Oral/gut microbiome and inflammatory markers in African American colon cancer patients

A thesis submitted to the faculty of
The School of Graduate Studies
The State University of New York
Downstate Health Sciences University

In partial fulfillment of the requirements
of the degree of Doctor of Philosophy

by

Sofia Tortora Morel, M.S, D.D.S

Program in Molecular and Cellular Biology
March 12, 2024
Thesis Advisor: Laura Martello-Rooney, Ph.D.
Department of Medicine, Division of Gastroenterology
Dedication

In loving memory of my abuela, Mabel Lara Yanes
12/1935-1/2023

You are truly missed.
Acknowledgments

I would like to thank my Principal Investigator, Dr. Laura Martello-Rooney, whose trust in my ideas and guidance was instrumental in turning our pilot project into a tangible reality. I am profoundly grateful to my mentors and role models, Drs. Constanza Camargo and Emily Vogtmann, whose invaluable expertise has been a guiding light at every step of my journey. A special thanks goes to former laboratory members Drs. Marzia Spagnardi, Stanley Soroka, Olalekan Lanipekun, and Jenny Paredes, for their indispensable contributions over the past four and a half years.

I extend my appreciation to the colon cancer Surgeon Dr. Henry Talus, surgical Pathologist Dr. Jianying Zeng, and resident Pathologist Dr. John Agboola. Also, to our collaborators Dr. Ji Qiu, Dr. Semir Beyaz, and Dr. Santhilal Subhash. A sincere thank you to my generous Thesis Committee, Drs. Janice Brissette (Chair), Patrick Geraghty, Christopher Roman, and John Carter. I also express my appreciation to former Dean Mark Stewart, Interim Dean David Christini, Program Director William Chirico, and the incredible School of Graduate Studies staff, Ed Throckmorton, Ellen Telesca, Denise Sheares, and Sharon Reid.

The most significant acknowledgment is reserved for my parents, Carolina and Vicente, and my brother Vicente, for the unconditional love, always nurturing my passion for science and supporting me throughout my (maybe too long) academic journey. I am also forever thankful for the camaraderie and support of my peers and amazing friends Ana Mejia-Bautista and João Moreira. Additionally, I extend heartfelt appreciation to my personal cheerleaders at every milestone: Xime, Cote, Fran, and Espi. Special recognition must be given to Ariën, who holds a special place in my heart. Your constant support and the inspiration drawn from witnessing your own achievements fuel my determination to grow as a scientist each day.

Lastly, heartfelt acknowledgment to the colon cancer patients whose resilience and bravery made this study possible.
TABLE OF CONTENTS

LIST OF TABLES ________________________________________________________________ 1
LIST OF FIGURES ______________________________________________________________ 2
LIST OF ABBREVIATIONS ________________________________________________________ 4
Abstract .................................................................................................................... 6
Introduction ............................................................................................................... 7
  I. Colorectal cancer overview ................................................................................ 7
  II. Racial/ethnic disparities in colon cancer .......................................................... 17
  III. Microbiome and colorectal carcinogenesis ..................................................... 24
  IV. The oral-gut-circulatory axis in colorectal cancer ........................................... 45
  V. Significance ....................................................................................................... 60
Specific Aims ............................................................................................................. 64
Aim 1 ......................................................................................................................... 64
  Research design and methods – Aim 1 ............................................................... 64
  Results – Aim 1 ...................................................................................................... 72
  Discussion- Aim 1 .................................................................................................. 99
Conclusions - Aim 1 ................................................................................................. 114
Aim 2 ......................................................................................................................... 115
  Research design and methods- Aim 2 ............................................................... 115
  Results – Aim 2 ...................................................................................................... 119
  Discussion – Aim 2 ............................................................................................... 136
Conclusions- Aim 2 ................................................................................................. 150
Summary .................................................................................................................. 151
  Microbiome in AA colon cancer patients .............................................................. 152
  Inflammatory mediators and antibody response in serum .................................... 153
  Gene expression analysis in tumors ..................................................................... 153
  F. nucleatum and colon cancer in AA patients ................................................... 154
References ................................................................................................................. 158
Appendix .................................................................................................................... 167
LIST OF TABLES

Table 1. Primer and probe sequences for each oral bacterial, pan bacteria 16S rRNA gene, and housekeeping gene __________________________________________________________ 69

Table 2. Patients’ characteristics________________________________________________ 73

Table 3. The list of selected bacteria for the NAPPA array___________________________________________ 91

