



A signature of *Prevotella copri* and *Faecalibacterium prausnitzii* depletion, and a link with bacterial glutamate degradation in the Kenyan colorectal cancer patients

Sarah Obuya^{1,2#}, Amr Elkholy^{3,4#}, Nagavardhini Avuthu^{5#}, Michael Behring³, Prachi Bajpai³, Sumit Agarwal³, Hyung-Gyoon Kim³, Nefertiti El-Nikhely⁵, Pamela Akinyi⁶, James Orwa⁷, Farrukh Afaq³, Mohammed Abdalla⁵, Anwar Michael², Mohamed Farouk², Lori Brand Bateman^{8,9}, Mona Fouad^{8,9}, Mansoor Saleh¹⁰, Chittibabu Guda⁵, Upender Manne^{3,9^}, Waleed Arafat²

¹Moi Teaching and Referral Hospital, Moi University, Kesses, Kenya; ²Department of Clinical Oncology and Nuclear Medicine, Faculty of Medicine, University of Alexandria, Alexandria, Egypt; ³Department of Pathology, University of Alabama at Birmingham, AL, USA; ⁴Department of Biotechnology, Institute of Graduate Studies and Research, Alexandria University, Alexandria, Egypt; ⁵Department of Genetics, Cell Biology and Anatomy, University of Nebraska Medical Center, Omaha, NE, USA; ⁶International Cancer Institute, Eldoret, Kenya; ⁷Department of Population Health, Aga Khan University, Nairobi, Kenya; ⁸Division of Preventive Medicine, University of Alabama at Birmingham, AL, USA; ⁹Department of Hematology-Oncology, Aga Khan University, Nairobi, Kenya; ¹⁰O'Neal Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL, USA

Contributions: (I) Conception and design: W Arafat, U Manne, M Saleh, C Guda, M Fouad, S Obuya; (II) Administrative support: None; (III) Provision of study materials or patients: S Obuya, N El-Nikhely, P Bajpai, S Agarwal, M Abdalla; (IV) Collection and assembly of data: S Obuya, A Elkholy, P Akinyi, J Orwa, A Michael, M Farouk, HG Kim, LB Bateman; (V) Data analysis and interpretation: A Elkholy, N Avuthu, M Behring; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Waleed Arafat, MBChB, PhD. Professor of Clinical Oncology, Department of Clinical Oncology and Nuclear Medicine, Faculty of Medicine, Alexandria University, Alexandria 21131, Egypt. Email: waleed.arafat@alexmed.edu.eg; Upender Manne, MS, PhD. Professor and Director of Translational Anatomic Pathology, Wallace Tumor Institute, Room # 420A, University of Alabama at Birmingham, Birmingham, AL 35233, USA. Email: upendermanne@uabmc.edu.

Background: Colorectal cancer (CRC) is the fifth most diagnosed cancer in Sub-Saharan Africa. In Kenya, CRC incidence rates tripled from 1997 to 2017. In the Moi Teaching and Referral Hospital, Moi University, there has been an increase in CRC cases, notably for younger patients. A suggested pathobiology for this increase is gut microbiome dysbiosis. Since, for the Kenyan CRC patient population, microbiome studies are rare, there is a need for a better understanding of how microbiome dysbiosis influences CRC epidemiology in Kenya. In this single-center study, the focus was on profiling the gut microbiome of Kenyan CRC patients and healthy volunteers and evaluating associations between microbiome profiles and the age of CRC patients.

Methods: The gut mucosa-associated microbiome of 18 CRC patients and 18 healthy controls were determined by 16S rRNA sequencing and analyzed for alpha and beta diversity, differential abundance, and microbial metabolic profiling.

Results: Alpha diversity metrics showed no significant differences, but beta diversity metrics showed dissimilarities in the microbial communities between CRC patients and healthy controls. The most underrepresented species in the CRC group were *Prevotella copri* (*P. copri*) and *Faecalibacterium prausnitzii* (*F. prausnitzii*), although *Bacteroides fragilis* (*B. fragilis*) and *Prevotella nigrescens* were overrepresented (linear discriminant analysis, LDA score >2, P < 0.05). Also, for CRC patients, significant metagenomic functional alterations were evident in microbial glutamate metabolic pathways (L-glutamate degradation VIII was enriched, and L-glutamate and L-glutamine biosynthesis were diminished) (P < 0.05, log₂ Fold Change > 1).

[^] ORCID: 0000-0002-1545-3032.

Moreover, the microbiome composition was different for patients under 40 years of age compared to older patients (LDA score >2, $P < 0.05$).

Conclusions: Microbiome and microbial metabolic profiles of CRC patients are different from those of healthy individuals. CRC microbiome dysbiosis, particularly *P. copri* and *F. prausnitzii* depletion and glutamate metabolic alterations, are evident in Kenyan CRC patients.

Keywords: Colorectal cancer (CRC); microbiome; *Prevotella copri*; *Faecalibacterium prausnitzii*; Kenya

Submitted Feb 09, 2022. Accepted for publication Jun 16, 2022.

doi: 10.21037/jgo-22-116

View this article at: <https://dx.doi.org/10.21037/jgo-22-116>

Introduction

In Sub-Saharan Africa (ssAfrica), colorectal cancer (CRC) is the fifth most diagnosed cancer. In Kenya, CRC incidence rates tripled from 1997 to 2017 (1), with cases predominantly of late-stage disease and with 20% of individuals 40 years of age or younger at diagnosis (2). A suggested pathobiology for this increase in CRC rates is gut dysbiosis, a change in the microbial community of the gut (3). Dysbiosis is linked to various colonic diseases, including CRC (4). For African populations, however, little work has been conducted to demonstrate a relationship between dysbiosis and CRC. Given the increase in the incidence of early-onset CRC in ssAfrica, there is a need to understand the link between dysbiosis and CRC in the Kenyan population.

In the Moi Teaching and Referral Hospital, Moi University (MTRH/MU), situated in the Rift Valley of Kenya, there has been an increased incidence of CRC, particularly for younger patients (<40 years of age). This observation is consistent with overall trends in ssAfrica. Over the last two decades, there has been an increase in the incidence of CRC for all ages in ssAfrica (5-8). However, the incidence of early-onset CRC is a concern specific to the region. Indeed, early onset CRC was reported in ssAfrica with rates higher than those of resource-rich countries (RRCs). The proportion of early onset cases from single centers in ssAfrica is between 19–38% as compared to 3–7% in RRCs (8). Although suboptimal data surveillance and collection may lead to a reporting bias (9), the reasons for this marked increase in cases are largely unexplained. In Kenya, westernized high-fat diet, a known risk factor for CRC has replaced the traditional diet with high-fiber, fruits, and vegetables. Additionally, other risk factors that are part of traditional Kenyan dietary habits include drinking of alcohol and soot-laced sour milk, reuse of fry oil, and aflatoxin- and pesticide-contaminated foods (10).

The traditional African high-fiber diet shapes the gut microbiota to protect the colon from inflammation and noninfectious colonic diseases, but the European high-fat diet does not (11). Although previous reports showed involvement of diet and gut dysbiosis in CRC (10,12), there are no comprehensive reports investigating the microbiome of Kenyan CRC patients.

Given an increasing CRC incidence rate, there is an unmet need for a better understanding of how the microbiome and age of CRC onset influence CRC epidemiology in Africa. In this single-center cross-sectional observational study, our objective was to establish the tumor microbiome profile of Kenyan CRC patients relative to the healthy population. We present the following article in accordance with the STROBE reporting checklist (available at <https://jgo.amegroups.com/article/view/10.21037/jgo-22-116/rc>).

Methods

Ethical consideration

The study was approved by the Institutional Research and Ethics Committee (IREC) of MTRH/MU, Kenya (No. IREC/2018/38, 0003114) to perform studies in collaboration with the School of Medicine, Alexandria University, Egypt. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Study participants were informed about the study. Those who agreed to participate signed the informed consent agreement. A step-by-step explanation in simple language was given to the participating patients. Confidentiality was provided according to international standards, and no information was used without consent of the patients. Coding was employed to conceal the identity of all patients. The tissues obtained were used only for the consented purpose.

