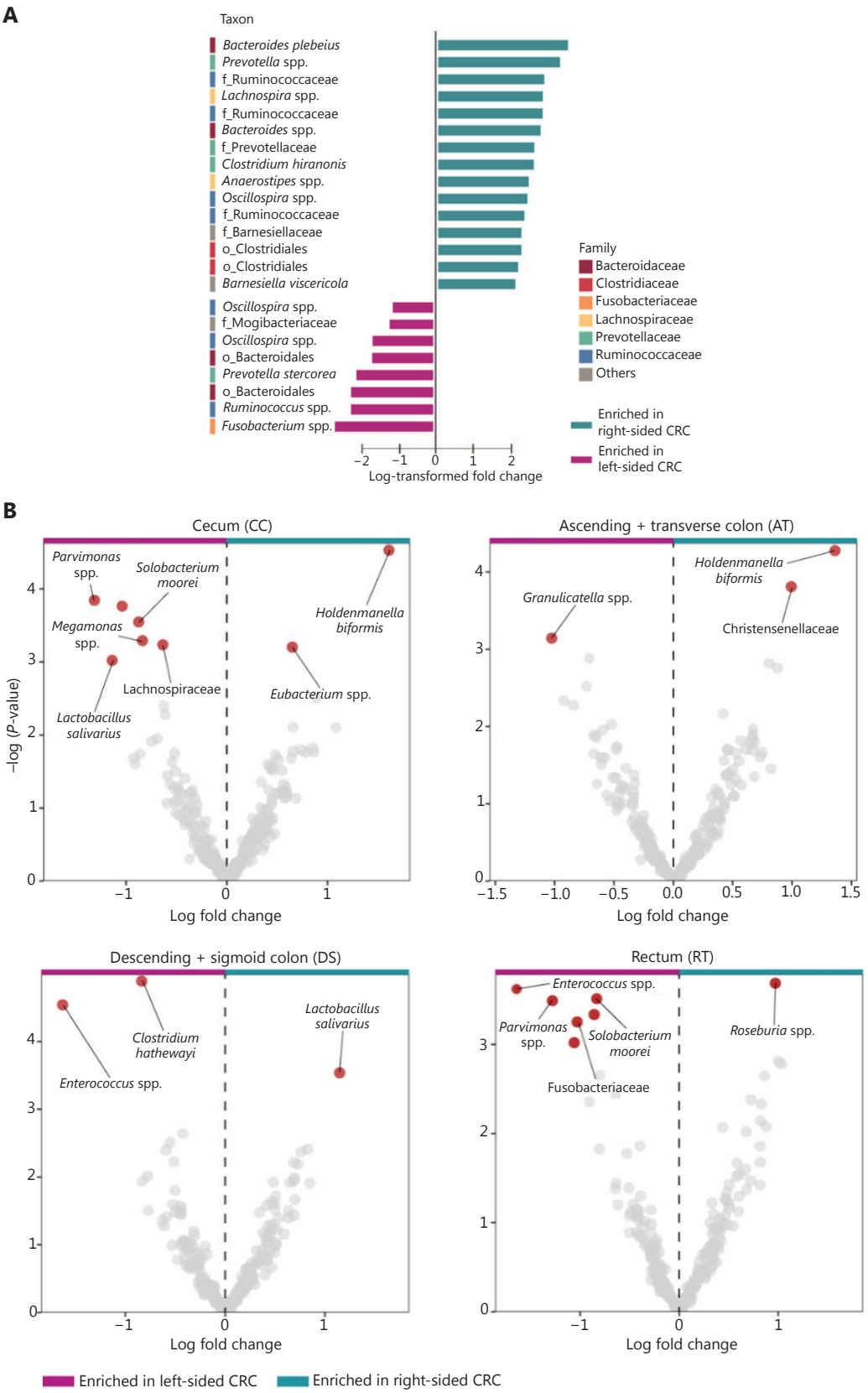


Figure 4 Continued

Figure 4 (A) Bar chart for differentially abundant taxa in fecal microbiome between left- and right-sided CRC patients. Significant differentially abundant taxa [*P*-value ≤ 0.05 and log-transformed fold change (lfc) ≥ 1.0, as determined by ANCOM-BC] are shown and color-coded according to the family. (B) Volcano plots for differentially abundant taxa in the mucosal microbiome between left- and right-sided CRC patients along the colorectum. Each circle represents a taxon. Each plot shows the lfc of relative abundance (x-axis) and negative log-transformed *P*-value (y-axis) for each taxon in mucosal biopsies sampled along the colorectum [cecum (CC), ascending colon, hepatic flexure, and transverse colon (AT), splenic flexure, descending colon, and sigmoid colon (DS), or rectum (RT)]. Significant differentially abundant taxa are highlighted in red color. Taxa with a negative lfc value were enriched in left-sided CRC, while taxa with a positive lfc value were enriched in right-sided CRC. *P*-values < 0.05 were considered statistically significant.

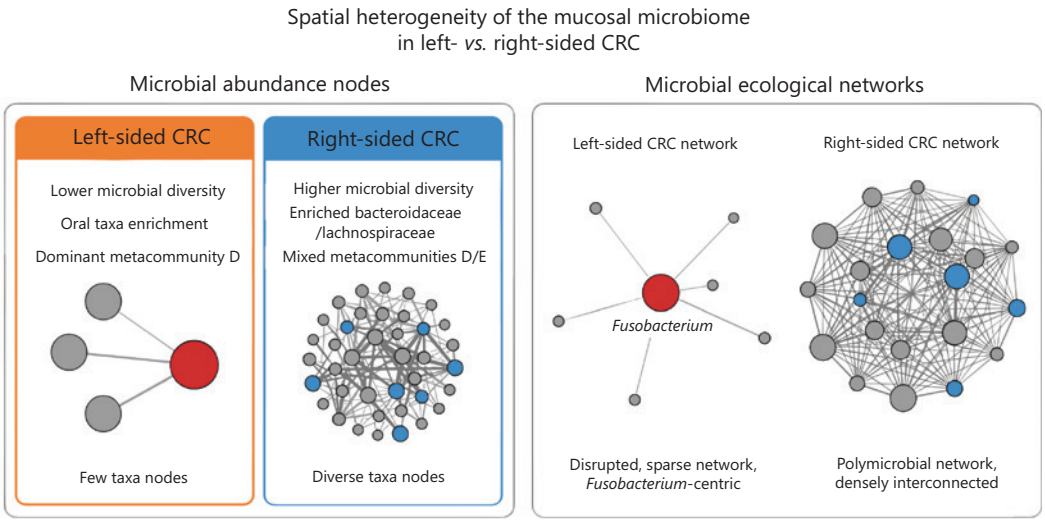


Figure 5 Summary of spatial microbiome heterogeneity in left- *vs.* right-sided CRC. The left panel illustrates the microbial abundance structures and the right panel depicts the ecologic interaction networks inferred from microbial correlations. CRC, colorectal cancer.

and normal-appearing parts of the colorectum. With the pervasive molecular changes in the cancer field, changes in host-microbial interactions may extend beyond the tumor site across the colorectum. Importantly, this dysbiosis may be reversible and represents a potential therapeutic target in corollary microbiome-directed studies.