Table 4. List of oral, gut and probiotic bacteria-NAPPA array that showed >10% seroprevalence in the samples_________________________________________________________ 93

Table 5. The genetic pathways associated with right- and left sided colon cancer from the RNA sequencing data__________________________________________________________ 121

Table 6. The genetic pathways associated with early and late colon cancer stages by RNA sequencing data__________________________________________________________ 123

Table 7. The genetic pathways associated with F. nucleatum positive and negative by RNA sequencing data__________________________________________________________ 126

Table 8. Correlation between F. nucleatum DNA amount in colon tumor tissue and density of T-cell subsets in tumor intraepithelial and stromal areas__________________________________________ 133

Table 9. Comparison of mean values of T-cell subset densities (cell count/mm²) between F. nucleatum-negative and -positive colon tumor tissues__________________________________________ 134

Table 10. Correlation between P. intermedia DNA amount in tumor tissue and density of T-cell subsets in tumor intraepithelial and stromal areas__________________________________________ 136
LIST OF FIGURES

Figure 1. Leading sites of cancer cases and deaths: 2023 Estimates___________________ 8

Figure 2. Colon cancer development pathways. CIMP: CpG island methylator phenotype; MMR: mismatch repair; MSI-H: microsatellite instability-high____________________________ 9

Figure 3. Biological differences between right-sided and left-sided colon cancers________12

Figure 4. Trends in death rates for all sites and colon and rectum cancer among Black and White people in the US for 1975-2019_________________________________________________18

Figure 5. The relationship of F. nucleatum presence in AA colon cancer patients with alterations to the gut microbiota, immune response, and tumorigenesis________________________ 62

Figure 6. Schematic overview of the primary analyses conducted in the study, the laboratory techniques employed, the range of samples included, and the collaborative efforts involved in the research________________________________________________________________ 63

Figure 7. Rarefaction curve for Taxon accumulation for each sample using R software____74

Figure 8. The relative abundance of the most frequent organisms at the phylum and genus levels in tumors when compared to non-tumor tissues_________________________________________75

Figure 9. Comparison of observed number of species and Shannon indices between different groups___________________________________________________________77

Figure 10. Differential abundance testing_________________________________________________________________________________________78

Figure 11. The relative abundance of Fusobacterium nucleatum DNA was determined using quantitative real-time PCR_________________________81

Figure 12. The relative abundance of Porphyromonas gingivalis DNA was determined using quantitative real-time PCR ____________________________82

Figure 13. The relative abundance of Prevotella intermedia DNA was determined using quantitative real-time PCR _________________________________83

Figure 15. Co-abundance of P. intermedia and P. gingivalis with F. nucleatum-high, -low, and -negative relative amount________________________________________85

Figure 16. Analysis of microbial relative abundance in genus-level between F. nucleatum presence (Fn-positive) and absence (Fn-negative) in tumors__________________________87

Figure 17. Comparative analysis to determine the differences in the “Oral pathogens” cluster and specific taxa of oral pathogens__________________________________________88

Figure 18. The relative abundance of specific taxa between Fn-positive and Fn-negative groups________________________________________________________89
Figure 19. NAPPA processing steps

Figure 20. Heatmaps of IgG antibodies on bacteria-NAPPA (with >10% seroprevalence) by location and cancer stage in 44 colon cancer patients

Figure 21. Seropositivity of IgG antibody response per bacterium by antigens between right- and left-sided location and colon cancer stage

Figure 22. Seropositivity of IgG antibody response against individual and clustered antigens of CRC-associated bacteria

Figure 23. Seropositivity of IgG antibody response against individual and clustered antigens of oral bacteria

Figure 24. Diagram for positive/negative detection of F. nucleatum and P. gingivalis by qPCR in non-tumor and tumor tissues and by IgG and IgA antibody response by each patient

Figure 25. Volcano plots for right-sided versus left-sided tumor location of colon cancer for RNA sequencing data

Figure 26. Volcano plots for colon cancer late versus early cancer stages for RNA sequencing data

Figure 27. Volcano plots for Fusobacterium nucleatum positive and negative presence in tumors for RNA sequencing data

Figure 28. Slide based multiplex arrays

Figure 29. Heatmap of serological levels of inflammatory mediators of each colon cancer patient including clinicopathological variables and absence/presence of oral pathogens

Figure 30. Summary of functionality and role in CRC of CRP, MMP-9, MMP-13, Osteopontin, and Osteoactivin

