

# Multi-omics analysis reveals associations between gut microbiota and host transcriptome in colon cancer patients

Yuling Qin,<sup>1</sup> Qiang Wang,<sup>2</sup> Qiumei Lin,<sup>1</sup> Fengfei Liu,<sup>1</sup> Xiaolan Pan,<sup>1</sup> Caibiao Wei,<sup>1</sup> Junxian Chen,<sup>1</sup> Taijun Huang,<sup>1</sup> Min Fang,<sup>2</sup> Weilong Yang,<sup>3,4</sup> Linghui Pan<sup>2</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 13.

**ABSTRACT** Colon cancer (CC) is one of the most common cancers globally, which is associated with the gut microbiota intimately. In current research, exploring the complex interaction between microbiomes and CC is a hotspot. However, the information on microbiomes in most previous studies is based on fecal, which does not fully display the microbial environment of CC. Herein, we collected mucosal and tissue samples from both the tumor and normal regions of 19 CC patients and clarified the composition of mucosal microbiota by 16S rRNA and metagenomic sequencing. Additionally, RNA-Seq was also conducted to identify the different expression genes between tumor and normal tissue samples. We revealed significantly different microbial community structures and expression profiles to CC. Depending on correlation analysis, we demonstrated that 1,472 genes were significantly correlated with CC tumor microbiota. Our study reveals a significant enrichment of *Campylobacter jejuni* in the mucosa of CC, which correlates with bile secretion. Additionally, we observe a negative correlation between *C. jejuni* and immune cells CD4+ Tem and mast cells. Finally, we discovered that metabolic bacterial endosymbiont of *Bathymodiolus* sp., *Bacillus wiedmannii*, and *Mycobacterium tuberculosis* had a significant survival value for CC, which was ignored by previous research. Overall, our study expands the understanding of the complex interplay between microbiota and CC and provides new targets for the treatment of CC.

**IMPORTANCE** This study contributes to our understanding of the interaction between microbiota and colon cancer (CC). By examining mucosal and tissue samples rather than solely relying on fecal samples, we have uncovered previously unknown aspects of CC-associated microbiota. Our findings reveal distinct microbial community structures and gene expression profiles correlated with CC progression. Notably, the enrichment of *Campylobacter jejuni* in CC mucosa, linked to bile secretion, underscores potential mechanisms in CC pathogenesis. Additionally, observed correlations between microbial taxa and immune cell populations offer new avenues for immunotherapy research in CC. Importantly, this study introduces CC-associated microbiota with survival implications for CC, expanding therapeutic targets beyond conventional strategies. By elucidating these correlations, our study not only contributes to uncovering the potential role of gut microbiota in colon cancer but also establishes a foundation for mechanistic studies of gut microbiota in colon cancer, emphasizing the broader impact of microbiota research on cancer biology.

**KEYWORDS** multi-omics, colon cancer, mucosal microbiota, correlation, bile secretion, immune, survival value

Colon cancer (CC) is the third most common cancer worldwide, causing nearly 1 million deaths annually (1). While advancements in diagnosing and treating CC have been achieved, leading to a reduction in mortality rates, the 5 years survival rate

**Editor** Sergio Baranzini, University of California, San Francisco, San Francisco, California, USA

Address correspondence to Min Fang, fangmin@sr.gxmu.edu.cn, Weilong Yang, yangweilong@whu.edu.cn, or Linghui Pan, linghuipangx@126.com.

Yuling Qin, Qiang Wang, and Qiumei Lin contributed equally to this article. Author order was determined by drawing straws.

The authors declare no conflict of interest.

See the funding table on p. 14.

**Received** 13 June 2024

**Accepted** 31 January 2025

**Published** 27 February 2025

Copyright © 2025 Qin et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

remains notably low (2). Environmental factors and hereditary factors both contribute to colon cancer. However, such as Lynch syndrome and familial adenomatous polyposis, the genetic susceptibility factors only account for a minority of CC cases (3, 4). Therefore, the impact of environmental factors on colon cancer is significant.

The role of microorganisms in the realm of cancer biology, within the influence of environmental factors, is progressively gaining recognition (5). Gut microbiota associates with colon cancer intimately, forming a unique tumor microbiota, suggesting the significant role of tumor-associated microbiota in CC (6). In general, CC patients frequently exhibit the characteristic of dysbiosis in the gut microbiome (7). The findings of the current study indicate a strong correlation between the occurrence and development of CC and certain bacterial species, such as *Fusobacterium*, *Porphyromonas*, and *Bifidobacterium*, among others (8–10). For instance, in colorectal cancer (CRC), *Fusobacterium nucleatum* promotes glycolysis and tumorigenesis by targeting lncRNA ENO1-IT1 (11). *Porphyromonas* promotes tumorigenesis utilizing the bacterial metabolite butyrate secretion (12). The activation of the hematopoietic NLRP3 inflammasome by *Porphyromonas gingivalis* is implicated in the promotion of colorectal cancer (13). Enterotoxigenic *Bifidobacterium fragilis* toxins promote tumor occurrence by activating STAT3 and Wnt pathways and stimulating IL-17 production (14).

In previous studies, the information on gut microbiota in CC was mostly based on fecal matter (15, 16). However, fecal samples can only partially reflect the composition of the intestinal microbiota and there are significant differences between fecal microbiota and intestinal microbiota (17, 18). The research by Li et al. suggests that certain genera, including *Fusobacterium* and *Gemella*, exhibit a greater propensity for adherence to tumor mucosa as opposed to being excreted in significant quantities within fecal matter (19). The evolution of intestinal mucosal microbiota is closely linked to the host and forms a specific local community. Intestinal mucosal microbiota is important for host health and disease development and can accurately describe the microbiota associated with colorectal cancer (20). Additionally, a complex interplay exists between the gut microbiota and host gene expression (21–23). The synergistic investigation of host transcriptomics and microbiome explains host-microbiota interaction at the gene expression level and shows gut microbiota-differential gene interaction mechanism. Therefore, in this study, we analyzed the different microbial community structures of CC based on mucosa using 16S rRNA and metagenomic sequencing. Subsequently, we identified differentially expressed genes (DEGs) and demonstrated their correlation with tumor-associated microbiota. Finally, we evaluated the survival value of tumor-associated microbiota in colon cancer. Overall, this comprehensive multi-omics approach helps us comprehend the complicated link between gut microbiota, host gene expression, and colon cancer, ultimately leading to potential insights for future therapeutic interventions.

## MATERIALS AND METHODS

### Study subjects and sample collection

The human specimens involved in this study were obtained from clinical tissue samples of patients who had radical resection of colon cancer tumors at the Cancer Hospital Affiliated to Guangxi Medical University from January to May 2022. The criteria for patient inclusion were as follows: (i) confirmed by colonoscopy pathology, the patient is diagnosed with colon cancer for the first time; (ii) the patient should be at least 18 years old. The criteria for exclusion are as follows: (i) recurrent colon cancer patients; (ii) having undergone any form of chemotherapy and radiation therapy before surgery; (iii) cancer concurrent with other systems. All patients included in this study have not undergone any antibiotic therapy within the past month. Paired normal and tumor tissues and mucosa were collected during open surgery, and normal tissues or mucosa were obtained at a distance of at least 10 cm from the tumor tissue. Table S1 details 38 samples from 19 patients in this study.

Postoperative fresh surgical specimens were washed with sterile saline solution to gently remove fecal material adhering to the mucosal surface. Sterile scalpels and cotton swabs were used to scrape the mucosal surface on the tumor site and at least 10 cm away from the tumor site. The samples were stored in  $-80^{\circ}\text{C}$  frozen tubes for 16S rRNA and metagenomic sequencing. Tissue samples from the same site were collected using sterile tissue scissors and stored in  $-80^{\circ}\text{C}$  frozen tubes for transcriptomic sequencing. After collection, the specimens were transported on dry ice to Novogene Co., Ltd., Beijing, for processing and sequencing.

### 16S rRNA sequencing and analysis

The bacterial DNA was isolated from samples of colon contents using a Magnetic Soil and Stool DNA Kit (Tiangen Biotech, China) following the instructions provided by the manufacturer. The concentration and purity of DNA were assessed using a 1% agarose gel. 16S rRNA/18S rRNA/ITS genes in distinct regions (16S V4/16S V3/16S V3-V4/16S V4-V5, 18S V4/18S V9, ITS1/ITS2, Arc V4) were amplified with specific primers (e.g., 16S V4: 515F-806R, 18S V4: 528F-706R, 18S V9: 1380F-1510R, etc.) and barcodes. The PCR-amplified products were detected and purified by 2% agarose gel electrophoresis with a Qiagen Gel Extraction Kit (Qiagen, Germany). Sequencing libraries were constructed using TruSeq DNA PCR-Free Sample Preparation Kit (Illumina, USA) according to the instructions provided by the manufacturer. The library underwent sequencing using an Illumina NovaSeq platform, resulting in the generation of 250 bp paired-end reads. The primer sequences used for sequencing were as follows: GTGCCAGCMGCCGCGGTAA, GGACTACHVGGGTWTCTAAT (24, 25).

High-quality clean data were obtained by subjecting the raw data to quality filtering using particular filtering conditions, following the QIIME (26, 27) (V1.9.1, [http://qiime.org/scripts/split\\_libraries\\_fastq.html](http://qiime.org/scripts/split_libraries_fastq.html)) quality-controlled process. Based on the clean data, the process of clustering operational taxonomic units (OTUs) and classifying species was evaluated using Uparse (28) (V7.0.1001, <http://www.drive5.com/uparse/>). The Silva Database (29) (<http://www.arb-silva.de/>) was utilized, employing the Mothur algorithm, to annotate taxonomic information for each representative sequence. The abundance information of OTUs was normalized using a standard based on the sequence number corresponding to the sample with the least sequences.