Regarding the geospatial distribution of the microbiome, pairwise heatmap comparison, metacommunity analysis, and microbial ecologic analysis showed that there were differences in the mucosal microbiome between left- and right-sided CRC. The left-sided CRC was largely constituted by metacommunity D, whereas the right-sided CRC had more metacommunities C and E. Patients with left-sided CRC had enriched *Fusobacterium* and members of Lactobacillaceae and Ruminococcaceae, while patients with right-sided CRC had enriched members of Bacteroidaceae, Lachnospiraceae, and Peptostreptococcaceae, forming a correlation-based polymicrobial network (**Figure 5**).

The cancer-associated microbiome is affected by the tumor microenvironment, which is regulated by host immunity, inflammation, molecular signatures, and signaling pathways in neoplasia. Generally, left-sided CRCs exhibit chromosomal instability with *APC* and *p53* mutations, while right-sided CRCs are frequently associated with MSI and a hypermethylation phenotype. Given the transcriptomic dichotomy of CRC corresponding to the two intrinsic subtypes (iCMS2 and iCMS3) arising preferentially in different parts of the colorectum[4], the distinct microbiome may constitute a part of the tumor microenvironment. Nevertheless, no noticeable difference was detected in the mucosal microbiome between adenomas and paired adjacent normal tissues or between left- and right-sided adenomas, despite the evident transcriptomic changes in conventional adenomas and serrated polyps[42]. This finding is consistent with previous studies that showed readily discernible microbiome changes in CRC yet only modest alterations in pre-malignant adenomatous polyps.