Figure 31. T cells densities analysis in digitalized whole colon tissue sections using QuPath Software

Figure 32. Immunohistochemistry staining for CD3+, CD4+, and CD8+ T cells for representative F. nucleatum-positive and -negative tumors. Red arrows indicate specific T cell subsets present in each tissue
LIST OF ABBREVIATIONS

Colorectal cancer: CRC
African Americans: AA
Caucasian Americans: CA
European Americans: EA
Fusobacterium nucleatum: Fn
The Cancer Genome Atlas: TCGA
Chromosomal instability: CIN
DNA mismatch repair: MMR
Microsatellite instability: MSI
Tumor microenvironment: TME
Consensus molecular subtypes: CMS
C-reactive protein: CRP
Interleukin: IL
Antigen presenting cells: APCs
Natural killer cells: NK
T helper cells: Th
T regulatory cells: Tregs
Dendritic cells: DC
Lipopolysaccharides: LPS
Toll-like receptors: TLRs
Reactive oxygen species: ROS
Interferon Gamma: IFN-γ
Tumor microenvironment: TME
Myeloid-derived suppressor cell: MDSC
Nuclear factor kappa-light-chain-enhancer of activated B cells: NF-kB
Mitogen-activated protein kinase: MAPK
Enterotoxigenic Bacteroides fragilis: ETBF
National Cancer Institute: NCI
Cycle threshold: Ct
Tumor-infiltrating lymphocytes: TILs
Ribosomal RNA: rRNA
Nucleic Acid Programmable Protein Array: NAPPA
Microsatellite instability-high: MSI-H
Programmed death-1: PD-1
RNA sequencing: RNAseq
Immunohistochemistry: IHC
Not Significant: NS
Abstract

Background: Colorectal cancer (CRC) is the third leading cause of cancer deaths in the US, with African American (AA) patients having the highest incidence and mortality rate. Periodontitis, a chronic oral inflammatory condition, increases the risk of CRC. *Fusobacterium nucleatum* (*Fn*), a periodontal pathogen, has been implicated in CRC pathogenesis by altering gut microbiota composition and promoting inflammation. Our study explored the complex interplay between the immune response and the microbiome as it relates to colon cancer in AA patients.

Methods: Our study included qPCR analysis, 16S rRNA gene sequencing, and RNA sequencing of tumor vs. non-tumor tissues, analysis of inflammatory mediators by using a quantitative slide-based array and antibody responses to specific bacteria by novel protein microarray NAPPA in serum, and immunohistochemistry staining to quantify T cell subset densities in tumors.

Results: Our analysis revealed a distinct microbial profile in colon tumor tissues, characterized by reduced microbial diversity and increased oral-origin bacterial abundance. Right-sided tumors revealed distinct inflammation and immune responses, in contrast to metabolic and signaling expression pathways in left tumors. *Fn* was significantly increased in tumors (p= 0.0003) and was associated with gut microbiota changes by co-aggregating with other oral pathogens and reducing probiotic and CRC-associated bacteria. Upregulated pathways in *Fn*-positive tumors were neutrophil chemotaxis, NF-κB, TLR signaling and Wnt signaling, and T cell differentiation with increased CD8+ T cell density within the tumor (p=0.036). Additionally, patients with early-stage and right-sided tumors exhibited higher anti-*Fn* seropositivity (p< 0.00001).

Conclusions: Our study reveals unique microbial and immunological profiles in AAs with colon cancer. These findings offer insights into genetic pathways and possible microbial contributions across different colon tumor locations and cancer stages, highlighting the need for further investigation into underlying mechanisms and potential new therapeutic targets.
Introduction

I. Colorectal cancer overview

Colorectal cancer (CRC) is the third most common cancer and the third leading cause of cancer deaths in males and females in the United States (US) (Figure 1)(1). In 2022, an estimated 106,180 cases of colon cancer and 44,850 cases of rectal cancer were diagnosed in the US, and a total of 52,580 people will be expected to die from these cancers (1). Even though incidence and mortality rates have decreased by about 2% per year in adults 50 and older, an increase has been shown in younger individuals, a trend that began in the mid-1990s for unknown reasons (1). Due to the increased proportion of individuals with CRC under 50 years, the American Cancer Society and the US Preventive Services Task Force began to modify recommendations for CRC screening commencement to 45 years (2).