Alpha diversity indices including observed species, Chao1, Shannon, Simpson, ACE, and Good's coverage were calculated with QIIME (V1.7.0) and displayed with R software (V2.15.3). Beta diversity among different groups was detected based on the unweighted UniFrac index using QIIME (V1.9.1) software.

### Metagenomic sequencing and analysis

The extraction of bacterial DNA from samples of colon contents was performed using a Magnetic Soil and Stool DNA Kit (Tiangen Biotech, China) in accordance with the manufacturer's instructions. The genomic DNA underwent random shearing, resulting in the formation of short fragments. The obtained fragments underwent end-repair, A-tailing, and subsequent ligation with an Illumina adapter. The adapter-containing fragments were PCR amplified, size chosen, and purified. Qubit and real-time PCR were used to quantify, and a bioanalyzer was used to detect size distribution in the library. The quantified libraries were pooled and subjected to sequencing using Illumina PE150 platforms, based on the optimal library concentration and data amount required.

Raw data from the Illumina sequencing platform were preprocessed with Readfq (V8, <https://github.com/cjfields/readfq>) to create clean data for analysis. The assembly analysis of clean data was conducted using the MEGAHIT software (V1.0.4-beta). Starting from the assembled scaffolds of individual samples, gene prediction was performed using MetaGeneMark (V3.05, <http://topaz.gatech.edu/GeneMark/>). The predicted genes from each sample were then compiled, and redundancies were removed to construct a gene catalog (Unigenes). DIAMOND (30) software (V0.9.9, <https://github.com/bbuch-fink/diamond/>) was used to align Unigenes with those in the functional database.

Functional databases include Kyoto Encyclopedia of Genes and Genomes (31, 32) (KEGG, V2018-01-01, <http://www.kegg.jp/kegg/>), Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups (33) (eggNOG, V4.5, (<http://eggnogdb.embl.de/#/app/home>), and Carbohydrate-active Enzymes Database (34) (CAZy, V201801, <http://www.cazy.org/>). Based on Bray-Curtis distance, non-metric multi-dimensional scaling (NMDS) analysis (35) was conducted using the vegan, permute, lattice, ade4, ggplot2, and grid packages within R (V1.9.1). Analysis of similarities (ANOSIM) was used to test the differences between groups (R vegan package, V2.15.3). The rarefaction curves, species accumulation boxplot, and principal coordinates analysis (PCoA) were generated using the FactoMineR, WGCNA, stats, and ggplot2 packages in R software (V3.2.1). The Linear discriminant analysis Effect Size software (LEfSe, V1.0) (36) was utilized to perform LEfSe analysis, with a Linear Discriminant Analysis (LDA) score threshold of 3, in order to identify the biomarkers.

### RNA sequencing and analysis

Colonic contents were subjected to RNA extraction using the Trizol method (Ambion, USA). The quantification and quality assessment of RNA samples was performed using the RNA Nano 6000 Assay Kit on the Bioanalyzer 2100 instrument (Agilent Technologies, CA, USA). The RNA sample preparations utilized total RNA as the input material. To isolate cDNA fragments with a desired length range of 370–420 bp, the library fragments underwent purification using the AMPure XP system (Beckman Coulter, Beverly, USA). Following the PCR amplification, the resulting PCR product underwent purification using AMPure XP beads, resulting in the library. After qualification, the libraries are pooled by effective concentration and the target amount of data off the machine, then sequenced by Illumina NovaSeq 6000. The end reading of 150 bp pairing is generated.

The sequencing data underwent quality control and filtration processes. Hisat2 (37) (V2.0.5) was used to create the reference genome index and match paired-end clean reads to it. FeatureCounts (38) (V1.5.0-p3) was employed for quantifying the number of mapped reads for each gene. Subsequently, each gene's FPKM (Fragments Per Kilobase of transcript per Million mapped fragments, FPKM) value was calculated from its length and read count. To analyze differential expression between two groups, the DESeq2 (39) R package was utilized.  $\text{Padj} \leq 0.05$  and  $|\log_2(\text{foldchange})| \geq 1$  were set as the threshold for significantly differential expression. Clusterprofiler (40) (V3.8.1) software was used to analyze the statistical enrichment of differentially expressed genes in the KEGG pathway. Based on the algorithm provided in R package xCell (41) (V4.2.1), the immune infiltration of each sample is calculated by using the markers of 64 kinds of immune cells provided internally.

### Correlation analysis

Based on the differentially expressed genes, immune infiltration score, and the absolute abundance of differential microbiota, the psych package in R (V4.4.0) software was used for Spearman correlation analysis and significance test. According to the correlation coefficient, corrplot package was used to draw the correlation heat map. The conditions for screening-related genes were  $P < 0.05$ ,  $r \geq 0.5$ , or  $r \leq -0.5$ .

### Survival analysis

Firstly, we used GEPIA2 (42) (<http://gepia2.cancer-pku.cn/#index>) to build the survivorship curve of DEGs in the passageways of bile secretion. To do a batch survival analysis of all DEGs, the phenotypic information and gene expression matrix of GDC TCGA Colon Cancer (COAD) were downloaded from the TCGA database. We took the intersection of the DEGs and the gene expression matrix from the TCGA database, and found that only 1,365 genes overlapped. Survival R package (43) was used to do survival analysis, and  $P < 0.05$  was set as the threshold for a significant survival rate.

## RESULTS

### Disordered mucosal microbiota structures of CC

To investigate the influence of microbial microenvironment on CC, a workflow was designed (Fig. 1A). Firstly, 21 pairs of tumor and normal mucosal samples were obtained to do 16S rRNA sequencing (Table S2). After quality-control filtering, 19 pairs of sequence data were remained. The species accumulation boxplot and rarefaction curve show that the sample size of our study is sufficient, and the amount of sequencing data is reasonable (Fig. S1). Clustering yielded 5,786 OTUs, 2,281 of which were unique to normal and 748 to tumor groups (Fig. 1B). According to the taxonomic annotations, at the phylum level, the two microbial communities are mainly composed of the *Proteobacteria* (N: 31.58%; T: 31.45%), *Bacteroidota* (N: 31.15%; T: 26.44%), *Firmicutes* (N: 23.70%; T: 19.69%), and *Fusobacteriota* (N: 5.83%; T: 14.68%) (Fig. 1C; Fig. S2).

Based on the analysis of alpha diversity using the observed species index, there was a significant difference in species richness between the normal group and the tumor group (Fig. 1D,  $P < 0.05$ ; Table S3). In addition, this study compared the diversity of gut bacteria using the  $\beta$ -diversity index and found significant differences in the composition of gut microbiota between the two groups (Fig. 1E,  $P < 0.05$ ).

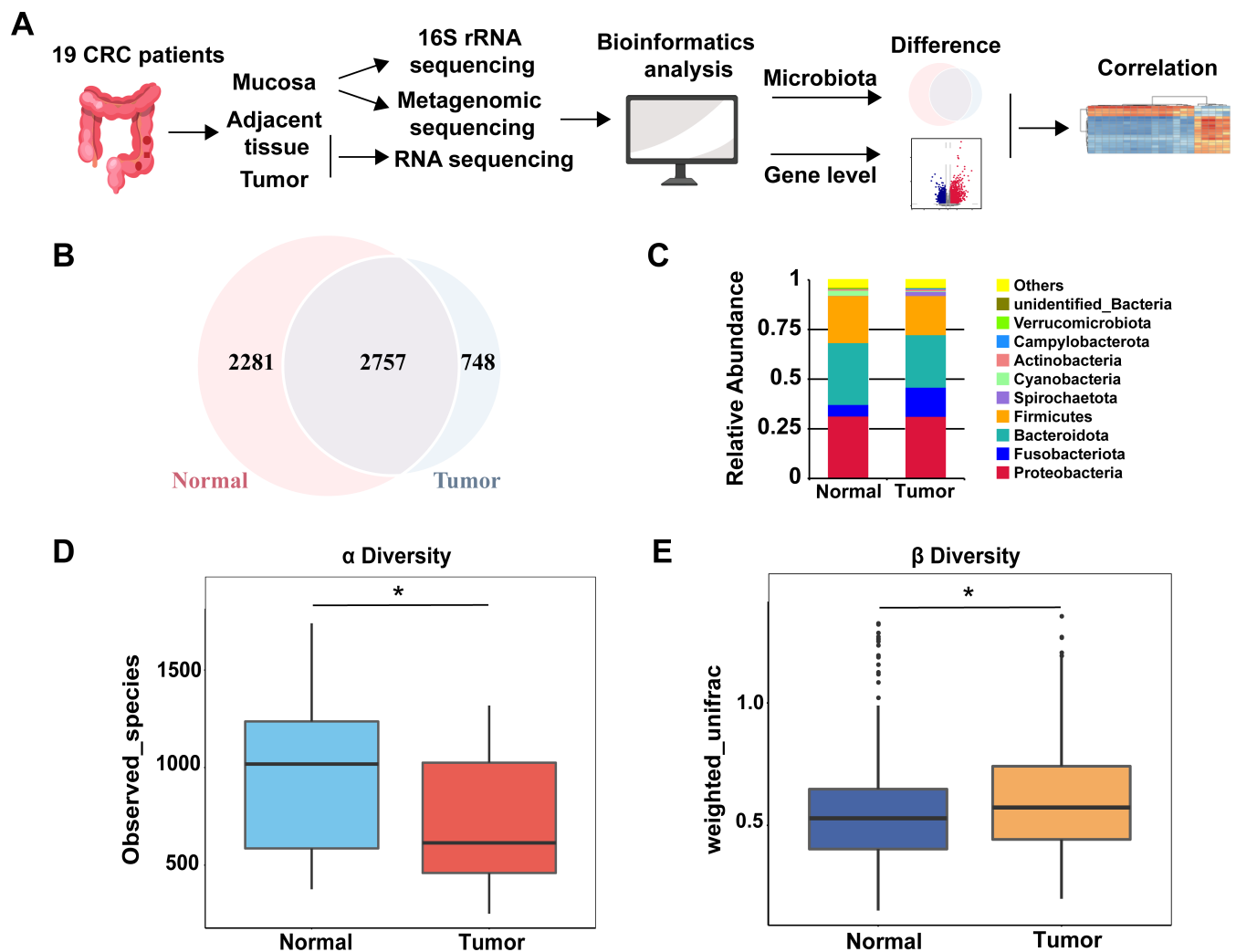
### Annotation of tumor-associated microbiota and functions