Table 1 Description of CRC patients and healthy controls

Characteristics	CRC (n=18)	Healthy controls (n=18)
Age (years)	49.6±18.3	40.8±15.1
Sex		
Male	12 (67%)	9 (50%)
Female	6 (33%)	9 (50%)
Colonoscopy		
Normal colon	–	13 (72%)
Colitis	–	3 (17%)
Polyp (benign)	–	1 (6%)
Roundworm	–	1 (6%)
Stage		
III	8 (44%)	–
IV	10 (56%)	–

Data are presented as mean ± standard deviation or numbers and percentages. CRC, colorectal cancer.

Patients

Colonic tissues from 18 CRC patients and 18 healthy controls were collected during colonoscopy. This was convenience sampling of a medically underserved population. The inclusion criteria for CRC patients were being FOLFOX chemotherapy naïve, having undergone colonoscopy from a clinician's indication, and having pathologically confirmed CRC at the MTRH/MU. All participants who had taken any probiotic or antibiotic within 4 weeks prior to colonoscopy, and those who previously had resectable CRC were excluded. The study characteristics are shown in *Table 1*. Study data were collected at the MTRH/MU, which mainly serves residents of Western Kenya region (representing at least 22 counties) (13) during the period of June-2018 to June-2019.

Taxonomic and metabolic profiling of gut microbiome and diversity analysis

DNA was extracted from tissues collected from 18 CRC patients and 18 healthy controls, using QIAamp DNA Microbiome kits (Qiagen, Hilden, Germany), according to manufacturer's instructions. For all samples, the V4 hypervariable region of the 16S rRNA gene was amplified using 16S V4 amplification primers (14) and sequenced with an Illumina MiSeq platform at the UAB Genomic

Core. For each sample, the raw paired-end sequencing reads were de-multiplexed into separate FASTQ files, and barcodes and primers were removed. Then the sequence data were imported into the *QIIME2 2019.10* environment (15) and analyzed for taxonomic composition and ecological diversity of the gut microbiome. The microbiome profiling procedure followed our previous publication with minor modifications (16). Briefly, the QIIME2-DADA2 denoising method was used to filter sequencing reads (*Table S1*) and to construct a feature table of amplicon sequence variants (ASVs). The QIIME2 feature-classifier plugin with pre-trained Naïve Bayes classifier, trained on 'Greengenes 13_8 99%' OTUs from the V4 region of sequences (17) was used to assign the comparative taxonomy for ASVs. Rare ASVs, contaminants, and unclassified ASVs were filtered from the identified taxonomy table. Low-depth samples were also excluded using the cut-off of 2,000 reads in each sample. The QIIME2 taxa plugin was used to group the features with the same taxonomic assignment and collapsed to species-level taxonomic tables for subsequent analysis. In species-level taxonomic profiles, the taxa that were not differentiated to species-level were represented at higher levels of classification.

We performed alpha (microbiome diversity) and beta (similarity or dissimilarity of two microbial communities) diversity analysis on taxonomic profiles of CRC and control samples. Using QIIME2's diversity plugin, we calculated alpha diversity indices (Shannon, Simpson, Chao1, and Observed OUTs) and beta diversity indices (Brays-Curtis, Jaccard, Unweighted Unifrac, and weighted Unifrac measures) on rarefied feature tables at a minimum sampling depth of 10,135 reads through the core-metrics-phylogenetic method. The alpha and beta diversity measures at species-level were compared among/between healthy controls and tumor samples using Wilcoxon rank-sum tests. Weighted and unweighted Unifrac distances were used for ordination of CRC and healthy controls through NMDS analysis using the VEGAN R package. The significance of NMDS results were evaluated using the PERMANOVA method implemented in QIIME2 diversity plugin with 10^5 permutations. For subsequent analyses, the taxonomic profiles were filtered to remove the microbial taxa with maximum relative abundance <0.01% and prevalence <5% across all samples.

The metabolic potentials of the healthy gut (control) and the CRC microbiome were predicted using *PICRUSt2 v2.2* software (18). The predicted MetaCyc pathway profiles were filtered to include pathways with relative abundance >0.1% and prevalence in >50% of the samples used for further analysis.

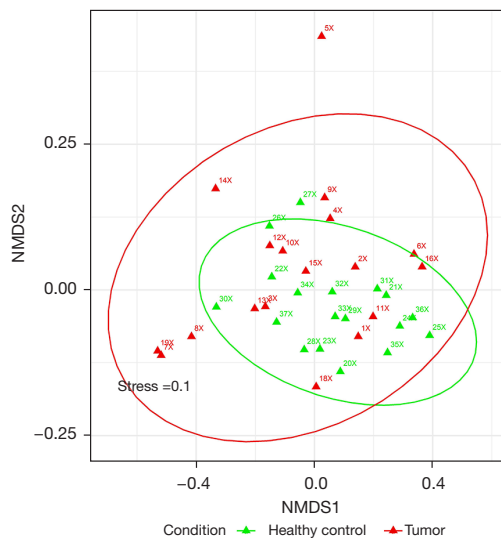


Figure 1 Beta diversity analysis: Non-metric Multidimensional Scaling (NMDS) ordination plots showing beta diversity comparisons between healthy control and CRCs. The beta diversities were measured based on weighted UniFrac distances. Ellipses represent the 95% confidence interval for healthy control and tumor groups, each dot in the ellipse represents a sample from that group. Statistics were calculated using Permutational Multivariate Analysis of Variance (PERMANOVA) with 999 permutations. A significant separation between healthy controls and CRC was found at $P=0.049$.

Statistical analysis

Differential analysis of microbial abundance was accomplished across two groups: (I) CRC cases *vs.* healthy controls, and (II) CRC patients under 40 years of age *vs.* CRC patients over 40 years. We tested the differential abundance at the species level using linear discriminant analysis (LDA) effect size (LEfSe) (19). P values <0.05 and LDA score (\log_{10}) >2 are considered significant for differential features; results were represented as barplots and cladograms. Differential analysis of pathways was accomplished using Wilcoxon rank-sum tests, and the results were plotted using the ggplot2 package of R. Of note, no sensitivity analyses were possible given the uniqueness of this population.

Results

Characteristics of study participants

Compared to healthy controls, CRC patients were slightly older (49.6 *vs.* 40.8 years), and mostly (67%) male.

Colonoscopy of healthy controls identified normal colon for 72% of patients, with 17% having colitis, and individual cases of benign colon polyps and roundworm. In general, CRC patients were more often males, with half of CRCs located in the colon, and with more cases of stage IV than III (Table 1).

Microbiome analysis

Gut microbial community structure and diversity

The 16S rRNA sequencing identified a total of 1,437 ASVs from the non-chimeric sequence reads ranging from 14,412 to 222,054 (Table S1). The rarefaction curves of observed ASVs/OUT confirmed that all samples had sufficient sequencing depth to explain the microbial diversity in each sample. The species richness evaluated by observed operational taxonomic units (OTUs) ($P=0.22$) and Chao1 estimator ($P=0.15$) showed no significant richness differences between CRC and healthy controls. Similar patterns of biodiversity were evident in measures of Simpson ($P=0.16$) and Shannon indices ($P=0.46$). These results suggested that there were no quantitative species differences between CRC and healthy controls (Figure S1).

Beta diversity measures showed dissimilarities between the CRC and healthy gut tissue microbial community structures. UniFrac matrices consider the phylogeny along with the taxonomic abundance to compute differences between communities. The UniFrac distance measures, visualized through the NMDS method (Figure 1), revealed that the microbiota composition for patients with CRC was significantly different from that for healthy controls (weighted UniFrac; stress =0.1, unweighted UniFrac; stress =0.181). Further analysis based on PERMANOVA on both UniFrac distances showed significant differences in community structures among healthy and CRC samples (on weighted UniFrac, $P=0.049$, on unweighted UniFrac, $P=0.0006$) (Figure S2).

Dominant microbial communities in CRC and healthy tissues were *Bacteroidetes*, *Firmicutes*, *Proteobacteria*, and *Fusobacteria*. Other phyla detected at low relative abundance were *Actinobacteria*, *Verrucomicrobia*, *Synergistetes*, *Cyanobacteria*, *Tenericutes*, *Spirochaetes*, *Euryarchaeota*, *Lentisphaerae*, and *Elusimicrobia*. Some CRC samples were dominated by one type of taxa such as *Fusobacteria* or *Proteobacteria* with greater than 60% relative proportions (Figure S3). Cumulative relative abundance analysis showed that the relative proportions of *Proteobacteria*, *Fusobacteria*, *Firmicutes*, *Actinobacteria*, and *Verrucomicrobia* were higher in CRC samples than in control samples, and the relative

proportion of *Bacteroidetes* was higher in healthy controls relative to CRC samples.