More than half of CRCs in the US are attributable to potentially modifiable risk factors, including excess body weight, physical inactivity, smoking, heavy alcohol consumption, and diet habits such as high consumption of meat, and very low intake of fruits, vegetables, and whole-grain fiber. Also, hereditary/genetic conditions such as personal or family history of CRC or adenomatous polyps, certain inherited genetic syndromes (e.g., Lynch syndrome), a personal history of chronic inflammatory bowel disease (ulcerative colitis or Crohn’s disease), and type 2 diabetes increase the risk of CRC (1).
Stage defines the extent or spread of cancer at diagnosis or during treatment, crucial for treatment planning and prognosis evaluation. It relies on primary tumor size or spread and the presence of metastasis in nearby lymph nodes or distant body sites. Clinicians predominantly use the TNM staging system, evaluating Tumor size, Nodal involvement, and Metastasis. TNM evaluates cancer growth and spread and assigns a stage from 0 (in situ) for the earliest stage up to I, II, III, or IV for more advanced metastatic disease (4).

Surgery is the most common treatment when cancer has not spread to distant sites. In metastatic CRC, treatments typically include chemotherapy and/or targeted therapy. Immunotherapy is a newer option that can be highly effective for advanced cancer stage where by activating or suppressing the immune system for immune cells to recognize and attack cancer cells (1).

Several genome-wide association studies have successfully identified susceptibility genes (common single-nucleotide polymorphisms) that are associated with CRC risk. There are three...
major molecular pathways that relate to CRC development: 1) the most predominant is chromosomal instability (CIN); 2) the second most predominant is CpG island methylator phenotype (CIMP), which includes sporadic microsatellite instability (MSI) high cancers; 3) and finally the pure MSI pathway resulting from germline mutation in a DNA mismatch repair (MMR) gene(s). Hereditary nonpolyposis colorectal cancer (HNPCC) develops via the pure MSI pathway (5) (Figure 2). These pathways characterize multiple genetic and epigenetic events in sequential order:

- The adenoma–carcinoma pathway or CIN leads to 70–90% of CRCs. Usually, progresses following genomic events initiated by an APC mutation, followed by RAS activation or loss of function of TP53.

- The serrated neoplasia pathway, represents 10–20% of CRCs. This pathway is associated with RAS and RAF mutations, and epigenetic instability, characterized by CIMP, leading to microsatellite stable (MSS) and MSI cancers. MSI is the result of defective DNA repair through inactivation of MMR genes which is also seen in Lynch syndrome (2).

Figure 2. Colon cancer development pathways. CIMP: CpG island methylator phenotype; MMR: mismatch repair; MSI-H: microsatellite instability-high. (Adapted from Dekker et al(2)).

However, although the genetic mutations driving CRC are well-documented, comprehending the mutational landscape of tumor cells alone is inadequate for stratifying tumor
subtypes or precisely predicting patient prognoses. Instead, the progression of colon carcinogenesis depends not only on key mutations but also on the complex interactions between tumor cells and their surrounding tumor microenvironment (TME). The TME includes three primary cell types: cancer-associated fibroblasts, vascular cells, and infiltrating immune cells. These diverse cell types interact closely in a reciprocal manner, exerting both positive and negative influences on tumor cell proliferation, apoptosis, evasion of growth suppression, energy metabolism, immune evasion, angiogenesis, and tumor cell invasion in a non-autonomous manner (6).

Within the TME of colon cancer, a variety of immune cell types play critical roles in influencing cancer initiation and progression. These immune cells encompass neutrophils, mast cells, natural killer (NK) cells, dendritic cells (DCs), tumor-associated macrophages, and a specific subset known as myeloid-derived suppressor cells (MDSCs). Their precise mechanisms of action in impacting tumor growth remain incompletely understood. T cells, crucial for both inflammation and anticancer immune responses, hold particular significance. In sporadic CRC, a delicate equilibrium exists between immune surveillance, mediated by cytotoxic CD8 T cells and NK cells, and tumor-promoting inflammation, orchestrated by macrophages, DCs, and other T cell subtypes (6,7).

The clinical significance of these stromal cells in CRC is exemplified in two instances. Firstly, the presence of tumor-infiltrating lymphocytes (TILs) correlates with improved prognosis in advanced CRC cases. Cytotoxic CD8 T cells directly target cancer cells by inducing apoptosis and necrosis. Studies have demonstrated that T helper 1 (Th1) cell-mediated immune responses and levels of IFN-γ within CRC tumors are associated with favorable outcomes, whereas immune responses mediated by Th17 cells are linked to less favorable prognoses (8). This has led to the development of an ‘immunoscore’ for CRC tumors based on the quantification of CD3+ and CD8+ lymphocytes in the core tumor and invasive margin (9). However, despite the acknowledged
significance of T cell infiltration in predicting CRC patient prognosis, the reasons why certain
tumors exhibit greater susceptibility to immune checkpoint blockade remains unknown.