To further elucidate the functional composition of colon cancer mucosal microbiota, we conducted metagenomic sequencing and functional annotation on 19 pairs of tumor and normal mucosal samples (Table S4). The core-pan gene dilution curve and the heatmap of gene abundance correlation among samples demonstrated the reliability of the experiment and the rationality of sample selection (Fig. S3). All functional annotation results, based on KEGG, eggNOG, and CAZy databases, showed that the functional abundance composition was different between the normal and tumor groups (Fig. 2A through C; Fig. S4). PCoA and ANOSIM illustrated the obvious separation of microbial functional composition between normal and tumor groups (Fig. S5), confirming that there were significant functional abundance differences between normal and tumor groups. Based on the CAZy database, we also explored the abundance differences of different enzymes (Tables S5 and S6). NMDS and ANOSIM based on species abundance revealed that there were significant differences in species composition between normal and tumor groups at the phylum level (Fig. 2D and E).

Next, in order to screen for species and functional biomarkers with significant differences between groups, we conducted LEfSe analysis. It is worth noting that we identified 20 differentially abundant species associated with the disease state, and we defined them as tumor-associated microbiota (Fig. 2F). The LEfSe analysis based on species abundance shows that species such as *Bacteroides vulgatus*, *Bifidobacterium pseudocatenulatum*, *Bifidobacterium breve*, *Bacteroides uniformis*, and *Bacteroides dorei* are enriched in the normal group. Species such as *Salmonella enterica*, *Klebsiella aerogenes*, *Bacillus cereus*, *Bacillus pseudomycoides*, and *Staphylococcus aureus* are enriched in the tumor group. Based on the KEGG databases, we found that starch and sucrose metabolism (ko00500), amino sugar and nucleotide sugar metabolism (ko00520), and galactose metabolism (ko00052) were significantly more abundant in the normal group than in the tumor group (Fig. 2G). However, the top three upregulated pathways in the tumor group were pathways in cancer (ko05200), endocytosis (ko04144), and ubiquitin-mediated proteolysis (ko04120).

### Altered gene expression profile in CC

In order to uncover the different gene expressions of CC, 19 pairs of tumor and adjacent non-cancerous tissues (normal) were obtained to do RNA-Seq (Table S7). After stringent data quality control and filtering, one pair of samples was excluded. Transcriptome analysis showed that there were 4,693 DEGs identified in the tumor group, compared to the normal group ( $|\log_2(\text{foldchange})| \geq 1$  and  $\text{Padj} \leq 0.05$ ) (Fig. 3A), including 2,169



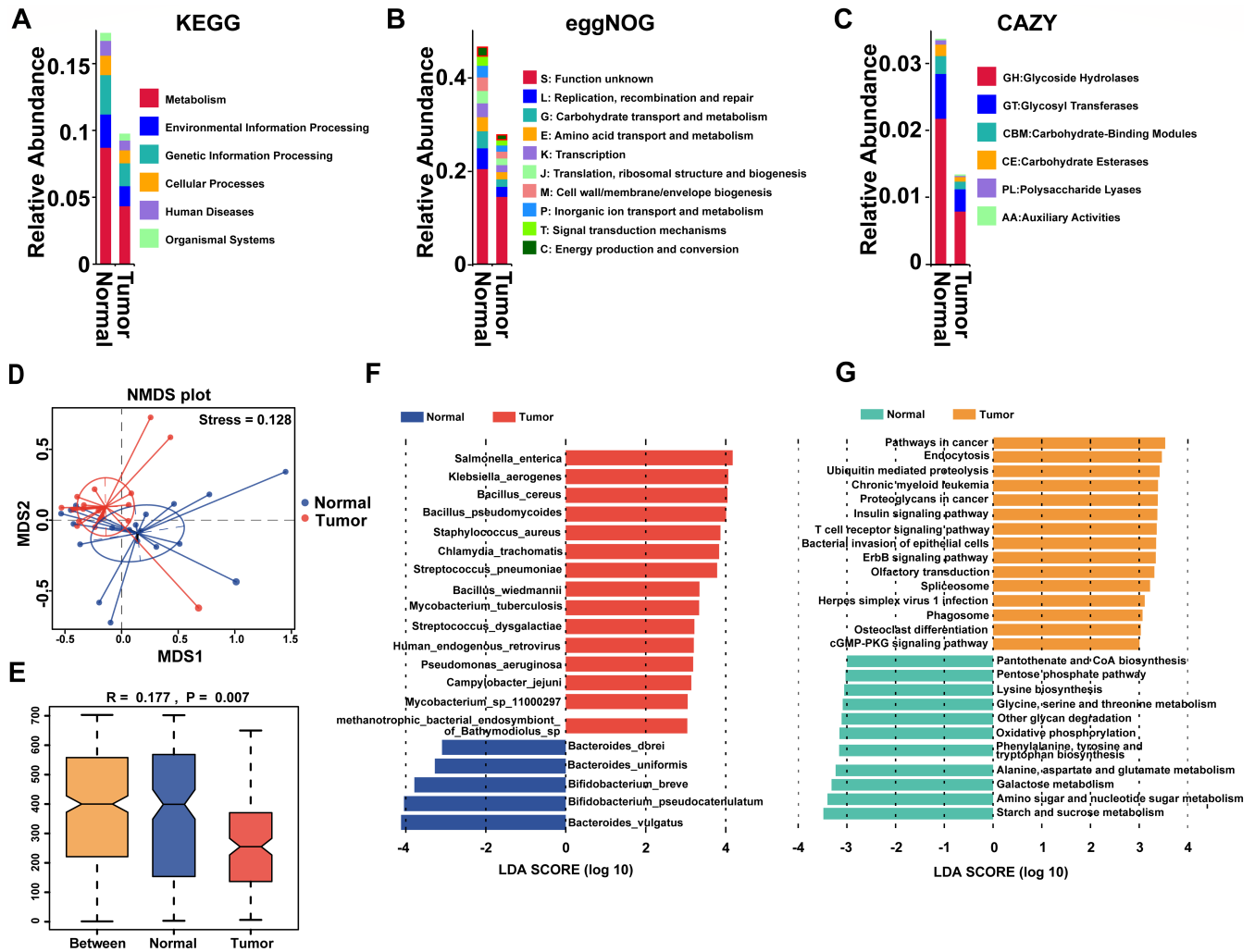
**FIG 1** Disordered mucosal microbiota of CC based on 16S rRNA sequencing. (A) The schematic of the overall approach. (B) Venn diagram for the number of unique and common OTUs in normal and tumor groups. (C) The mucosal microbiota composition of normal and tumor groups at the phylum level. (D) Alpha diversity index between groups difference analysis. (E) Beta diversity index between groups difference analysis. \* $P < 0.05$ .

up-expression genes and 2,524 down-expression genes. Next, we also do KEGG pathway clustering analysis to up-expression genes and down-expression genes. We found that up-expression genes were mainly enriched in cytokine-cytokine receptor interaction, cell cycle, IL-17 signaling pathway, etc., and down-expression genes were mainly enriched in neuroactive ligand-receptor interaction, calcium signaling pathway, bile secretion, etc. (Fig. 3B and C). In addition, differential analysis based on immune infiltration showed differences in 30 types of immune cells between the normal group and the tumor group (Fig. 3D).