This study had several limitations. First, recruitment was challenged by the need for two separate colonoscopes to minimize contamination and perform multiple additional samplings. This imposed additional procedural time and consideration for patient consent. Second, the present study enrolled 93 participants and incorporated extensive multi-site samplings. While this design enabled detailed microbiome

characterization across intestinal niches, the design resulted in relatively small sample sizes per subgroup. A formal *a priori* power calculation was not performed because the study was designed as an exploratory spatial microbiome profiling study involving extensive multi-site sampling. However, *post-hoc* estimation based on the observed cohort sizes (normal, n = 32; adenoma, n = 30; CRC, n = 31) suggests that two-group would provide approximately 80% power to detect standardized effect sizes of Cohen's $d \approx 0.7–0.8$ at an $\alpha = 0.05$, corresponding to moderate-to-large differences. Smaller subgroup analyses (e.g., left- *vs*. right-sided CRC) would require larger detectable effects (Cohen's $d \approx 1.0$). In addition, a metacommunity-based ecologic stratification approach was applied that reduced the dimensionality of microbiome data with better statistical power. The distribution of mucosal metacommunities differed significantly between left- and right-sided CRC, corresponding to a large ecologic effect size (Cramér's $V \approx 0.5$) and estimated *post-hoc* power of >80%. Third, this was a cross-sectional study and the results were largely based on associations or correlations. Whether the observed microbiome differences represent a driver or a consequence of the distinct molecular carcinogenesis pathways cannot be determined from this cross-sectional study. Experimental models are required to investigate casual effects of the observed changes. Functional profiling inferred from the 16S rRNA data represents predicted metabolic potential rather than direct measurement of microbial gene activity. Fourth, detailed dietary data were not available for this cohort and therefore the potential influence of dietary patterns on gut microbiome composition could not be evaluated. Finally, the 16S rRNA gene sequencing approach provides bacterial genus-, and in some cases, species-level resolution, but cannot distinguish between subspecies or clades within a species. This approach is particularly relevant for *F. nucleatum*, in which only the *Fna* C2 clade has been shown to specifically dominate the CRC tumor niche[43]. Detection of *Fusobacterium* by 16S sequencing does not distinguish between carcinogenic and non-carcinogenic clades. With the recent characterization of fungi across various cancers[44], the mycobiome and its interaction with the bacteriome may play an important role in cancer formation and therapeutic response.

This study provides unique insights into the microbial profiles of patients with right- and left-sided CRC. Our data indicated a clear presence of the CRC microbiome as a distinctive profile and our results suggested sufficient differences in the mucosal microbiome between right- and left-sided CRCs that may form part of the tumor microenvironment. These microbiome differences may have implications for CRC-related onco-therapeutic responses, although this requires validation in future studies. Future studies with larger cohorts, longitudinal samples, shotgun sequencing, and molecular characterization will be required to validate the observed microbiome patterns and clarify the effects on cancer biology. We anticipate that the results of this study will provide important information for using microbiomes as biomarkers to identify CRC and predict the therapeutic response to this disease.

Acknowledgements

The authors are grateful to Dr. Siu-Kin Ng for his assistance with the data analysis.

Grant support

The work was supported by the Singapore National Medical Research Council Clinician Scientist Individual Research Grant (Grant No. CIRG23jan-0004), Large Collaborative Grant (Grant Nos. OFLCG22may-0009, OFLCG23may-0031, and OFLCG24may-0025), and the Wang Lee Wah Memorial Fund.

Conflict of interest statement

No potential conflicts of interest are disclosed.

Authorship contributions

SHW, TNYK, and RZ contributed to the formal analysis, experimentation, data collection, and drafting of the manuscript. TYTL, RZ, AMYH, WWL, XK, HCHL, ESHC, MTLW, LHSL, RNSL, RSYT, JYWL, and SSMN contributed to the provision of the study materials and laboratory samples. JY and JJYS contributed to the conceptualization, oversight, and planning of this project, and critically reviewed the manuscript.