A second example highlighting the importance of the TME comes from recent efforts to
classify CRC into consensus molecular subtypes (CMS). This classification, based on gene
expression profiles from over 4,000 CRC samples, has identified four distinct subtypes labeled
as CMS1 through CMS4. This system groups colon tumors into distinct subtypes based on their
molecular features with the idea that these subtypes may have different biological behaviors and
may respond differently to treatment. The specific molecular features used to classify tumors into
CMS subtypes can vary depending on the cancer type. In the case of colon cancer classification
is based on the presence of specific mutations, the levels of MSI, and the mutational burden (the
number of mutations present in the genome of the tumor). Based on these features, colon cancers
can be classified: CMS1, enriched for inflammatory/immune genes; CMS2, canonical; CMS3,
metabolic; and CMS4, mesenchymal (10).

Differentiating colon cancer location by right (proximal to the splenic flexure) and left
(distal) is important for several reasons. Emerging evidence suggests that right- and left-sided
colon cancers may have unique molecular characteristics and responses to treatment,
highlighting the importance of site-specific management strategies (2). A study (11) analyzed data
from The Cancer Genome Atlas (TCGA) and found that right-sided colon cancers had significantly
lower rates of mutations in the KRAS and TP53 genes and higher rates of mutations in the BRAF
gene compared to left-sided colon cancers. They also found that right-sided tumors were more
likely to be classified as CMS1 (a subtype characterized by high microsatellite instability and high
mutational burden) and left-sided colon cancers were more likely to be classified as CMS3 (a
subtype characterized by low microsatellite instability and low mutational burden). The authors
concluded that there are significant biological differences between right- and left-sided colon
cancers and that these differences may have implications for treatment (Figure 3)(11). Another
study (12) found a gradient in the prevalence of certain molecular features (e.g. TP53 mutations)
across the different tumor locations (such as sigmoid colon to cecum), suggesting a continuum in the molecular characteristics of CRCs. The authors concluded that classifying CRCs by location rather than sidedness may provide a more accurate depiction of the molecular heterogeneity of these tumors and may have implications for treatment decisions (12).

Figure 3. Biological differences between right-sided and left-sided colon cancers. Adapted from Dekker et al (2).

Beyond molecular disparities, the right and left colon exhibit variances in embryological origins, biological characteristics, and anatomical structures. The right-sided colon encompasses the cecum, ascending colon, hepatic flexure, and a majority of the transverse colon, while the left-sided colon comprises the splenic flexure, descending colon, sigmoid, and rectosigmoid junction. These anatomical distinctions are fundamental, as they influence blood supplies and lymphatic drainage patterns, which are pivotal factors in determining accurate diagnosis and formulating effective treatment strategies (Figure 3) (2). Rectal cancer is often examined separately from colon cancer since localized stage II–III rectal cancers have a different treatment paradigm than colon cancers (2).
Studies have shown that patients with right-sided colon cancer may have worse prognoses and survival rates compared to those with left-sided colon cancer. Identifying markers that can aid or complement colonoscopy in detecting right-sided colon cancer holds immense significance for several reasons. Firstly, colonoscopy, while effective in reducing the incidence and mortality of left-sided CRC, has shown limitations in detecting and preventing right-sided CRC to the same degree. Gupta et al (13) found that most right-sided polyps with high-grade dysplasia, adenocarcinoma, or any advanced neoplasia were ≤9 mm, while most left-sided were >9 mm. Secondly, right-sided colon cancers tend to be diagnosed at more advanced stages compared to left-sided cancers, which can significantly impact treatment outcomes and prognosis. The delayed detection often leads to a poorer prognosis due to the cancer being more advanced at the time of diagnosis.