Differential microbial associations between CRC patients and healthy controls

Using LEfSe analysis, we identified potential microbiome biomarkers associated with the CRC patients (n=18) and the healthy group (n=18). CRC patient samples showed high abundance of *Bacillales* at the order level and *Tissierellaceae*, *Staphylococcaceae*, *Helicobacteraceae*, and *Mycoplasmataceae* at family level compared to healthy controls. In CRC samples, there were also several abundant genera *Fusobacterium*, *Porphyromonas*, *Staphylococcus*, and *Veillonella*. Furthermore, the species *Prevotella copri* (*P. copri*), *Faecalibacterium prausnitzii* (*F. prausnitzii*), *Roseburia faecis*, *Collinsella aerofaciens*, *Haemophilus parainfluenzae*, *Eubacterium bifforme*, *Dorea formicigenerans*, *Ruminococcus lactaris*, and *Bacteroides eggertii* were enriched in healthy controls, whereas *Bacteroides fragilis* (*B. fragilis*), *Prevotella nigrescens*, *Veillonella dispar*, *Veillonella parvula*, *Moryella indoligenes*, *Bulleidia moorei*, *Lactobacillus iners*, *Peptostreptococcus anaerobius*, and *Campylobacter ureolyticus* were enriched in CRC samples (LDA score >2, P<0.05) (Figure 2, Figure S4).

Differential microbial associations between CRC patients under 40 years of age vs. CRC patients over 40

Since studies from Kenya report a higher incidence of CRC in individuals of age 40 years or younger (2), we analyzed the data in CRC patients under 40 years of age and those over 40 years.

CRC patients under 40 years of age (n=8) showed high abundance of taxa belong to *Xanthomonadales*, *Oxalobacteraceae*, *Xanthomonadaceae*, *Herbaspirillum*, *Enhydrobacter*, and *Veillonella* compared to CRC patients over 40 (n=10) (Figure 3). At the species level, *P. copri* and *Lactobacillus salivarius* were enriched for CRC patients under 40, whereas *Clostridium ramosum*, *Clostridium symbiosum*, and *Bacteroides ovatus* were enriched for CRC patients over 40 (LDA score >2, P<0.05) (Figure 3, Figure S5).

Gut microbiome function is altered for CRC patients

In addition to microbiome dysbiosis for CRC patients, there were significant metagenomic alterations in microbial function for CRC patients compared to healthy patients. Notably, for CRC patients, L-glutamate degradation VIII pathway were enriched, and L-glutamate and L-glutamine biosynthesis were diminished (P<0.05, |log₂FC| >1) (Figure 4).

Discussion

Although there is a multifactorial etiology of CRC carcinogenesis, there is evidence that dysbiosis of the gut microbiota is linked to CRC. In our study, we revealed differing mucosa-associated microbiome and microbiome-functional profiles for Kenyan CRC patients. *P. copri* and *F. prausnitzii* depletion may be a signature of these patients. The results also indicate a link between CRC and the microbial glutamate metabolic pathways. Moreover, results showed high abundance of the genus *Herbaspirillum* in early-onset CRC. We propose future development of non-invasive microbial diagnostic tests.

We found that multiple microbiota are linked to CRC. CRC patients had a high abundance of *B. fragilis* (4); such abundance was linked with CRC (20–23). *B. fragilis* induces the stemness of CRC through Wnt/ β -catenin signaling (21,22,24,25) and initiates colonic inflammation as it induces irreversible oxidative damage to DNA, epithelial barrier disruption, activation of STAT3, and differentiation of T-helper (Th17) cells (20,26). *Prevotella nigrescens* is enriched for patients with CRC (27), and *Veillonella dispar* is abundant for patients with adenocarcinoma (28).

For the Kenyan CRC patients, *F. prausnitzii* was depleted. This microbe, a member of Clostridium cluster IV, is a prominent butyrate-producing bacteria. It contributes to CD4 T cell differentiation into CD4CD8 $\alpha\alpha$ T cells (Tregs) in the colonic mucosa, where they also prevent excessive inflammatory responses. For CRC patients, *F. prausnitzii* and CD4CD8 $\alpha\alpha$ T cells are concomitantly depleted (29), suggesting an involvement of reduced *F. prausnitzii*/DP8 α T cells in colonic carcinogenesis. Other studies showed that *F. prausnitzii* confers anti-inflammatory effects in the colon [99], and oral administration of *F. prausnitzii* as a probiotic shows an immunomodulatory effect in murine models of inflammatory colonic diseases and CRC (30). The anti-inflammatory effects are attributed to the microbial anti-inflammatory molecule (MAM) protein, secreted by *F. prausnitzii*, which inhibits the nuclear factor- κ B pathway and reduces expression of interferon γ and interleukin 17 (31). Therefore, this butyrate-producing microbiota should be considered as a candidate adjuvant probiotic to prevent CRC and to improve its treatment outcomes. In our study, a high prevalence of *F. prausnitzii* and *P. copri* found in healthy individuals could be linked to butyrate production and to anti-inflammatory effects that contribute to their protective effect on/decreased risk of CRC development.

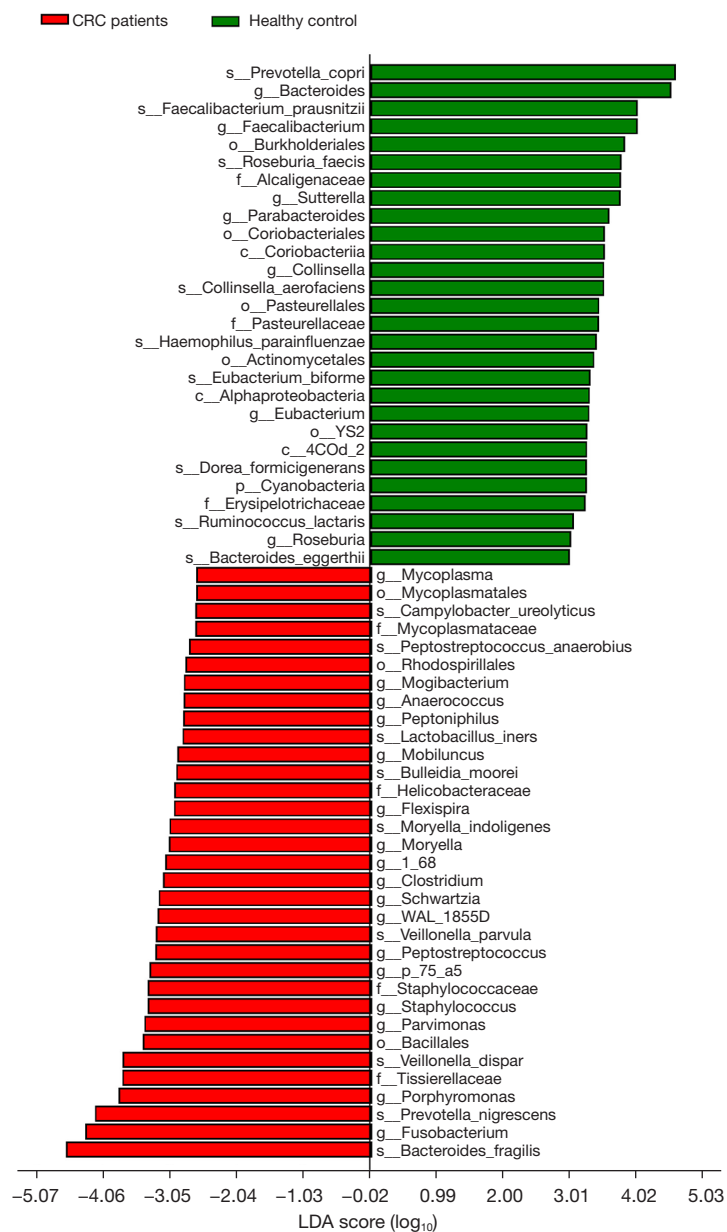


Figure 2 CRC patients and healthy controls differ in gut microbial composition: bar graph of LEfSe analysis CRC vs. healthy controls. Each bar represents the log₁₀ transformed LDA score for each differential microbe marked along with it (P<0.05). The CRC group is indicated by red bars and the control group is indicated by green bars. The name of the taxon level is abbreviated as p-phylum, c-class, o-order, f-family, and g-genus. CRC, colorectal cancer; LEfSe, Linear Discriminant Analysis (LDA) Effect Size.