### Correlation between tumor-associated microbiota and DEGs in CC

The tumor-associated microbiota is strongly related to CC; however, its relationship to host gene expression is still unclear (44). Therefore, we determined the correlation between the tumor-associated microbiota and the 3,632 DEGs ( $\text{Padj} < 0.01$ ) identified previously. Among them, 1,472 DEGs were significantly correlated to the tumor-associated microbiota (20 species,  $\text{LDA} > 3$ ) (Table S8). Notably, the number of correlated DEGs in each tumor-associated microbiota genus was different, ranging from 86 to 872 (Fig. S6). KEGG pathway clustering analysis was performed on these 1,472 correlated DEGs, indicating that they were significantly enriched in eight functional pathways (bile





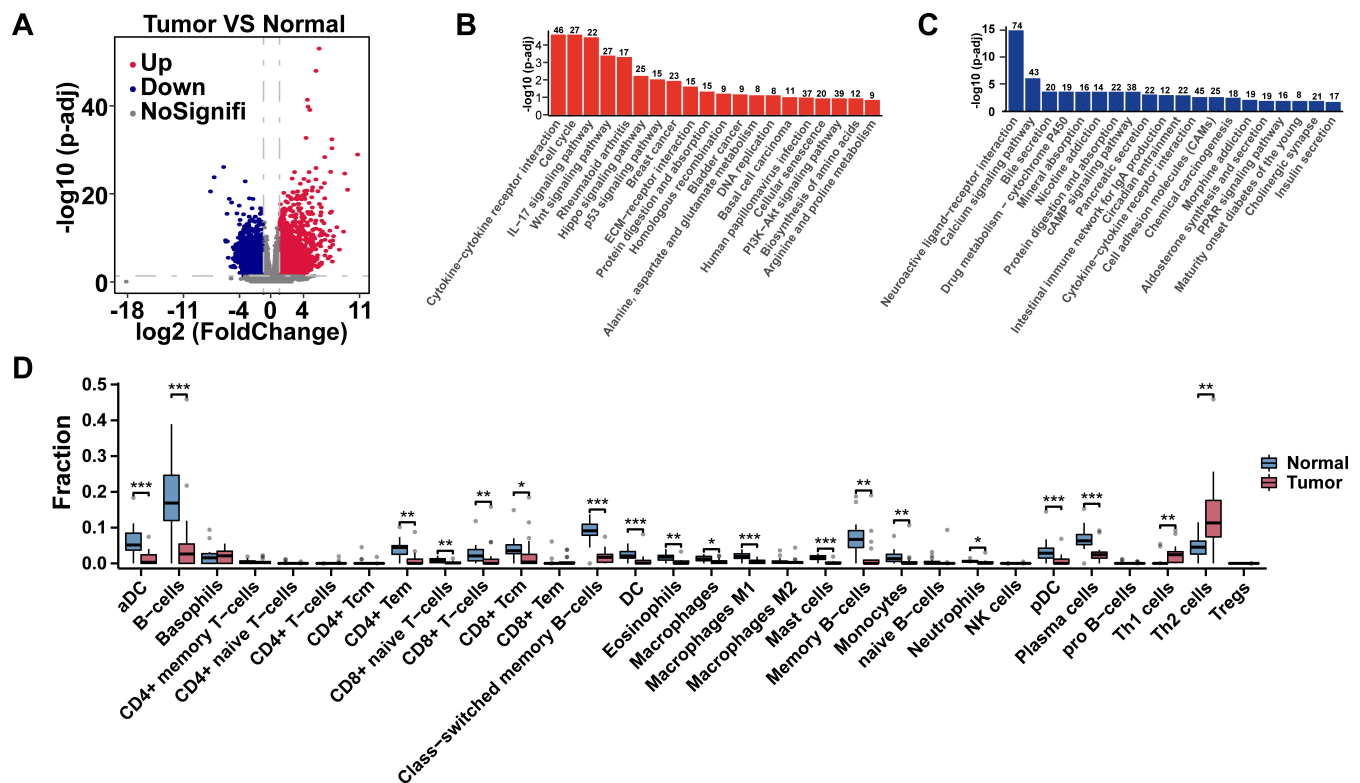
**FIG 2** Function analysis of mucosal microbiota based on metagenomic sequencing. (A–C) At the first classification level, the functional composition of mucosal microbiota in normal and tumor groups at the KEGG (A), eggNOG (B), and CAZy (C) databases. (D) NMDS analysis of normal and tumor groups based on the species of phylum level. (E) ANOSIM of normal and tumor groups based on the species of phylum level. (F) Distribution histogram of LDA values of different species (LDA > 3). (G) Distribution histogram of LDA based on functional abundance (LDA > 3).

secretion, neuroactive ligand-receptor interaction, cytokine-cytokine receptor interaction, arginine and proline metabolism, protein digestion and absorption, PPAR signaling pathway, nicotine addiction, Nitrogen metabolism) (Fig. 4A).

Previous studies showed that bile secretion had a relationship with CC (45). Thus, we picked up all 18 DEGs related to bile secretion and highlighted their correlation to 20 tumor-associated microbiota (Fig. 4B). It demonstrated that *Campylobacter jejuni* and methanotrophic bacterial endosymbiont of *Bathymodiolus* sp. had strong correlations to bile secretion (Fig. S7). In addition, based on species abundance and transcriptome immune infiltration results, we also explored the correlation between different species and host immunity (Fig. 4C). It is worth mentioning that there is a negative correlation between *Campylobacter jejuni* and CD4+ Tem and master cells (Fig. 4D and E).

### Tumor-associated microbiota has a significant survival value for CC

Although 1,472 DEGs were found significantly correlated to the tumor-associated microbiota above, most of these DEGs did not influence the survival of CC patients. Therefore, we did the survival analysis of all DEGs in CC based on the TCGA database, and 100 DEGs were identified as related to the survival of CC patients ( $P < 0.05$ ) (Table S9).



**FIG 3** Different transcriptome profiles between normal and tumor groups. (A) The DEGs between normal and tumor groups. (B) The KEGG enrichment analysis of up-expression genes in tumor groups. (C) The KEGG enrichment analysis of down-expression genes in tumor groups. (D) Analysis of immune infiltration between the normal group and the tumor group. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

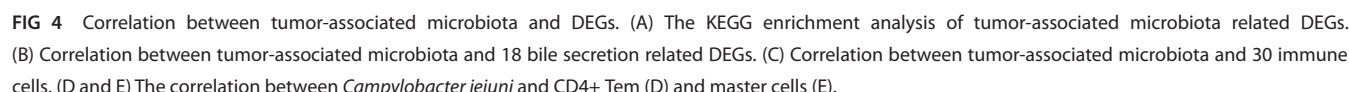
Next, we calculated the number of correlated DEGs related to the survival of CC patients in each tumor-associated microbiota. The number of each species had an obvious difference (Fig. 5A; Table S10). Among them, compared to the other tumor-associated microbiota species, methanotrophic bacterial endosymbiont of *Bathymodiolus* sp., *Bacillus wiedmannii*, and *Mycobacterium tuberculosis* have more DEGs related to the survival of CC patients. This means that they have significant survival value for CC. Importantly, we also performed the survival analysis of all 18 DEGs related to bile secretion in CC, revealing that AQP8 and SLC4A4 were related to the survival of CC patients (Fig. 5B and C).

## DISCUSSION

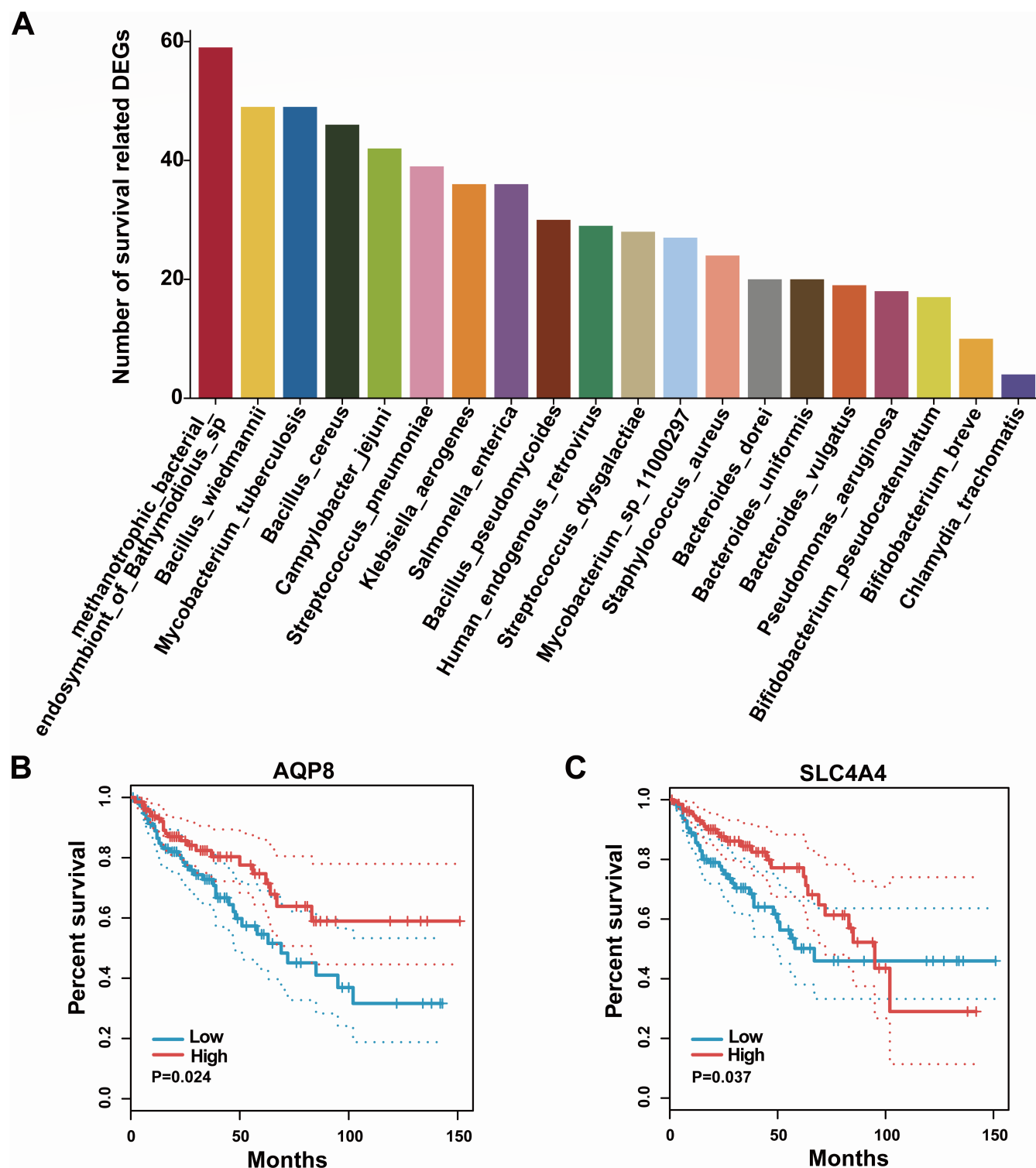
### Portray tumor-associated microbiota accurately