Other studies have focused on the relationship between the tumor location and other factors that may affect patient outcomes, such as the relationship of the TME and systemic inflammation. Patel et al (14) analyzed a large database of patients who had undergone surgery for colon or rectal cancer and systemic inflammation was measured by analyzing the levels of certain biomarkers in the blood, such as C-reactive protein (CRP) and interleukin-6 (IL-6). The authors of the paper hypothesized that the higher density of immune cells in the TME that were associated with higher levels of systemic inflammation in patients with right-sided tumors and this may contribute to their worse clinical outcomes. Inflammation has been linked to the development and progression of cancer where chronic inflammation may promote the growth and spread of cancer cells (14), suggesting that targeting systemic inflammation may be a potential strategy for improving outcomes for patients with right-sided colon cancer.

Recent evidence highlights a complex interplay between the gut microbiome, genetic variations, and environmental factors. Variables such as host genetics, antibiotic exposure, and environmental and social determinants, including race/ethnicity, geographic location, and urbanization, are recognized as influential factors shaping the composition and function of the gut
microbiome in CRC (15). Emerging research suggests that patients with CRC exhibit dysbiosis, an imbalance in the gut microbiota (16,17). This dysbiosis can lead to chronic inflammation and the production of carcinogenic substances, thereby elevating the risk of CRC. Patients with CRC showed increased levels of *Bacteroides, Prevotella, Escherichia coli, enterotoxigenic Bacteroides fragilis, Streptococcus gallolyticus, Enterococcus faecalis, Fusobacterium nucleatum,* and *Clostridium difficile* (18).

It has been also shown that microbial communities present in left-sided and right-sided colon cancer tumors are distinct (19–22). Zhong et al (21) found that the microbial profiles of left-sided colon cancer tumors were more like those of healthy colon tissue, while right-sided tumors' microbiome was more like colon tumors in general. The authors concluded that the microbiome may be a useful tool for distinguishing left-sided and right-sided colon tumors and may have implications for the diagnosis and treatment of these tumors (21).

Recently, Dejea et al (19) described the novel link between bacterial biofilms and CRC. These colon biofilms are dense communities of bacteria encased in a likely complex matrix adhering to colon epithelial cells in right-side tumors that were found to correlate with tissue invasion and increased cell growth. Unexpectedly, individuals with biofilm-positive tumors, regardless of whether they had cancerous growths or adenomas, exhibited biofilms on their healthy mucosa far away from their tumors. These bacterial biofilms were linked to reduced E-cadherin in colonic epithelial cells and increased activation of IL-6 and Stat3. Additionally, they were associated with increased proliferation of epithelial cells in the normal colon mucosa (19). Drewes et al (20) supported these findings showing that the consistent presence of bacterial biofilm structures within the mucus layer of the gut is a distinguishing characteristic found in right-sided CRC, whereas it is not observed in left-sided CRC. This study unveils a new idea: how the structural arrangement of microbial communities might influence disease advancement, marking a significant finding in understanding disease progression. Importantly, colon mucosal biofilm detection may predict an increased risk for the development of sporadic CRC (19). These biofilms
were found to cover large areas of the colonic mucosa, detectable in even the distant normal mucosa of patients with CRC, and consisted predominantly of \textit{Bacteroidetes}, \textit{Lachnospiraceae}, \textit{Fusobacterium} spp, and the phylum \textit{Proteobacteria} (20).

Findings from both the Malaysian cohort and the broader meta-analysis (20) confirm that CRC tissues, especially in right-sided tumors, exhibit an abundance of invasive biofilms and the vast majority (>80%) of CRC cases contain aberrant microbial signatures indicative of dysbiosis. By high-resolution 16S rRNA gene meta-analysis captured species-level taxonomic assignments, this data backs various pathways through which bacteria might contribute to cancer development. These include the increase of a symbiotic bacterium with enterotoxigenic potential \textit{B. fragilis}, the growing understanding of oral microbes as potentially harmful residents in the gut, and the formation of complex pro-cancerous biofilms involving multiple microorganisms. These tissues also show an increase in oral pathogens such as \textit{F. nucleatum}, \textit{Parvimonas micra}, and \textit{Peptostreptococcus stomatis} (20). In agreement, a meta transcriptomic study showed that \textit{Fusobacterium}, \textit{Leptotrichia}, and \textit{Campylobacter} spp. commonly coexist in the examined samples (23). Pairwise correlation of these microbial sequence counts revealed that these three species exist in a cooccurrence network along with \textit{Porphyromonas} in CRC tissue (23). This occurrence isn't surprising considering they are all Gram-negative anaerobic microbes that typically occupy the same space within the oral cavity in the subgingival plaque (23).