Although precise reason for depletion of *F. prausnitzii* in CRC patients of this study is not known, a prior study reported that alcohol dependence depletes *Ruminococcaceae* bacteria, particularly *F. prausnitzii*, which results in gut leakiness (32). Moreover, in the study area of western Kenya, a traditional drink with a high-alcohol content called

“mursik”, a fermented milk consumed by people in this area; it has a possible etiology for esophageal cancer (33). In addition, as described earlier, the inhabitants of the study area typically consume maize as a main meal. Maize is a poor source of B-complex vitamins, particularly riboflavin and niacin (34). *F. prausnitzii* survive by exploiting extracellular

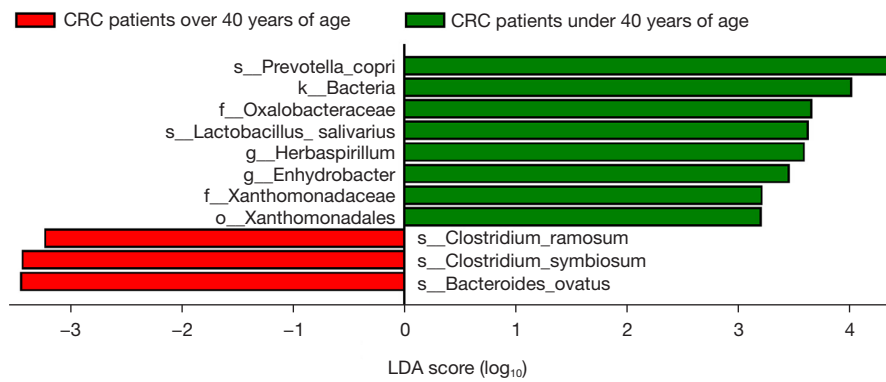


Figure 3 CRC patients under 40 years of age and CRC patients over 40 differ in gut microbial composition: bar graph of LEfSe analysis for CRC patients under 40 years of age *vs.* over 40 years of age. Each bar represents the log₁₀ transformed LDA score for each differential microbe marked along with it (P<0.05). The group with patients under 40 years of age is indicated by green bars and the group with patients over 40 years of age is indicated by red bars. The name of the taxon level is abbreviated as p-phyllum, c-class, o-order, f-family, and g-genus. CRC, colorectal cancer; LEfSe, Linear Discriminant Analysis (LDA) Effect Size.

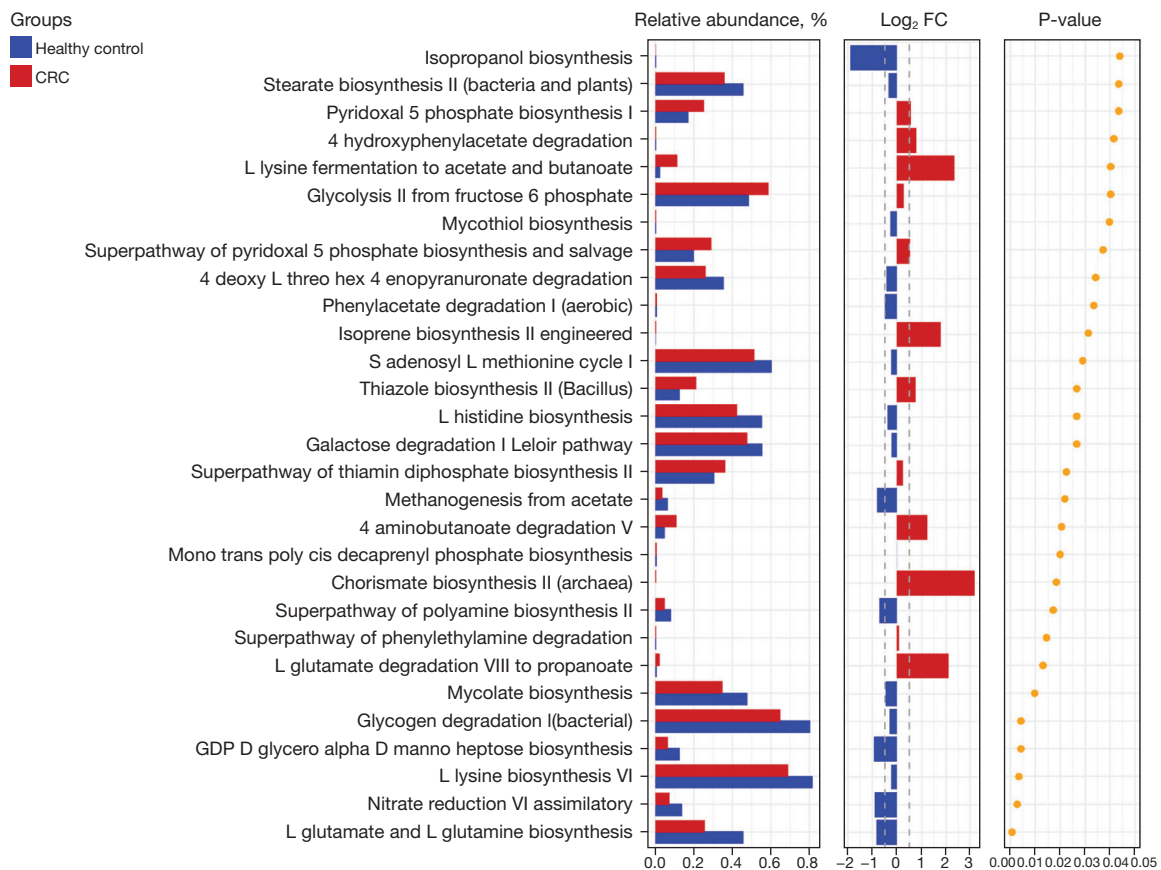


Figure 4 Bar plots showing the comparison of the significantly different *MetaCyc* pathways between CRC and healthy gut microbiome (P<0.05). Log₂ FC is CRC *vs.* Healthy control. CRC, colorectal cancer; FC, fold change.

antioxidants such as riboflavin (35). Studies to investigate the use of riboflavin as a prebiotic to restore *F. prausnitzii* depletion in Kenyan CRC patients may be of clinical interest because of this anti-inflammatory microbiota. Moreover, the cases were from a lower socio-economic status, and patients were less likely to have a healthy balanced diet, which could be a factor contributing to cancer. However, for our study participants, information on alcohol intake and diet is not available. Further studies linking microbiome, alcohol intake, and diet are needed to assess their impact on CRC risk.

In our results, *P. copri* was of high prevalence in healthy Kenyan individuals compared to CRC patients. *P. copri* was depleted in CRC patients, particularly those over 40 years of age compared to CRC patients under 40. Although murine colitis models show that *P. copri* is involved in inflammatory colonic diseases (36), no human studies have reported such findings. In fact, human studies suggest either reduced abundance of *P. copri* or no association of *P. copri* with Crohn's disease (37,38). Perhaps, in mice, the co-occurrence of other microbiota with *Prevotella* is the cause of initiation of intestinal inflammation, suggesting a need for more detailed investigations.

In Morocco, a North African country, the most significantly overrepresented species in healthy individuals compared to CRC patients were *P. copri* and *F. prausnitzii* (39). This suggests that depletion of these two microbiota is a signature of the CRC microbiome in some African populations. Moreover, meta-analysis studies show that *P. copri* is of high prevalence in healthy non-Westernized populations, suggesting that their diet is responsible for its low prevalence in Westernized populations (40).

The present results showed high abundance of the genus *Herbaspirillum* in early-onset CRC (under 40 years) compared to late-onset CRC (over 40 years). Higher abundances of *Herbaspirillum* are associated with NRAS mutations in CRC (41), suggesting the need for studies to investigate the link between the gut microbiome and host gene mutations in early-onset CRC in Kenya. Beta-diversity analysis revealed significant community-level separation between CRC patients and healthy individuals.