The tumor-associated microbial community is closely linked to the occurrence and development of tumors, and the presence of these microorganisms has an impact on tumor progression and treatment response (46, 47). Fecal microbes are used to do most of the related research; however, many studies have shown that there are significant differences in the bacterial composition between fecal and intestinal mucosal samples, which are directly connected to the tumor, allowing for a more accurate reflection of the composition of tumor-associated microbiota (48). Thus, to research the tumor-associated microbiota and their functional composition on the mucosal surface, we sequenced 16S rRNA amplicon and metagenomics on the mucosal tissues of CC patients. This study confirmed significant microbial community differences between tumor sites and paired normal mucosa (Fig. 1). In line with other studies, the predominant phyla in the gut microbiota are *Firmicutes*, *Bacteroidetes*, *Fusobacteria*, and *Proteobacteria* (49, 50).





10.1128/msystems.00805-24 9



**FIG 5** Survival analysis of DEGs and tumor-associated microbiota. (A) The number of DEGs which were both associated with survival and correlated to the tumor-associated microbiota at the species level. (B and C) Survival analysis of AQP8 (B) and SLC4A4 (C).

to the occurrence and progression of colon cancer (56). However, the association between *B. cereus*, *S. aureus*, and colon cancer has not been thoroughly studied and confirmed. *M. tuberculosis* is the causative agent of pulmonary tuberculosis (57). Interestingly, Kim et al. discovered a negative relationship between CRC and the incidence of tuberculosis (58). They found that heat-inactivated *M. tuberculosis* could

improve inflammation-induced colorectal cancer by altering the composition of intestinal microorganisms. Unlike previous findings, our study reveals that *M. tuberculosis* is enriched in tumor mucosa. Various factors such as sample sources, research methodologies, and experimental conditions could account for this disparity. The abnormal transcriptional activation of *Human endogenous retrovirus (HERV)* is associated with the occurrence and progression of cancer (59). Ko et al. found that the expression of *HERV-K env* protein was significantly increased in CRC, and knocking out the *HERV-K env* gene significantly reduced the tumorigenic characteristics of colorectal cancer cells, including proliferation, invasion, migration, and tumor colonization (60). At present, there is insufficient research evidence to suggest a direct association between *B. pseudomycoides*, *C. trachomatis*, *S. pneumoniae*, *B. wiedmannii*, *S. dysgalactiae*, *P. aeruginosa*, *Mycobacterium* sp. 11000297, and metamorphotropic bacterial endosymbiont of *Bathymodiolus* sp. with colon cancer. Pathogenic bacteria will be an important target for treating CC, and further research is needed to establish any potential relationship between these tumor-associated microbiota and colon cancer. *Bacteroides vulgatus*, *Bifidobacterium pseudocatenulatum*, *Bifidobacterium breve*, *Bacteroides uniformis*, and *Bacteroides dorei* are enriched in normal mucosa. *B. vulgatus* is one of the predominant species of the *Bacteroides* genus in the human intestinal tract. Research has shown that *B. vulgatus* exhibits tumor-suppressive effects *in vitro*. *B. vulgatus* is one of the predominant species of the *Bacteroides* genus in the human intestinal tract and is generally considered a beneficial symbiotic bacterium. Research has shown that *B. vulgatus* exhibits tumor-suppressive effects *in vitro*, and it can alleviate colitis in mice by modulating the gut microbiota and immune response (61, 62). *B. pseudocatenulatum* G7 also exhibits anti-colitis effects (63), but further research is needed to elucidate its association with CC. *B. breve* is also a potential probiotic, and its derived indole-3-lactic acid can improve colitis-associated tumorigenesis by guiding the differentiation of immature colonic macrophages (64). Currently, there is no direct evidence linking *B. uniformis* and *B. dorei* to colon cancer. However, their depletion of tumor mucosa suggests their significant role in intestinal ecological balance. Therefore, further research is needed to substantiate this claim.

To further explore the functional composition of mucosal microbiota, we performed functional annotation of the metagenome. We found that glycoside hydrolases accounted for the majority of differential enzymes, and the abundance of glycoside hydrolase families was significantly decreased in normal mucosa. Significantly, the tumor group exhibited downregulation of genes associated with starch and sucrose metabolism pathways. This suggests the potential role of tumor-associated microbial communities in the pathogenesis of cancer, which may involve starch and sucrose metabolism as well as glycoside hydrolases. Moreover, previous studies also implicated enzymes such as  $\beta$ -galactosidase and  $\beta$ -glucuronidase as key players in the impact of gut microbiota on colorectal cancer (65, 66). In a word, multi-omics analyses can provide a deeper understanding of the role of tumor-associated microbial communities in cancer pathogenesis, and the incorporation of metabolomics and proteomics will be a focus of our future mechanistic studies.

### Identification of the correlated tumor-associated microbiota to DEGs

Consistent with other studies (67, 68), our transcriptomic sequencing results indicate differential gene expression between tumor tissue and adjacent normal tissue. The top three microorganisms most significantly associated with DEGs were metamorphotropic bacterial endosymbiont of *Bathymodiolus* sp., *Bacillus wiedmannii*, and *Campylobacter jejuni* (Fig. S6). However, their role in colon cancer is unknown, and further experimental validation is required to determine whether they impact the host's transcriptome.

There is a close relationship between metabolic pathways and colon cancer. Abnormalities in metabolic pathways may lead to an imbalance in cell growth, proliferation, and apoptosis, thereby promoting cancer development (69). Increasing evidence suggests that the human gut microbiota exerts carcinogenic effects through its

metabolic repertoire, particularly in bile acid metabolism (70). In this study, we observed the enrichment of DEGs that related to differential microbial taxa in bile secretion pathways (Fig. 4). Moreover, we found a strong negative correlation between bacteria belonging to *C. jejuni* and genes associated with bile secretion. Therefore, we hypothesize that *C. jejuni* may promote the development of CC by influencing bile secretion. However, the current research on the relationship between *C. jejuni* and bile secretion is relatively limited. Further in-depth research is required to understand the specific mechanisms and impact of *Campylobacter jejuni* on disease development.

Inflammation mechanisms are key drivers in tumorigenesis, and some CC patients also suffer from inflammatory bowel disease (71). Increased pathogenic or pathobiont microbiota in the gut may lead to excessive activation of the immune system's inflammatory response, disrupting immune balance and thereby promoting CC development (72). In our study, we analyzed the correlation between immune infiltration and gut microbiota. We found a negative correlation between *C. jejuni* and CD4<sup>+</sup> Tem cells, as well as mast cells. Previous studies have indicated that the interaction between the host and the microbiota can influence various aspects of host immunity, including the development, differentiation, and function of immune cells (73, 74). Consistent with this notion, our research has revealed interactions between the microbiota and metabolic pathways. Specifically, we found a correlation between *C. jejuni* and the bile secretion pathway. Therefore, we hypothesize that *C. jejuni* may modulate the host's immune response by affecting bile metabolism, and their interaction promotes the onset of colon cancer tumors. However, prior studies have suggested that *C. jejuni* can promote the development of colorectal tumors through the action of its cytotoxic distending toxin (56). Our study undoubtedly presents a novel discovery regarding the pro-carcinogenic effect of *C. jejuni*, yet further research is needed to elucidate its underlying mechanism.

### Tumor-associated microbiota—potential therapeutic direction for prolonging CC survival

Microbiota-based therapy is gaining increasing attention in cancer treatment (75, 76). Microbiota-based therapy involves targeted reconstruction of the tumor microbiota or management of microbial metabolites to intervene in cancer, and it is becoming an important component of cancer treatment (6, 47). For example, the addition of physiologically relevant doses of aspirin inhibits the expression of virulence factors FadA/Fap2 in *Fusobacterium nucleatum*, thereby suppressing tumor formation in ApcMin/+ mice induced by *F. nucleatum* (77). Administration of *Bifidobacterium CGMCC 15068* effectively prevents tumor formation in a mouse model of colitis-associated colorectal cancer, and this intervention leads to a marked increase in the abundance of beneficial bacteria in the mouse intestine and induces notable modifications to the metabolic profile of the intestine (78). Our study found that the metabolic bacterial endosymbiont of *Bathymodiolus* sp., *Bacillus wiedmannii*, and *Mycobacterium tuberculosis* are enriched in the tumor mucosa, and they have significant survival value in CC (Fig. 5). However, there is currently no definitive research or evidence to suggest a direct link between these species and CC. Our research findings provide preliminary evidence demonstrating a close association between the metabolic bacterial endosymbiont of *Bathymodiolus* sp., *B. wiedmannii*, *M. tuberculosis*, and survival-related genes in CC. Further exploring this association can help deepen our understanding of their involvement in the development of colon cancer and provide new clues for the treatment and prognosis assessment of CC patients. Future studies could investigate the interactions between these species and survival-related genes in CC, as well as their potential roles in development and treatment response. However, it is important to note that our study is observational and cannot establish a causal relationship; further experimental and clinical research is needed to validate these findings.