Data availability statement

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022; 72: 7-33.
2. Nawa T, Kato J, Kawamoto H, Okada H, Yamamoto H, Kohno H, et al. Differences between right- and left-sided colon cancer in patient characteristics, cancer morphology and histology. J Gastroenterol Hepatol. 2008; 23: 418-23.
3. Carethers JM. One colon lumen but two organs. Gastroenterology. 2011; 141: 411-2.

4. Joanito I, Wirapati P, Zhao N, Nawaz Z, Yeo G, Lee F, et al. Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and refines the consensus molecular classification of colorectal cancer. Nat Genet. 2022; 54: 963-75.
5. Yu J, Chen Y, Fu X, Zhou X, Peng Y, Shi L, et al. Invasive *Fusobacterium nucleatum* may play a role in the carcinogenesis of proximal colon cancer through the serrated neoplasia pathway. Int J Cancer. 2016; 139: 1318-26.
6. Liang L, Kong C, Li J, Liu G, Wei J, Wang G, et al. Distinct microbes, metabolites, and the host genome define the multi-omics profiles in right-sided and left-sided colon cancer. Microbiome. 2024; 12: 274.
7. Dejea CM, Wick EC, Hechenbleikner EM, White JR, Mark Welch JL, Rossetti BJ, et al. Microbiota organization is a distinct feature of proximal colorectal cancers. Proc Natl Acad Sci U S A. 2014; 111: 18321-6.
8. Mima K, Cao Y, Chan AT, Qian ZR, Nowak JA, Masugi Y, et al. *Fusobacterium nucleatum* in colorectal carcinoma tissue according to tumor location. Clin Transl Gastroenterol. 2016; 7: e200.
9. Mima K, Nishihara R, Qian ZR, Cao Y, Sukawa Y, Nowak JA, et al. *Fusobacterium nucleatum* in colorectal carcinoma tissue and patient prognosis. Gut. 2016; 65: 1973-80.
10. Younginger BS, Mayba O, Reeder J, Nagarkar DR, Modrusan Z, Albert ML, et al. Enrichment of oral-derived bacteria in inflamed colorectal tumors and distinct associations of *Fusobacterium* in the mesenchymal subtype. Cell Rep Med. 2023; 4: 100920.
11. Ito M, Kanno S, Nosho K, Sukawa Y, Mitsuhashi K, Kurihara H, et al. Association of *Fusobacterium nucleatum* with clinical and molecular features in colorectal serrated pathway. Int J Cancer. 2015; 137: 1258-68.
12. Miyake T, Mori H, Yasukawa D, Hexun Z, Maehira H, Ueki T, et al. The comparison of fecal microbiota in left-side and right-side human colorectal cancer. Eur Surg Res. 2021; 62: 248-54.
13. Kneis B, Wirtz S, Weber K, Denz A, Gittler M, Geppert C, et al. Colon cancer microbiome landscaping: differences in right- and left-sided colon cancer and a tumor microbiome-ileal microbiome association. Int J Mol Sci. 2023; 24: 3265.
14. Zhong M, Xiong Y, Ye Z, Zhao J, Zhong L, Liu Y, et al. Microbial community profiling distinguishes left-sided and right-sided colon cancer. Front Cell Infect Microbiol. 2020; 10: 498502.
15. Rubinstein MR, Baik JE, Lagana SM, Han RP, Raab WJ, Sahoo D, et al. *Fusobacterium nucleatum* promotes colorectal cancer by inducing Wnt/β-catenin modulator Annexin A1. EMBO Rep. 2019; 20: e47638.
16. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. Cell Host Microbe. 2013; 14: 195-206.
17. Lin Y, Lau HC, Liu C, Ding X, Sun Y, Rong J, et al. Multi-cohort analysis reveals colorectal cancer tumor location-associated fecal microbiota and their clinical impact. Cell Host Microbe. 2025; 33: 589-601.e3.
18. Piccinno G, Thompson KN, Manghi P, Ghazi AR, Thomas AM, Blanco-Miguez A, et al. Pooled analysis of 3,741 stool metagenomes from 18 cohorts for cross-stage and strain-level reproducible microbial biomarkers of colorectal cancer. Nat Med. 2025; 31: 2416-29.
19. Xiang H, Shen B, Lao W, Bai B, Chen M, Zhou L, et al. The landscape of tissue-resident microbiota across normal, polyp, and colorectal cancer tissues. Nat Commun. 2026; 17: 3099.
20. Roelands J, Kuppen PJK, Ahmed EI, Mall R, Masoodi T, Singh P, et al. An integrated tumor, immune and microbiome atlas of colon cancer. Nat Med. 2023; 29: 1273-86.
21. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019; 37: 852-7.
22. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: high-resolution sample inference from illumina amplicon data. Nat Methods. 2016; 13: 581-3.
23. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. 2006; 72: 5069-72.
24. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One. 2013; 8: e61217.
25. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome. 2018; 6: 226.
26. Weyrich LS, Farrer AG, Eisenhofer R, Arriola LA, Young J, Selway CA, et al. Laboratory contamination over time during low-biomass sample analysis. Mol Ecol Resour. 2019; 19: 982-96.
27. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara R, et al. Vegan: community ecology package, version 2.0-10. J Stat Softw. 2013; 48: 103-32.
28. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2019; 35: 526-8.
29. Kaufman L, Rousseeuw PJ. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley & Sons Ltd; New York; 1990.
30. Holmes I, Harris K, Quince C. Dirichlet multinomial mixtures: generative models for microbial metagenomics. PLoS One. 2012; 7: e30126.
31. Brock G, Pihur V, Datta S, Datta S. clValid: an R package for cluster validation. J Stat Softw. 2008; 25: 1-22.
32. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. Nat Commun. 2020; 11: 3514.
33. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. J Stat Softw. 2010; 36: 1-13.
34. Watts SC, Ritchie SC, Inouye M, Holt KE. FastSpar: rapid and scalable correlation estimation for compositional data. Bioinformatics. 2019; 35: 1064-6.
35. Csárdi G, Nepusz T. The Igraph Software package for complex network research. InterJournal, Complex Syst. 2006; 1695: 1-9.
36. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models