CRC microbiome functional analysis showed upregulation in L-glutamate degradation and downregulation of L-glutamate and L-glutamine biosynthesis, leading to low L-glutamate levels in the colon. L-glutamate metabolism, dietary L-glutamate degradation, and L-glutamate biosynthesis by gut microbiota modulate L-glutamate signaling (42). Alterations of glutamatergic transmission

in the microbiota-gut-brain axis may contribute to the pathogenesis of inflammation of the colon (42). In a neuro-inflammatory response in the gut, L-glutamate signaling induces oxidative and nitrosative stress pathways (43). Along with our results, this suggests that CRC programs the gut microbiota to protect against oxidative stress and assure cancer cell survival. Reactive oxygen species (ROS) are considered to induce oncogene activation, and tumor cells increase their antioxidant status by avoiding ROS thresholds that would trigger apoptosis (44). In this context, cancer cells not only display an endogenous adaptive response to oxidative stress by increasing expression of colon antioxidant enzymes and molecules, but they also cross-talk with the tumor microbiota to decrease oxidative stress by degrading L-glutamate. CRC patients with low serum glutamine levels have poorer survival than those with high glutamine levels (45). Further, low dietary intake of glutamine/glutamate is associated with high cancer mortality (46,47).

One of the limitations of this studies is the small sample size; thus, future large studies are needed to generalize our findings.

In conclusion, the CRC microbiome and microbiome function showed a significant difference relative to healthy volunteers. Additionally, various microbiota are candidate markers for CRC. CRC microbiome dysbiosis, particularly *P. copri* and *F. prausnitzii* depletion, may be a signature for Kenyan CRC patients, suggesting that further studies are needed to investigate the use of prebiotics/probiotics to restore *P. copri* and *F. prausnitzii* depletion in these patients. This may be of clinical interest because of the anti-inflammatory microbiota. The present results reflect a link between CRC and the microbial L-glutamate metabolic pathways. Thus, we propose that it is a microbiome functional signature for Kenyan CRC patients with possible alterations of glutamatergic transmission in the microbiota-gut axis that participate in the pathogenesis of CRC. Although this study was not focused on understanding the mechanisms of the microbiome in CRC carcinogenesis and progression, various studies are investigating the roles of gut dysbiosis in tumorigenesis and tumor progression as described in recent reviews (48,49).

Acknowledgments

We acknowledge the help provided by the Comprehensive Genomic Shared Facility of the O'Neal Comprehensive Cancer Center at UAB (P30 CA013148), and the Bioinformatics and Systems Biology Core (BSBC) Facility,

of University of Nebraska Medical Center, Omaha, Nebraska for supporting the computational infrastructure that is used to carry out data analysis in this study. BSBC receives support from Nebraska Research Initiative and NIH awards (2P20GM103427 and 5P30CA036727). We thank Dr. Donald Hill, a faculty member of the UAB O'Neal Comprehensive Cancer Center, for his editorial assistance.

Funding: These studies were partly supported by funding from the AMPATH and Moi Teaching and Referral Hospital, Kenya awarded to SO and by institutional funds (Department of Pathology and School of Medicine of the University of Alabama at Birmingham, UAB), awarded to UM. This article is also derived from the Subject Data funded in part by USAID and National Academy of Science (NAS) through sub-award 2000007148 (MF), and Science and Technology Development Fund-US cycle 17 (STDF-USC17-144) of Egypt awarded to MF and WA, respectively. Any opinions, findings, conclusions, or recommendations expressed in such article are those of the authors alone, and do not necessarily reflect the views of USAID or NAS.

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at <https://jgo.amegroups.com/article/view/10.21037/jgo-22-116/rc>

Data Sharing Statement: Available at <https://jgo.amegroups.com/article/view/10.21037/jgo-22-116/dss>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (Available at <https://jgo.amegroups.com/article/view/10.21037/jgo-22-116/coif>). UM serves as an unpaid editorial board member of Journal of Gastrointestinal Oncology from January 2021 to December 2022, and is supported by the Department of Pathology, the School of Medicine, and University of Alabama at Birmingham. MF is supported by the USAID and National Academy of Science (NAS) through sub-award 2000007148 (MF), and Science and Technology Development Fund-US cycle 17 (STDFUSC17-144) of Egypt awarded to MF and WA, respectively. CG is supported by the Bioinformatics and Systems Biology Core (BSBC) Facility, of University of Nebraska Medical Center, Omaha, Nebraska for supporting the computational infrastructure that is used to carry out data analysis in this study. BSBC receives support from Nebraska Research Initiative and NIH awards (2P20GM103427 and 5P30CA036727). WA is supported

by the USAID and National Academy of Science (NAS) through sub-award 2000007148 (MF), and Science and Technology Development Fund-US cycle 17 (STDF-USC17-144) of Egypt awarded to MF and WA, respectively. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Research and Ethics Committee (IREC) of Moi Teaching and Referral Hospital, Moi University (MTRH/MU), Kenya (NO. IREC/2018/38, 0003114). Candidates for enrollment were informed about the study while attending a colonoscopy outpatient clinic. Those who agreed to participate signed the informed consent agreement.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Graham A, Adeloje D, Grant L, et al. Estimating the incidence of colorectal cancer in Sub-Saharan Africa: A systematic analysis. *J Glob Health* 2012;2:020404.
- Parker RK, Ranketi SS, McNelly C, et al. Colorectal cancer is increasing in rural Kenya: challenges and perspectives. *Gastrointest Endosc* 2019;89:1234-7.
- Sears CL, Garrett WS. Microbes, microbiota, and colon cancer. *Cell Host Microbe* 2014;15:317-28.
- O'Keefe SJ, Li JV, Lahti L, et al. Fat, fibre and cancer risk in African Americans and rural Africans. *Nat Commun* 2015;6:6342.
- Bah E, Parkin DM, Hall AJ, et al. Cancer in the Gambia: 1988-97. *Br J Cancer* 2001;84:1207-14.
- Irabor D, Adedeji OA. Colorectal cancer in Nigeria: 40 years on. A review. *Eur J Cancer Care (Engl)* 2009;18:110-5.

7. Irabor DO. Emergence of Colorectal Cancer in West Africa: Accepting the Inevitable. *Niger Med J* 2017;58:87-91.
8. Katsidzira L, Gangaidzo I, Thomson S, et al. The shifting epidemiology of colorectal cancer in sub-Saharan Africa. *Lancet Gastroenterol Hepatol* 2017;2:377-83.
9. May FP, Anandasabapathy S. Colon cancer in Africa: Primetime for screening? *Gastrointest Endosc* 2019;89:1238-40.
10. Maiyoh GK, Tui VC. Rising Cancer Incidence and Role of the Evolving Diet in Kenya. *Nutr Cancer* 2019;71:531-46.
11. De Filippo C, Cavalieri D, Di Paola M, et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A* 2010;107:14691-6.
12. Song M, Chan AT, Sun J. Influence of the Gut Microbiome, Diet, and Environment on Risk of Colorectal Cancer. *Gastroenterology* 2020;158:322-40.
13. Moi Teaching and Referral Hospital. About Us - Moi Teaching and Referral Hospital 2020 [Moi Teaching and Referral Hospital (MTRH) started in 1917 with a bed capacity of 60 to cater for the Health]. Available online: http://www.mtrh.go.ke/?page_id=598.
14. Caporaso JG, Lauber CL, Walters WA, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* 2011;108 Suppl 1:4516-22.
15. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37:852-7.
16. Kumar SRP, Wang X, Avuthu N, et al. Role of Small Intestine and Gut Microbiome in Plant-Based Oral Tolerance for Hemophilia. *Front Immunol* 2020;11:844.
17. McDonald D, Price MN, Goodrich J, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 2012;6:610-8.
18. Douglas GM, Maffei VJ, Zaneveld JR, et al. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* 2020;38:685-8.
19. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;12:R60.
20. Yu LC, Wei SC, Ni YH. Impact of microbiota in colorectal carcinogenesis: lessons from experimental models. *Intest Res* 2018;16:346-57.
21. Toprak NU, Yagci A, Gulluoglu BM, et al. A possible role of *Bacteroides fragilis* enterotoxin in the aetiology of colorectal cancer. *Clin Microbiol Infect* 2006;12:782-6.
22. Wu S, Morin PJ, Maouyo D, et al. *Bacteroides fragilis* enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology* 2003;124:392-400.
23. Dahmus JD, Kotler DL, Kastenber DM, et al. The gut microbiome and colorectal cancer: a review of bacterial pathogenesis. *J Gastrointest Oncol* 2018;9:769-77.
24. Viljoen KS, Dakshinamurthy A, Goldberg P, et al. Quantitative profiling of colorectal cancer-associated bacteria reveals associations between *Fusobacterium* spp., enterotoxigenic *Bacteroides fragilis* (ETBF) and clinicopathological features of colorectal cancer. *PLoS One* 2015;10:e0119462.
25. Liu CJ, Zhang YL, Shang Y, et al. Intestinal bacteria detected in cancer and adjacent tissue from patients with colorectal cancer. *Oncol Lett* 2019;17:1115-27.
26. Bundgaard-Nielsen C, Baandrup UT, Nielsen LP, et al. The presence of bacteria varies between colorectal adenocarcinomas, precursor lesions and non-malignant tissue. *BMC Cancer* 2019;19:399.
27. Zeller G, Tap J, Voigt AY, et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol* 2014;10:766.
28. Kasai C, Sugimoto K, Moritani I, et al. Comparison of human gut microbiota in control subjects and patients with colorectal carcinoma in adenoma: Terminal restriction fragment length polymorphism and next-generation sequencing analyses. *Oncol Rep* 2016;35:325-33.
29. Toucheffeu Y, Duchalais E, Bruley des Varannes S, et al. Concomitant decrease of double-positive lymphocyte population CD4CD8 $\alpha\alpha$ and *Faecalibacterium prausnitzii* in patients with colorectal cancer. *Eur J Gastroenterol Hepatol* 2021;32:149-56.
30. Martín R, Chain F, Miquel S, et al. The commensal bacterium *Faecalibacterium prausnitzii* is protective in DNBS-induced chronic moderate and severe colitis models. *Inflamm Bowel Dis* 2014;20:417-30.
31. Quévrain E, Maubert MA, Michon C, et al. Identification of an anti-inflammatory protein from *Faecalibacterium prausnitzii*, a commensal bacterium deficient in Crohn's disease. *Gut* 2016;65:415-25.
32. Leclercq S, Matamoros S, Cani PD, et al. Intestinal permeability, gut-bacterial dysbiosis, and behavioral markers of alcohol-dependence severity. *Proc Natl Acad Sci U S A* 2014;111:E4485-93.
33. Nieminen MT, Novak-Frazer L, Collins R, et al. Alcohol and acetaldehyde in African fermented milk mursik--a possible etiologic factor for high incidence of esophageal