However, this study still has some limitations. Firstly, the sample size in this study is limited. In future studies, we will continue to expand the recruitment scope of research subjects, extend the data collection period, and collaborate on multicenter

studies to increase the sample size, so that the research results can better describe the microbial landscape of CC. Secondly, the integration of metabolomics and proteomics data into this study could provide a more comprehensive understanding of the molecular mechanisms, revealing the intricate relationships between the gut microbiota and colorectal cancer from multiple levels, identifying potential biomarkers and therapeutic targets, and driving advancements in personalized medicine and translational research. Finally, despite the association between the microbiome and gene expression being discovered, due to the observational nature of the study, the causality of these relationships cannot be determined, necessitating further research for validation. In future studies, we will conduct more in-depth experimental model studies and *in vitro* experiments, integrating multi-omics data to explore the causal mechanisms of microbiome-host interactions.

## Conclusions

Totally, we systematically investigated the special mucosal microbiome and transcriptome profiles in CC. Importantly, the significant correlation between 3,632 DEGs and tumor-associated microbiota in CC was proved. We also found that *Campylobacter jejuni* is related to bile secretion and CD4+ Tem and master cells of immune cells, which may contribute to the occurrence and development of CC. Finally, we also explored the survival value of tumor-associated microbiota on colon cancer, thus providing a new strategy for the treatment of CC. Our integrated analysis of transcriptome and microbiome has revealed potential connections between gut microbiota and host gene expression, contributing to a comprehensive understanding of the interactions between host and microbiota, and providing better strategies for the diagnosis and prevention of CC through further experiments. Moving forward, our research emphasizes the necessity for further studies to unravel the mechanisms by which these microbiota promote cancer. Future research should focus on investigating the complex interactions between these microbiota and host metabolism and immune system, and exploring the potential of harnessing microbiota in cancer treatment.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge each editor and reviewer for their efforts on this study.

This work was supported by grants from Guangxi Clinical Research Center for Anesthesiology (AD22035214), National Science Foundation of Guangxi (2022GXNSFAA035510), National Science Foundation of China (8236080196, 81760530), and Postdoctoral Science Foundation of China (2021M693803).

Y.Q., Q.W., and Q.L. collected the samples. F.L., X.P., C.W., J.C., and T.H. collected and performed the data curation. Y.Q., Q.W., Q.L., W.Y., and L.P. analyzed the data and interpreted the results. Y.Q., W.Y., M.F., and L.P. drafted and reviewed the manuscript. All authors read and approved the final manuscript.

## AUTHOR AFFILIATIONS

<sup>1</sup>Department of Clinical Laboratory, Guangxi Medical University Cancer Hospital, Nanning, China

<sup>2</sup>Guangxi Clinical Research Center for Anesthesiology, Guangxi Medical University Cancer Hospital, Nanning, China

<sup>3</sup>Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

<sup>4</sup>Institute of Advanced Biotechnology and School of Medicine, Southern University of Science and Technology, Shenzhen, China

AUTHOR ORCID*s*

Min Fang  <http://orcid.org/0009-0008-7336-9606>

Weilong Yang  <http://orcid.org/0000-0002-0327-4272>

Linghui Pan  <http://orcid.org/0000-0003-1107-1261>

## FUNDING

Funder	Grant(s)	Author(s)
<a href="#">Guangxi Clinical Research Center for Anesthesiology</a>	AD22035214	Min Fang
<a href="#">National Science Foundation of Guangxi</a>	2022GXNSFAA035510	Min Fang
<a href="#">National Science Foundation of China</a>	8236080196	Min Fang
<a href="#">National Science Foundation of China</a>	81760530	Min Fang
<a href="#">Postdoctoral Science Foundation of China</a>	2021M693803	Min Fang

## DATA AVAILABILITY

A STORMS (Strengthening The Organizing and Reporting of Microbiome Studies) checklist (79) is available at [10.5281/zenodo.14747658](https://doi.org/10.5281/zenodo.14747658). The 16S rRNA, metagenomic, and transcriptome sequencing data have been submitted to the Sequence Read Archive database at the National Center for Biotechnology Information website (<https://www.ncbi.nlm.nih.gov/>) under accession numbers [PRJNA911471](#), [PRJNA1028103](#) and [PRJNA1028158](#).

## ETHICS APPROVAL

Our study was approved by Guangxi Medical University Cancer Hospital Ethical Review Committee (approval no. LW2024072). We confirm that all methods have been implemented in accordance with Declaration of Helsinki and have obtained the informed consent of all subjects and/or their legal guardian.

## ADDITIONAL FILES

The following material is available [online](#).

## Supplemental Material

**Supplemental Figures (mSystems00805-24-s0001.docx).** Figures S1 to S8.

**Supplemental Tables (mSystems00805-24-s0002.xlsx).** Tables S1 to S10.

## REFERENCES

- Morgan E, Arnold M, Gini A, Lorenzoni V, Cabasag CJ, Laversanne M, Vignat J, Ferlay J, Murphy N, Bray F. 2023. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. *Gut* 72:338–344. <https://doi.org/10.1136/gutjnl-2022-327736>
- Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, Cercek A, Smith RA, Jemal A. 2020. Colorectal cancer statistics, 2020. *CA Cancer J Clin* 70:145–164. <https://doi.org/10.3322/caac.21601>
- Boland PM, Yurgelun MB, Boland CR. 2018. Recent progress in Lynch syndrome and other familial colorectal cancer syndromes. *CA Cancer J Clin* 68:217–231. <https://doi.org/10.3322/caac.21448>
- Graff RE, Möller S, Passarelli MN, Witte JS, Skytthe A, Christensen K, Tan Q, Adami H-O, Czene K, Harris JR, Pukkala E, Kaprio J, Giovannucci EL, Mucci LA, Hjelmborg JB. 2017. Familial risk and heritability of colorectal cancer in the nordic twin study of cancer. *Clin Gastroenterol Hepatol* 15:1256–1264. <https://doi.org/10.1016/j.cgh.2016.12.041>
- Sepich-Poore GD, Zitvogel L, Straussman R, Hasty J, Wargo JA, Knight R. 2021. The microbiome and human cancer. *Science* 371:eabc4552. <https://doi.org/10.1126/science.abc4552>
- Boesch M, Horvath L, Baty F, Pircher A, Wolf D, Spahn S, Straussman R, Tilg H, Brutsche MH. 2022. Compartmentalization of the host microbiome: how tumor microbiota shapes checkpoint immunotherapy outcome and offers therapeutic prospects. *J Immunother Cancer* 10:e005401. <https://doi.org/10.1136/jitc-2022-005401>
- Janney A, Powrie F, Mann EH. 2020. Host-microbiota maladaptation in colorectal cancer. *Nature New Biol* 585:509–517. <https://doi.org/10.1038/s41586-020-2729-3>
- El Tekle G, Garrett WS. 2023. Bacteria in cancer initiation, promotion and progression. *Nat Rev Cancer* 23:600–618. <https://doi.org/10.1038/s41568-023-00594-2>
- Zhao L-Y, Mei J-X, Yu G, Lei L, Zhang W-H, Liu K, Chen X-L, Kolat D, Yang K, Hu J-K. 2023. Role of the gut microbiota in anticancer therapy: from molecular mechanisms to clinical applications. *Signal Transduct Target Ther* 8:201. <https://doi.org/10.1038/s41392-023-01406-7>
- Clay SL, Fonseca-Pereira D, Garrett WS. 2022. Colorectal cancer: the facts in the case of the microbiota. *J Clin Invest* 132:e155101. <https://doi.org/10.1172/JCI155101>
- Hong J, Guo F, Lu S-Y, Shen C, Ma D, Zhang X, Xie Y, Yan T, Yu T, Sun T, Qian Y, Zhong M, Chen J, Peng Y, Wang C, Zhou X, Liu J, Liu Q, Ma X, Chen Y-X, Chen H, Fang J-Y. 2021. *F. nucleatum* targets lncRNA ENO1-IT1 to promote glycolysis and oncogenesis in colorectal cancer. *Gut* 70:2123–2137. <https://doi.org/10.1136/gutjnl-2020-322780>