Etiological investigations have further elucidated the role of \textit{F. nucleatum} as a bacterium that promotes carcinogenesis at various stages of CRC development. Two different groups were the first to show an increased abundance of this oral pathogen in CRC tissues when compared to normal tissues (24,25). These collective findings suggest that \textit{F. nucleatum} infection has been implicated as an added environmental risk factor for CRC by promoting a proinflammatory microenvironment conducive to the progression of colorectal neoplasia (18,26–32). Moreover, the high abundance of \textit{F. nucleatum} in CRC is associated with poorer survival (33–35) and recurrence after chemotherapy (36). A recent study highlighted that identical strains of \textit{F. nucleatum} were
identified in both the saliva and tumors of patients with CRC (35), in addition to the findings that oral \textit{F. nucleatum} translocate to the colon through the hematogenous route (37). This suggests that CRC-associated \textit{F. nucleatum} likely originates from the oral cavity, underscoring a potential link between oral bacteria and the development of CRC.

Metagenomic sequencing analysis revealed the presence of four species enriched in tumor samples in CRC patients with oral cavity origin: \textit{Porphyromonas asaccharolytica}, \textit{F. nucleatum}, \textit{Prevotella intermedia}, and \textit{Parvimonas micra}. Remarkably, these bacteria were found to form mutually beneficial networks within the microbial community where \textit{F. nucleatum} occupied a central position within this network.

Certain oral pathogens, such as \textit{F. nucleatum}, play a pivotal role in the development of periodontitis (38,39). Periodontitis is characterized as a chronic inflammatory disease initiated by a dysbiotic biofilm formation within the gingival pocket, ultimately resulting in the destruction of the tooth-supporting tissues. Periodontitis increases risk of CRC incidence and mortality (40,41). The Nurses’ Health Study revealed that periodontitis, including tooth loss can increase systemic inflammation and cause immune dysregulation which may subsequently alter gut microbiota, thereby influencing colorectal carcinogens (40). They observed a slight rise in the likelihood of proximal (right-sided) colon cancer and rectal cancer among women with a lower number of teeth (40).

Another study investigated the link between periodontal disease, tooth loss, and the development of colorectal precursor lesions using the Nurses’ Health Study (1992-2002) and the Health Professionals Follow-up Study (1992-2010), involving 17,904 women and 24,582 men. The study found that individuals with periodontal disease had a 17\% higher risk of serrated polyps and an 11\% higher risk of conventional adenomas. Additionally, those who lost $\geq 4$ teeth had a 20\% greater risk of serrated polyps. These associations remained significant even among never smokers (42). Other studies (43)(44) also observed an increased risk for CRC adenomas among individuals with periodontitis, especially proximal advanced adenomas. The findings suggest that
a history of periodontal disease and significant tooth loss may slightly elevate the risk of developing colorectal precursor lesions, contributing to our understanding of the intricate relationship between oral health, the microbiome, and early CRC carcinogenesis.

II. Racial/ethnic disparities in colon cancer

There is abundant evidence that suggests that there are racial disparities in the incidence and outcomes of colon cancer. AA have a higher incidence of colon cancer compared to other racial/ethnic groups, and they are also more likely to be diagnosed at later stages of the disease. The non-Hispanic Black population is the third-largest racial/ethnic group following non-Hispanic white and Hispanic people. The population includes ancestors of individuals brought to the US as slaves who largely identify as AA, as well as nearly 10% who are recent immigrants, mostly of African or Caribbean descent (45,46). Structural racism is the underlying source of health disparities among people of color, which is discrimination perpetuated through interconnected institutions and reinforced through culture, history, ideology, and sanctioned practices. These impact all components of life to limit the accumulation of wealth and overall standard of living through unequal access to work, education, housing, healthy food, and quality health care (47). In addition, studies have shown that AA people have greater mistrust in the medical system and biomedical research process compared to Caucasian American (CA)/European American (EA) individuals (48). The terms Caucasian American (CA) and European American (EA) are often used interchangeably and generally refer to individuals of European descent residing in the US. Both terms are used to describe the same racial group, but White is preferred now.

AAs with colon cancer are also more likely to have worse clinical outcomes, including a higher risk of cancer recurrence and death, compared to other racial groups. AAs have the second-highest incidence of CRC in the US, following the Alaska Native/American Indian population. From 2009 to 2018, incidence rates for CRC declined by about 3% per year among AAs versus 2% among EAs. Like incidence rates, mortality rates are historically higher in AAs (Figure 4) (49).
Racial disparities in stage-specific survival essentially reveal inequalities in treatment, and differences in comorbidities and tumor characteristics (49–52).