- cancer in western Kenya. *Cancer Epidemiol Biomarkers Prev* 2013;22:69-75.
34. Lopez-Siles M, Khan TM, Duncan SH, et al. Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. *Appl Environ Microbiol* 2012;78:420-8.
 35. Steinert RE, Sadaghian Sadabad M, Harmsen HJ, et al. The prebiotic concept and human health: a changing landscape with riboflavin as a novel prebiotic candidate? *Eur J Clin Nutr* 2016;70:1461.
 36. Elinav E, Strowig T, Kau AL, et al. NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. *Cell* 2011;145:745-57.
 37. Lewis JD, Chen EZ, Baldassano RN, et al. Inflammation, Antibiotics, and Diet as Environmental Stressors of the Gut Microbiome in Pediatric Crohn's Disease. *Cell Host Microbe* 2015;18:489-500.
 38. Gevers D, Kugathasan S, Knights D, et al. A Microbiome Foundation for the Study of Crohn's Disease. *Cell Host Microbe* 2017;21:301-4.
 39. Allali I, Boukhatem N, Bouguenouch L, et al. Gut microbiome of Moroccan colorectal cancer patients. *Med Microbiol Immunol* 2018;207:211-25.
 40. Tett A, Huang KD, Asnicar F, et al. The *Prevotella copri* Complex Comprises Four Distinct Clades Underrepresented in Westernized Populations. *Cell Host Microbe* 2019;26:666-679.e7.
 41. Sarhadi V, Lahti L, Saberi F, et al. Gut Microbiota and Host Gene Mutations in Colorectal Cancer Patients and Controls of Iranian and Finnish Origin. *Anticancer Res* 2020;40:1325-34.
 42. Baj A, Moro E, Bistoletti M, et al. Glutamatergic Signaling Along The Microbiota-Gut-Brain Axis. *Int J Mol Sci* 2019;20:1482.
 43. Kaszaki J, Erces D, Varga G, et al. Kynurenines and intestinal neurotransmission: the role of N-methyl-D-aspartate receptors. *J Neural Transm (Vienna)* 2012;119:211-23.
 44. Hayes JD, Dinkova-Kostova AT, Tew KD. Oxidative Stress in Cancer. *Cancer Cell* 2020;38:167-97.
 45. Ling HH, Pan YP, Fan CW, et al. Clinical Significance of Serum Glutamine Level in Patients with Colorectal Cancer. *Nutrients* 2019;11:898.
 46. Ma W, Heianza Y, Huang T, et al. Dietary glutamine, glutamate and mortality: two large prospective studies in US men and women. *Int J Epidemiol* 2018;47:311-20.
 47. Viana Veloso GG, Franco OH, Ruitter R, et al. Baseline dietary glutamic acid intake and the risk of colorectal cancer: The Rotterdam study. *Cancer* 2016;122:899-907.
 48. Chattopadhyay I, Dhar R, Pethusamy K, et al. Exploring the Role of Gut Microbiome in Colon Cancer. *Appl Biochem Biotechnol* 2021;193:1780-99.
 49. Park EM, Chelvanambi M, Bhutiani N, et al. Targeting the gut and tumor microbiota in cancer. *Nat Med* 2022;28:690-703.

Cite this article as: Obuya S, Elkholy A, Avuthu N, Behring M, Bajpai P, Agarwal S, Kim HG, El-Nikhely N, Akinyi P, Orwa J, Afaq F, Abdalla M, Michael A, Farouk M, Bateman LB, Fouad M, Saleh M, Guda C, Manne U, Arafat W. A signature of *Prevotella copri* and *Faecalibacterium prausnitzii* depletion, and a link with bacterial glutamate degradation in the Kenyan colorectal cancer patients. *J Gastrointest Oncol* 2022;13(5):2282-2292. doi: 10.21037/jgo-22-116

Table S1 Details of quality-score based filtering of the input sequencing reads using the QIIME2-DADA2 algorithm

Sample-id	Input	Filtered	% of input passed filter	Denoised	Merged	% of input merged	Non-chimeric	% of input non-chimeric
221-100	80286	59136	73.66	58556	56264	70.08	55860	69.58
221-101	88160	66844	75.82	66333	64360	73	62743	71.17
221-102	93058	69105	74.26	68590	66476	71.44	65741	70.65
221-103	89703	68455	76.31	67553	63809	71.13	63014	70.25
221-104	59400	44958	75.69	44185	41900	70.54	41624	70.07
221-105	94989	71068	74.82	70203	66587	70.1	65211	68.65
221-106	65142	50770	77.94	50321	48751	74.84	46577	71.5
221-107	40474	31077	76.78	30687	29449	72.76	29154	72.03
221-108	26617	21420	80.47	20939	17934	67.38	14412	54.15
221-109	102069	78168	76.58	77507	74872	73.35	72888	71.41
221-110	127464	95837	75.19	94711	91373	71.69	89452	70.18
221-111	81604	59106	72.43	58564	56084	68.73	56024	68.65
221-112	82951	65849	79.38	65124	62516	75.36	61721	74.41
221-113	111257	77847	69.97	77378	73035	65.65	70983	63.8
221-114	83759	62849	75.04	62423	59705	71.28	57503	68.65
221-115	48545	33139	68.26	32750	31050	63.96	31002	63.86
221-79	147535	108266	73.38	106740	101214	68.6	99155	67.21
221-80	71022	51056	71.89	50411	48508	68.3	48107	67.74
221-81	139977	82930	59.25	81612	72080	51.49	65837	47.03
221-82	152177	113246	74.42	111886	107572	70.69	106329	69.87
221-83	80978	65730	81.17	65365	61740	76.24	61365	75.78
221-84	128180	99043	77.27	98164	95137	74.22	93080	72.62
221-85	118558	86956	73.34	85953	83638	70.55	83356	70.31
221-86	256847	194346	75.67	192665	179599	69.92	177151	68.97
221-87	218101	164309	75.34	162494	155499	71.3	151978	69.68
221-88	100298	76134	75.91	72208	60250	60.07	51202	51.05
221-89	147559	113430	76.87	111505	100120	67.85	82576	55.96
221-90	332554	237411	71.39	234488	218460	65.69	198576	59.71
221-91	60731	45225	74.47	44503	42811	70.49	41349	68.09
221-92	81810	64543	78.89	63991	62274	76.12	61769	75.5
221-93	87123	63182	72.52	61277	54542	62.6	50938	58.47
221-94	113726	80326	70.63	79168	75810	66.66	72102	63.4
221-96	75753	57818	76.32	57275	53904	71.16	51942	68.57
221-97	149305	107617	72.08	106831	103282	69.18	100340	67.2
221-98	384311	251555	65.46	250730	228567	59.47	222054	57.78
221-99	67416	50611	75.07	50055	47780	70.87	46749	69.34