12. Okumura S, Konishi Y, Narukawa M, Sugiura Y, Yoshimoto S, Arai Y, Sato S, Yoshida Y, Tsuji S, Uemura K, et al. 2021. Gut bacteria identified in colorectal cancer patients promote tumorigenesis via butyrate secretion. *Nat Commun* 12:5674. <https://doi.org/10.1038/s41467-021-25965-x>
13. Wang X, Jia Y, Wen L, Mu W, Wu X, Liu T, Liu X, Fang J, Luan Y, Chen P, Gao J, Nguyen K-A, Cui J, Zeng G, Lan P, Chen Q, Cheng B, Wang Z. 2021. *Porphyromonas gingivalis* promotes colorectal carcinoma by activating the hematopoietic NLRP3 inflammasome. *Cancer Res* 81:2745–2759. <https://doi.org/10.1158/0008-5472.CAN-20-3827>
14. Chung L, Thiele Orberg E, Geis AL, Chan JL, Fu K, DeStefano Shields CE, Dejea CM, Fathi P, Chen J, Finard BB, Tam AJ, McAllister F, Fan H, Wu X, Ganguly S, Lebid A, Metz P, Van Meerbeke SW, Huso DL, Wick EC, Pardoll DM, Wan F, Wu S, Sears CL, Housseau F. 2018. *Bacteroides fragilis* toxin coordinates a pro-carcinogenic inflammatory cascade via targeting of colonic epithelial cells. *Cell Host Microbe* 23:203–214. <https://doi.org/10.1016/j.chom.2018.01.007>
15. Kong C, Liang L, Liu G, Du L, Yang Y, Liu J, Shi D, Li X, Ma Y. 2023. Integrated metagenomic and metabolomic analysis reveals distinct gut-microbiome-derived phenotypes in early-onset colorectal cancer. *Gut* 72:1129–1142. <https://doi.org/10.1136/gutjnl-2022-327156>
16. Cai P, Xiong J, Sha H, Dai X, Lu J. 2023. Tumor bacterial markers diagnose the initiation and four stages of colorectal cancer. *Front Cell Infect Microbiol* 13:1123544. <https://doi.org/10.3389/fcimb.2023.1123544>
17. Kneis B, Wirtz S, Weber K, Denz A, Gittler M, Geppert C, Brunner M, Krautz C, Siebenhüner AR, Schierwagen R, Tyc O, Agaimy A, Grützmann R, Trebicka J, Kersting S, Langheinrich M. 2023. Colon cancer microbiome landscaping: differences in right- and left-sided colon cancer and a tumor microbiome-ileal microbiome association. *Int J Mol Sci* 24:3265. <https://doi.org/10.3390/ijms24043265>
18. Zhu Y, Ma L, Wei W, Li X, Chang Y, Pan Z, Gao H, Yang R, Bi Y, Ding L. 2022. Metagenomics analysis of cultured mucosal bacteria from colorectal cancer and adjacent normal mucosal tissues. *J Med Microbiol* 71. <https://doi.org/10.1099/jmm.0.001523>
19. Li Y, Cao H, Fei B, Gao Q, Yi W, Han W, Bao C, Xu J, Zhao W, Zhang F. 2022. Gut microbiota signatures in tumor, para-cancerous, normal mucosa, and feces in colorectal cancer patients. *Front Cell Dev Biol* 10:916961. <https://doi.org/10.3389/fcell.2022.916961>
20. Gao Z, Guo B, Gao R, Zhu Q, Qin H. 2015. Microbiota disbiosis is associated with colorectal cancer. *Front Microbiol* 6:20. <https://doi.org/10.3389/fmicb.2015.00020>
21. Sommer F, Nookaew I, Sommer N, Fogelstrand P, Bäckhed F. 2015. Site-specific programming of the host epithelial transcriptome by the gut microbiota. *Genome Biol* 16:62. <https://doi.org/10.1186/s13059-015-0614-4>
22. Sakai K, De Velasco MA, Kura Y, Nishio K. 2021. Transcriptome profiling and metagenomic analysis help to elucidate interactions in an inflammation-associated cancer mouse model. *Cancers (Basel)* 13:3683. <https://doi.org/10.3390/cancers13153683>
23. Chang Y, Huang Z, Hou F, Liu Y, Wang L, Wang Z, Sun Y, Pan Z, Tan Y, Ding L, Gao H, Yang R, Bi Y. 2023. Parvimonas micra activates the Ras/ERK/c-Fos pathway by upregulating miR-218-5p to promote colorectal cancer progression. *J Exp Clin Cancer Res* 42:13. <https://doi.org/10.1186/s13046-022-02572-2>
24. Gupta S, Hjelmsø MH, Lehtimäki J, Li X, Mortensen MS, Russel J, Trivedi U, Rasmussen MA, Stokholm J, Bisgaard H, Sørensen SJ. 2020. Environmental shaping of the bacterial and fungal community in infant bed dust and correlations with the airway microbiota. *Microbiome* 8:115. <https://doi.org/10.1186/s40168-020-00895-w>
25. Zhang K, Delgado-Baquerizo M, Zhu YG, Chu H. 2020. Space is more important than season when shaping soil microbial communities at a large spatial scale. *mSystems* 5. <https://doi.org/10.1128/msystems.00783-19>
26. Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JL, Knight R, Mills DA, Caporaso JG. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 10:57–59. <https://doi.org/10.1038/nmeth.2276>
27. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JL, et al. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <https://doi.org/10.1038/nmeth.f303>
28. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10:996–998. <https://doi.org/10.1038/nmeth.2604>
29. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596. <https://doi.org/10.1093/nar/gks1219>
30. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>
31. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M. 2006. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* 34:D354–D357. <https://doi.org/10.1093/nar/gkj102>
32. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 42:D199–205. <https://doi.org/10.1093/nar/gkt1076>
33. Powell S, Forslund K, Szklarczyk D, Trachana K, Roth A, Huerta-Cepas J, Gabaldón T, Rattei T, Creevey C, Kuhn M, Jensen LJ, von Mering C, Bork P. 2014. eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Res* 42:D231–D239. <https://doi.org/10.1093/nar/gkt1253>
34. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* 37:D233–D238. <https://doi.org/10.1093/nar/gkn663>
35. Noval Rivas M, Burton OT, Wise P, Zhang Y, Hobson SA, Garcia Lloret M, Chehoud C, Kuczynski J, DeSantis T, Warrington J, Hyde ER, Petrosino JF, Gerber GK, Bry L, Oettgen HC, Mazmanian SK, Chatila TA. 2013. A microbiota signature associated with experimental food allergy promotes allergic sensitization and anaphylaxis. *J Allergy Clin Immunol* 131:201–212. <https://doi.org/10.1016/j.jaci.2012.10.026>
36. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60. <https://doi.org/10.1186/gb-2011-12-6-r60>
37. Goldstein LD, Cao Y, Pau G, Lawrence M, Wu TD, Seshagiri S, Gentleman R. 2016. Prediction and quantification of splice events from RNA-Seq data. *PLoS One* 11:e0156132. <https://doi.org/10.1371/journal.pone.0156132>
38. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30:923–930. <https://doi.org/10.1093/bioinformatics/btt656>
39. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>
40. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. 2021. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation* 2:100141. <https://doi.org/10.1016/j.xinn.2021.100141>
41. Wang J, Wang Y, Wan L, Chen X, Zhang H, Yang S, Zhong L. 2023. Identification of lactate regulation pattern on tumor immune infiltration, therapy response, and DNA methylation in diffuse large B-cell lymphoma. *Front Immunol* 14:1230017. <https://doi.org/10.3389/fimmu.2023.1230017>
42. Tang Z, Kang B, Li C, Chen T, Zhang Z. 2019. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res* 47:W556–W560. <https://doi.org/10.1093/nar/gkz430>
43. Guo Y, Liu N. 2022. Systematic analysis and identification of molecular subtypes of TRP-related genes and prognosis prediction in lung adenocarcinoma. *J Oncol* 2022:1–19. <https://doi.org/10.1155/2022/5388283>
44. Schwabe RF, Jobin C. 2013. The microbiome and cancer. *Nat Rev Cancer* 13:800–812. <https://doi.org/10.1038/nrc3610>
45. Jia W, Xie G, Jia W. 2018. Bile acid-microbiota crosstalk in gastrointestinal inflammation and carcinogenesis. *Nat Rev Gastroenterol Hepatol* 15:111–128. <https://doi.org/10.1038/nrgastro.2017.119>
46. Gopalakrishnan V, Helmink BA, Spencer CN, Reuben A, Wargo JA. 2018. The influence of the gut microbiome on cancer, immunity, and cancer immunotherapy. *Cancer Cell* 33:570–580. <https://doi.org/10.1016/j.ccell.2018.03.015>
47. Park EM, Chelvanambi M, Bhutiani N, Kroemer G, Zitvogel L, Wargo JA. 2022. Targeting the gut and tumor microbiota in cancer. *Nat Med* 28:690–703. <https://doi.org/10.1038/s41591-022-01779-2>