of biomolecular interaction networks. Genome Res. 2003; 13: 2498-504.
37. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, et al. PICRUST2 for prediction of metagenome functions. Nat Biotechnol. 2020; 38: 685-8.
38. Escapa IF, Chen T, Huang Y, Gajare P, Dewhirst FE, Lemon KP. New insights into human nostril microbiome from the expanded Human Oral Microbiome Database (eHOMD): a resource for the microbiome of the human aerodigestive tract. mSystems. 2018; 3: e00187-18.
39. Nakatsu G, Li X, Zhou H, Sheng J, Wong SH, Wu WK, et al. Gut mucosal microbiome across stages of colorectal carcinogenesis. Nat Commun. 2015; 6: 8727.
40. Gazdar AF, Minna JD. Multifocal lung cancers—clonality vs field cancerization and does it matter? J Natl Cancer Inst. 2009; 101: 541-3.
41. Slaughter DP, Southwick HW, Smejkal W. "Field cancerization" in oral stratified squamous epithelium. Clinical implications of multicentric origin. Cancer. 1953; 6: 963-8.
42. Chen B, Scurrah CR, McKinley ET, Simmons AJ, Ramirez-Solano MA, Zhu X, et al. Differential pre-malignant programs and microenvironment chart distinct paths to malignancy in human colorectal polyps. Cell. 2021; 184: 6262-80.e26.
43. Zepeda-Rivera M, Minot SS, Bouzek H, Wu H, Blanco-Miguez A, Manghi P, et al. A distinct *fusobacterium nucleatum* clade dominates the colorectal cancer niche. Nature. 2024; 628: 424-32.
44. Narunsky-Haziza L, Sepich-Poore GD, Livyatan I, Asraf O, Martino C, Nejman D, et al. Pan-cancer analyses reveal cancer-type-specific fungal ecologies and bacteriome interactions. Cell. 2022; 185: 3789-806.e17.

Cite this article as: Wong SH, Kwong TN, Zhang R, Lam TY, Ho AM, Leung WW, et al. Comprehensive microbiome profiling reveals mucosal microbiome heterogeneity in patients with left- and right-sided colorectal neoplasia. Cancer Biol Med. 2026; 23: 755-768. doi: 10.20892/j.issn.2095-3941.2025.0846