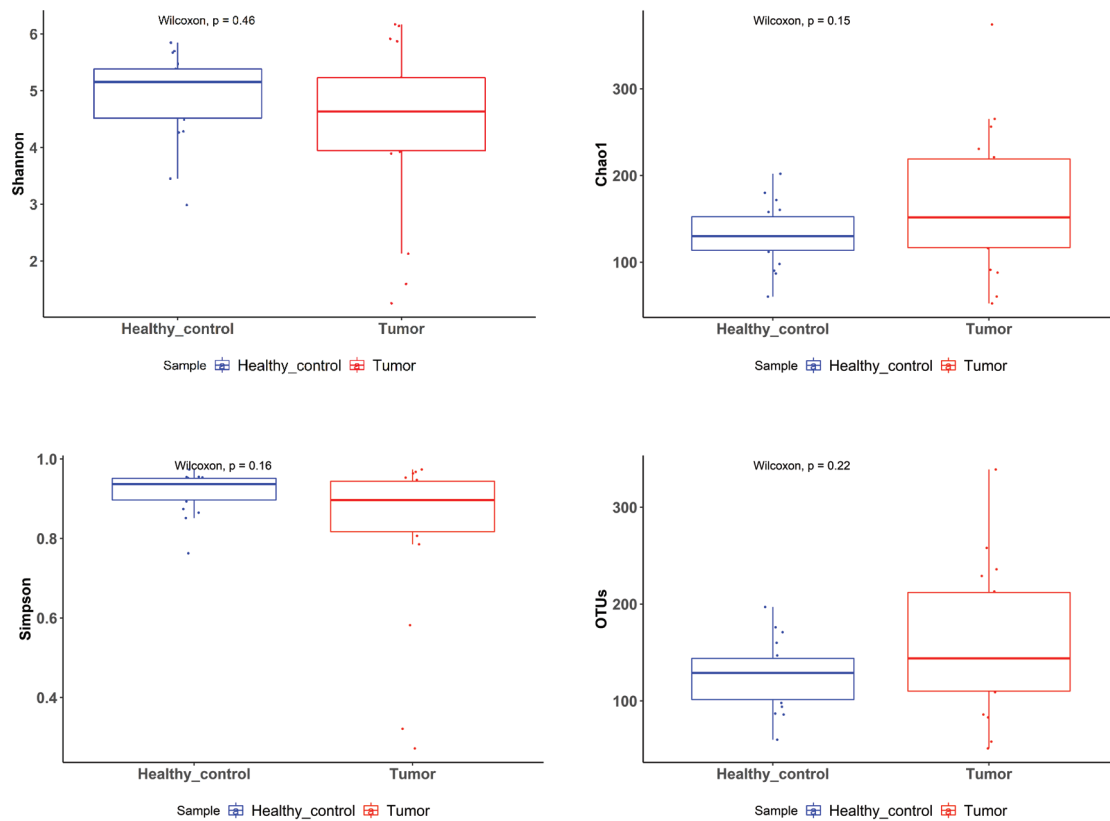


Figure S1 Reflects the alpha diversity measurements for CRC patients *vs.* healthy controls, measured by observed species, Chao1 richness estimator, Simpson diversity, and Shannon diversity and plotted for patients with CRC (red) and healthy controls (blue). The line inside the box represents the median; the whiskers represent the lowest and highest values within the 1.5 interquartile range (IQR). Outliers as well as individual sample values are shown as dots. CRC, colorectal cancer.

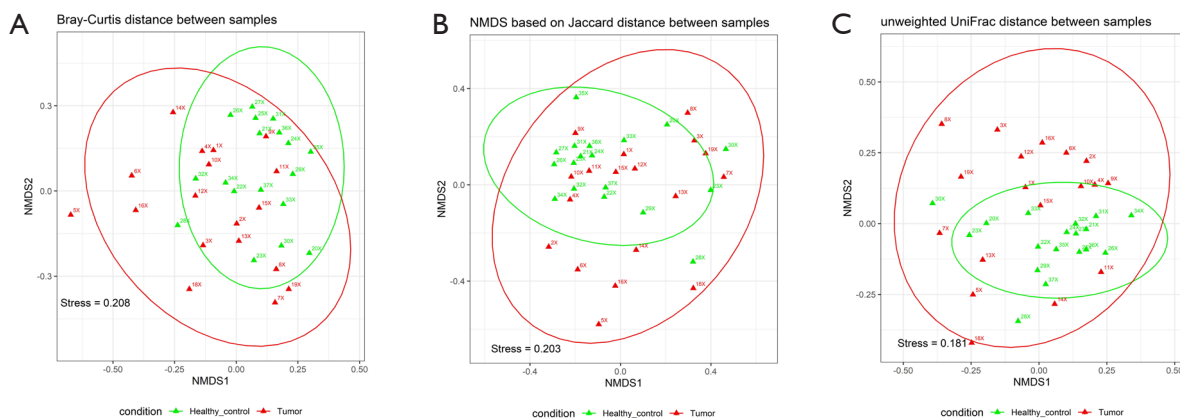


Figure S2 Beta diversity analysis: Non-metric Multidimensional Scaling (NMDS) ordination plots showing beta diversity comparisons between healthy control and tumor (CRC) samples. Ellipses represent the 95% confidence interval for healthy control and tumor groups, each dot in the ellipse represents a sample from that group. Statistics were calculated using PERMANOVA with 999 permutations. Beta diversity based on, (A) *Bray-Curtis* distance ($P=0.004$), (B) *Jaccard* distance ($P=0.002$), and (C) *unweighted UniFrac* distance ($P=0.0006$).