48. Flemer B, Lynch DB, Brown JMR, Jeffery IB, Ryan FJ, Claesson MJ, O'Riordain M, Shanahan F, O'Toole PW. 2017. Tumour-associated and non-tumour-associated microbiota in colorectal cancer. *Gut* 66:633–643. <https://doi.org/10.1136/gutjnl-2015-309595>
49. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto J-M, et al. 2011. Enterotypes of the human gut microbiome. *Nature New Biol* 473:174–180. <https://doi.org/10.1038/nature09944>
50. Dai Z, Coker OO, Nakatsu G, Wu WKK, Zhao L, Chen Z, Chan FKL, Kristiansen K, Sung JJY, Wong SH, Yu J. 2018. Multi-cohort analysis of colorectal cancer metagenome identified altered bacteria across populations and universal bacterial markers. *Microbiome* 6:70. <https://doi.org/10.1186/s40168-018-0451-2>
51. Henstra C, van Praagh J, Olinga P, Nagelkerke A. 2021. The gastrointestinal microbiota in colorectal cancer cell migration and invasion. *Clin Exp Metastasis* 38:495–510. <https://doi.org/10.1007/s10585-021-10130-x>
52. Wang X, Meng M, Sun J, Gao W, Lin C, Yu C. 2024. *Klebsiella aerogenes* exacerbates colon tumorigenesis in the AOM/DSS-induced C57BL/6J mouse. *Biochem Biophys Res Commun* 694:149410. <https://doi.org/10.1016/j.bbrc.2023.149410>
53. Stenfor Arnesen LP, Fagerlund A, Granum PE. 2008. From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol Rev* 32:579–606. <https://doi.org/10.1111/j.1574-6976.2008.00112.x>
54. Liu C, Shen Y, Yang M, Chi K, Guo N. 2022. Hazard of staphylococcal enterotoxins in food and promising strategies for natural products against virulence. *J Agric Food Chem* 70:2450–2465. <https://doi.org/10.1021/acs.jafc.1c06773>
55. Young KT, Davis LM, DiRita VJ. 2007. *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat Rev Microbiol* 5:665–679. <https://doi.org/10.1038/nrmicro1718>
56. He Z, Gharaibeh RZ, Newsome RC, Pope JL, Dougherty MW, Tomkovich S, Pons B, Mirey G, Vignard J, Hendrixson DR, Jobin C. 2019. *Campylobacter jejuni* promotes colorectal tumorigenesis through the action of cytolethal distending toxin. *Gut* 68:289–300. <https://doi.org/10.1136/gutjnl-2018-317200>
57. Koch A, Mizrahi V. 2018. *Mycobacterium tuberculosis*. *Trends Microbiol* 26:555–556. <https://doi.org/10.1016/j.tim.2018.02.012>
58. Kim YM, Choi JO, Cho YJ, Hong BK, Shon HJ, Kim BJ, Park JH, Kim WU, Kim D. 2023. *Mycobacterium* potentiates protection from colorectal cancer by gut microbial alterations. *Immunology* 168:493–510. <https://doi.org/10.1111/imm.13586>
59. Jansz N, Faulkner GJ. 2021. Endogenous retroviruses in the origins and treatment of cancer. *Genome Biol* 22:147. <https://doi.org/10.1186/s13059-021-02357-4>
60. Ko EJ, Ock MS, Choi YH, Iovanna JL, Mun S, Han K, Kim HS, Cha HJ. 2021. Human endogenous retrovirus (HERV)-K *env* gene knockout affects tumorigenic characteristics of *nupr1* gene in DLD-1 colorectal cancer cells. *Int J Mol Sci* 22:3941. <https://doi.org/10.3390/ijms22083941>
61. Kang X, Ng SK, Liu C, Lin Y, Zhou Y, Kwong TNY, Ni Y, Lam TYT, Wu WKK, Wei H, Sung JJY, Yu J, Wong SH. 2023. Altered gut microbiota of obesity subjects promotes colorectal carcinogenesis in mice. *EBioMedicine* 93:104670. <https://doi.org/10.1016/j.ebiom.2023.104670>
62. Liu L, Xu M, Lan R, Hu D, Li X, Qiao L, Zhang S, Lin X, Yang J, Ren Z, Xu J. 2022. *Bacteroides vulgatus* attenuates experimental mice colitis through modulating gut microbiota and immune responses. *Front Immunol* 13:1036196. <https://doi.org/10.3389/fimmu.2022.1036196>
63. Sun Y, Zhou J, Du H, Zhou Z, Han Y, Luo M, Guo X, Gu M, Yang H, Xiao H. 2024. The anti-inflammatory potential of a strain of probiotic *Bifidobacterium pseudocatenulatum* G7: *in vitro* and *in vivo* evidence. *J Agric Food Chem* 72:10355–10365. <https://doi.org/10.1021/acs.jafc.3c07935>
64. Li Y, Li Q, Yuan R, Wang Y, Guo C, Wang L. 2024. *Bifidobacterium breve*-derived indole-3-lactic acid ameliorates colitis-associated tumorigenesis by directing the differentiation of immature colonic macrophages. *Theranostics* 14:2719–2735. <https://doi.org/10.7150/thno.92350>
65. Li Q, Hu W, Liu WX, Zhao LY, Huang D, Liu XD, Chan H, Zhang Y, Zeng JD, Coker OO, Kang W, Ng SSM, Zhang L, Wong SH, Gin T, Chan MTV, Wu JL, Yu J, Wu WKK. 2021. *Streptococcus thermophilus* inhibits colorectal tumorigenesis through secreting  $\beta$ -galactosidase. *Gastroenterology* 160:1179–1193. <https://doi.org/10.1053/j.gastro.2020.09.003>
66. Cheng K-W, Tseng C-H, Chen I-J, Huang B-C, Liu H-J, Ho K-W, Lin W-W, Chuang C-H, Huang M-Y, Leu Y-L, Roffler SR, Wang J-Y, Chen Y-L, Cheng T-L. 2022. Inhibition of gut microbial  $\beta$ -glucuronidase effectively prevents carcinogen-induced microbial dysbiosis and intestinal tumorigenesis. *Pharmacol Res* 177:106115. <https://doi.org/10.1016/j.phrs.2022.106115>
67. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, Etienne-Grimaldi M-C, Schiappa R, Guenot D, Ayadi M, Kirzin S, Chazal M, Fléjou J-F, Benchimol D, Berger A, Lagarde A, Pencreach E, Piard F, Elias D, Parc Y, Olschwang S, Milano G, Laurent-Puig P, Boige V. 2013. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med* 10:e1001453. <https://doi.org/10.1371/journal.pmed.1001453>
68. Wu Y, Wang X, Wu F, Huang R, Xue F, Liang G, Tao M, Cai P, Huang Y. 2012. Transcriptome profiling of the cancer, adjacent non-tumor and distant normal tissues from a colorectal cancer patient by deep sequencing. *PLoS ONE* 7:e41001. <https://doi.org/10.1371/journal.pone.0041001>
69. Ocvirk S, O'Keefe SJD. 2021. Dietary fat, bile acid metabolism and colorectal cancer. *Semin Cancer Biol* 73:347–355. <https://doi.org/10.1016/j.semcancer.2020.10.003>
70. Wang S, Dong W, Liu L, Xu M, Wang Y, Liu T, Zhang Y, Wang B, Cao H. 2019. Interplay between bile acids and the gut microbiota promotes intestinal carcinogenesis. *Mol Carcinog* 58:1155–1167. <https://doi.org/10.1002/mc.22999>
71. Schmitt M, Greten FR. 2021. The inflammatory pathogenesis of colorectal cancer. *Nat Rev Immunol* 21:653–667. <https://doi.org/10.1038/s41577-021-00534-x>
72. Hanus M, Parada-Venegas D, Landskron G, Wielandt AM, Hurtado C, Alvarez K, Hermoso MA, López-Köstner F, De la Fuente M. 2021. Immune system, microbiota, and microbial metabolites: the unresolved triad in colorectal cancer microenvironment. *Front Immunol* 12:612826. <https://doi.org/10.3389/fimmu.2021.612826>
73. Zhang J, Zhu S, Ma N, Johnston LJ, Wu C, Ma X. 2021. Metabolites of microbiota response to tryptophan and intestinal mucosal immunity: a therapeutic target to control intestinal inflammation. *Med Res Rev* 41:1061–1088. <https://doi.org/10.1002/med.21752>
74. Bachem A, Makhlof C, Binger KJ, de Souza DP, Tull D, Hochheiser K, Whitney PG, Fernandez-Ruiz D, Dähling S, Kastenmüller W, Jönsson J, Gressier E, Lew AM, Perdomo C, Kupz A, Figgett W, Mackay F, Oleshansky M, Russ BE, Parish IA, Kallies A, McConville MJ, Turner SJ, Gebhardt T, Bedoui S. 2019. Microbiota-derived short-chain fatty acids promote the memory potential of antigen-activated CD8<sup>+</sup> T cells. *Immunity* 51:285–297. <https://doi.org/10.1016/j.immuni.2019.06.002>
75. Wong CC, Yu J. 2023. Gut microbiota in colorectal cancer development and therapy. *Nat Rev Clin Oncol* 20:429–452. <https://doi.org/10.1038/s41571-023-00766-x>
76. Fernandes MR, Aggarwal P, Costa RGF, Cole AM, Trinchieri G. 2022. Targeting the gut microbiota for cancer therapy. *Nat Rev Cancer* 22:703–722. <https://doi.org/10.1038/s41568-022-00513-x>
77. Brennan CA, Nakatsu G, Gallini Comeau CA, Drew DA, Glickman JN, Schoen RE, Chan AT, Garrett WS. 2021. Aspirin modulation of the colorectal cancer-associated microbe *Fusobacterium nucleatum*. *MBio* 12:e00547-21. <https://doi.org/10.1128/mBio.00547-21>
78. Wang Q, Wang K, Wu W, Lv L, Bian X, Yang L, Wang Q, Li Y, Ye J, Fang D, Wu J, Jiang X, Xie J, Lu Y, Li L. 2020. Administration of *Bifidobacterium bifidum* CGMCC 15068 modulates gut microbiota and metabolome in azoxymethane (AOM)/dextran sulphate sodium (DSS)-induced colitis-associated colon cancer (CAC) in mice. *Appl Microbiol Biotechnol* 104:5915–5928. <https://doi.org/10.1007/s00253-020-10621-z>
79. Mirzayi C, Renson A, Furlanello C, Sansone S-A, Zohra F, Elsafoury S, Geistlinger L, Kasselmann LJ, Eckenrode K, van de Wijgert J, et al. 2021. Reporting guidelines for human microbiome research: the STORMS checklist. *Nat Med* 27:1885–1892. <https://doi.org/10.1038/s41591-021-1552-x>