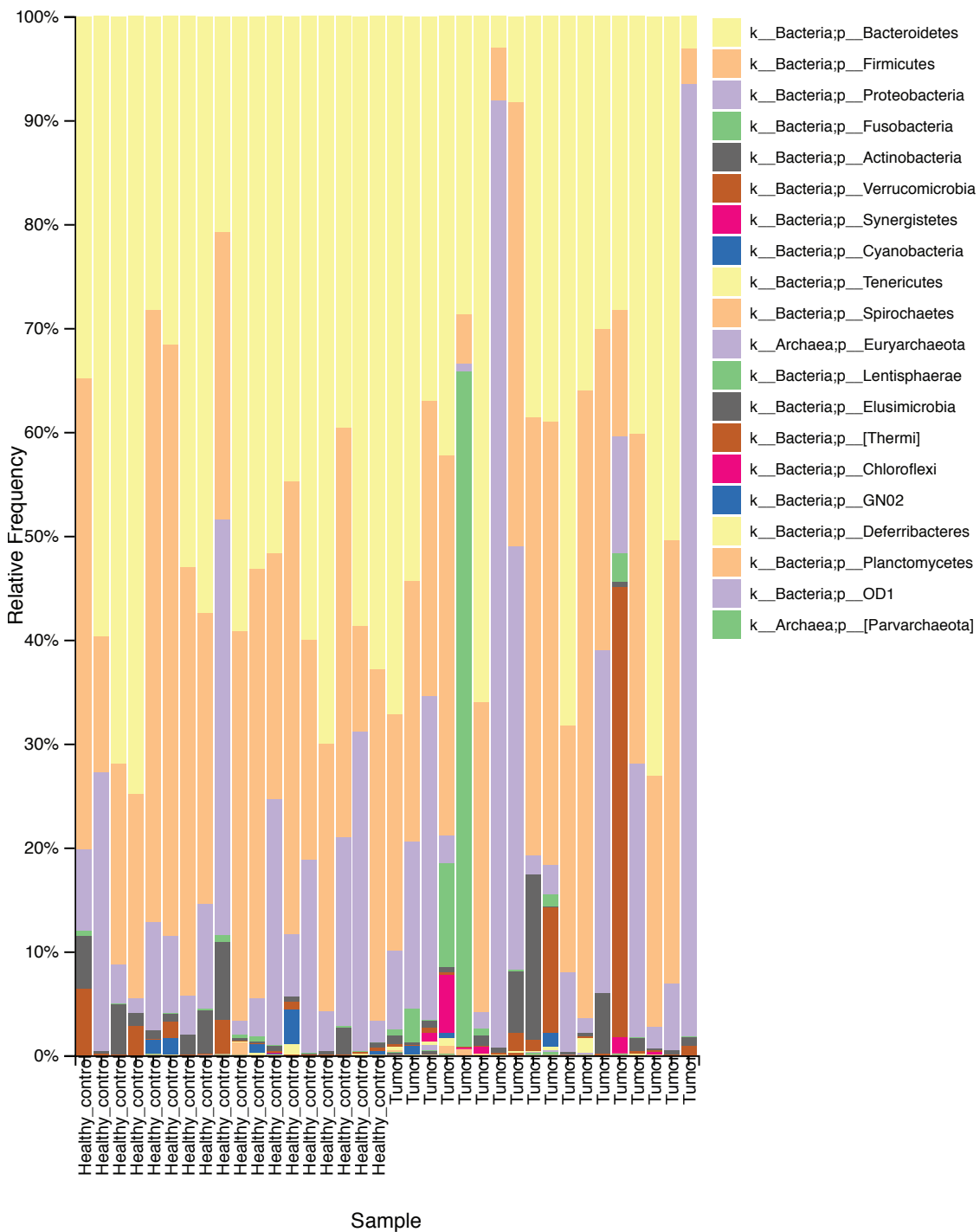


Figure S3 Relative frequencies of microbial taxa in CRC and healthy control samples at the phylum level.

Cladogram

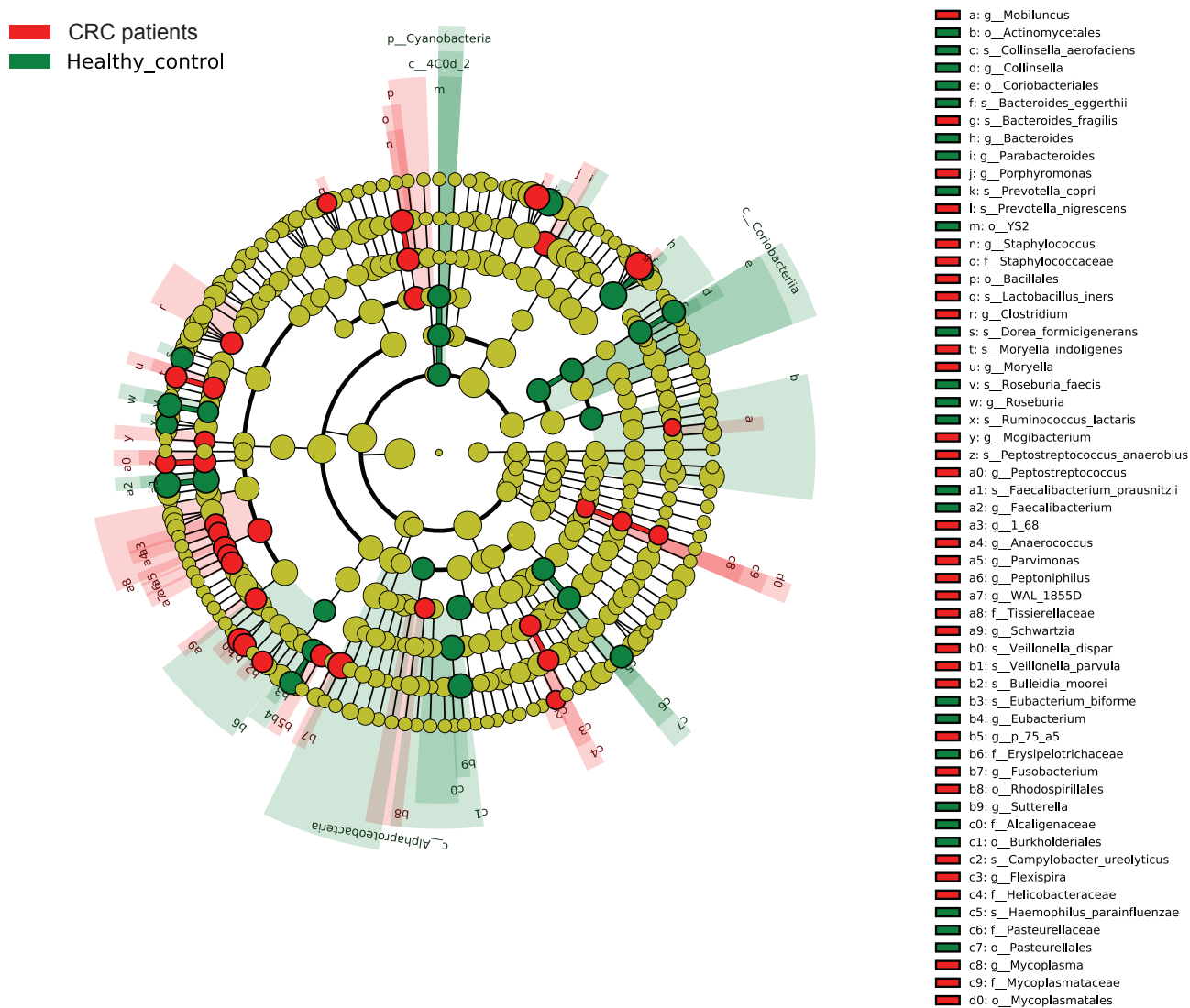


Figure S4 CRC patients and healthy controls differ in gut microbial composition. Cladogram for differentially distributed taxa ($P < 0.05$, $LDA > 2$) based on taxonomic abundance at the species level. The circles in the cladogram represent the phylogenetic tree from phylum to the species level. Each colored circle/dot on each phylogenetic level represents the taxa, and the diameter of the dot is proportional to the relative abundance of the taxa. Taxa with no significant differences are represented by the color yellow, whereas the other significant different taxa colored according to the group. CRC, colorectal cancer; LDA, Linear Discriminant Analysis.

Cladogram

- CRC patients over 40 years of age
- CRC patients under 40 years of age

- a: s_*Bacteroides_ovatus*
- b: s_*Prevotella_copri*
- c: s_*Lactobacillus_salivarius*
- d: s_*Clostridium_symbiosum*
- e: s_*Clostridium_ramosum*
- f: g_*Herbaspirillum*
- g: f_*Oxalobacteraceae*
- h: g_*Enhydrobacter*
- i: f_*Xanthomonadaceae*
- j: o_*Xanthomonadales*

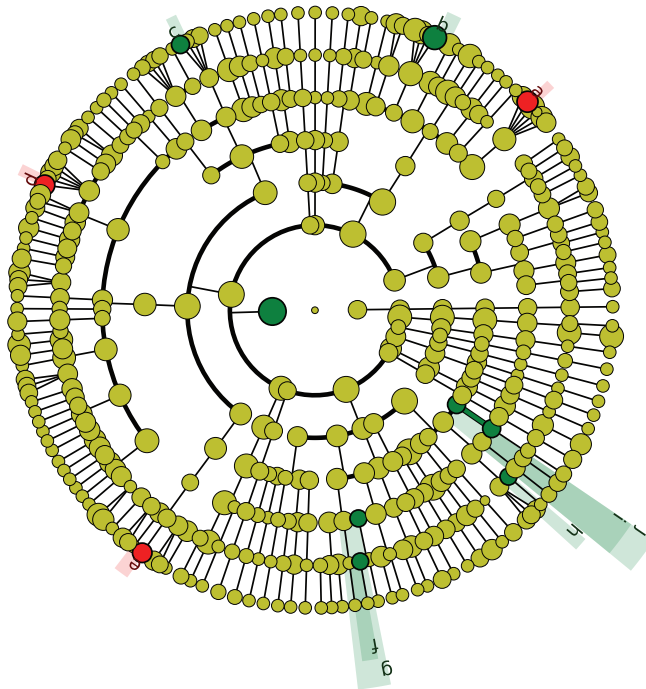


Figure S5 CRC patients under 40 years of age and CRC patients over 40 years of age differ in gut microbial composition. Cladogram for differentially distributed taxa ($P < 0.05$, $LDA > 2$) based on taxonomic abundance at the species level. The circles in the cladogram represent the phylogenetic tree from phylum to the species level. Each colored circle/dot on each phylogenetic level represents the taxa, and the diameter of the dot is proportional to the relative abundance of the taxa. Taxa with no significant differences are represented by the color yellow, whereas the other significant different taxa colored according to the group. CRC, colorectal cancer; LDA, Linear Discriminant Analysis.