**Figure 4.** Trends in death rates for all sites and colon and rectum cancer among Black and White people in the US for 1975-2019 (49).

Jovov et al (51) analyzed the gene expression between tumor and non tumor tissues from the same individuals of AA and EA CRC patients. They found that 95 genes were expressed differently in AAs and EAs where three of the six pathways affected are related to inflammatory and immune responses. Seven of the nine genes associated with antigen presentation and inflammatory response were up-regulated in AA patients (HLA-DQB1, IL33, PAK2, PROKR1, SAA2, TLR4, ZNF234), and two genes were down-regulated (DHX58, IL27). From the expression of these 95 genes, a group of 10 genes could in 94% of cases accurately predict from which racial group the tumor tissue was obtained. These results suggest that there may be differences in the inflammation and immune responses to CRC in AAs and EAs, which could contribute to the disparities in incidence and mortality of the disease between these groups and these findings could also have implications for cancer immunotherapy (51).
Accumulating evidence suggests that the gut microbiota is an important component to consider when it comes to CRC (18,53,54) and may serve as an effective and potentially modifiable molecular mechanism underlying racial and ethnic disparities (55). Moreover, the gut microbiome interacts with environmental factors, potentially triggering inflammation and influencing genetic alterations that contribute to the colon carcinogenesis.

The examination of the gut microbiota from the stool of 2084 participants from the Healthy Living in an Urban Environment Study revealed a trend: individuals residing in the same city often exhibited comparable gut microbiota profiles to others from their racial or ethnic background (56). This suggests that where people live can influence the similarity of their gut microbiota with others from the same racial or ethnic background. Despite some connections to factors like lifestyle and diet, differences in gut microbial diversity seem more related to racial or ethnic identity rather than solely being attributed to these other elements. Therefore, when studying the gut microbiome in relation to CRC, understanding and considering an individual's racial or ethnic background could be crucial. Our patient population lives in neighboring areas in Brooklyn and are predominantly African descendants.

Several studies have described race-associated variations in microbial abundance, specifically in the colon (18,52,57). But importantly, there have been a limited number of studies that have collected and analyzed colon tissues from CRC patients and healthy controls, especially in minority groups. Fecal samples are the most common targeted sample, even when evidence indicates that stool microbiota only partially reflects mucosal microbiota in CRC.

When analyzing gut microbiota in disparity groups, AA cancer patients had decreased microbial diversity compared to CA patients (52). Higher levels of Bacteroides have been found in AA compared to CA patients. Pro-inflammatory bacteria such as F. nucleatum and Enterobacter species were significantly higher in AA patients (58). Bacteroides, Prevotella, E. coli, ETBF, S. gallolyticus, E. faecalis, and Clostridium difficile have been found in patients with CRC (18).
Hester et al. (59) compared bacteria in stool samples of AA and CA in a small pilot study. They found lower acetate, butyrate, total SCFAs content, and a higher pH in AA compared to the other racial groups. Acetate, butyrate, and total short-chain fatty acids (SCFAs) play crucial roles in maintaining colon health. These SCFAs are produced through the fermentation of dietary fibers by gut microbiota in the colon. They serve as an energy source for colonocytes, contributing to the maintenance of epithelial integrity and barrier function in the colon (59). Butyrate has beneficial effects in reducing colon cancer risk with anti-inflammatory, immunomodulatory effects and downregulating Wnt signaling, which inhibits cell proliferation and migration. Wnt/B-catenin signaling plays a fundamental role in several biological processes such as development and cell proliferation related to tumorigenesis (59).

In addition, AA had higher levels of Firmicutes bacteria compared to CA and Hispanics. Moreover, the ratio of Firmicutes compared to Bacteriodes, which has been associated with obesity, was higher in AA (59). Additionally, AA patients with CRC had higher concentrations of sulfidogenic bacteria compared to healthy EAs and AAs without CRC (60). Sulfidogenic bacteria, such as Fusobacterium, Desulfovibrio and Bilophila wadsworthia, have been implicated in CRC development through the production of hydrogen sulfide (61). Hydrogen sulfide is a genotoxic compound that has been shown to damage DNA leading to genomic or chromosomal instability (CIN), effecting DNA repair in a multistep carcinogenic process. One hypothesis is that hydrogen sulfide diffuses into intestinal epithelial cells and interferes with mitochondrial function, ultimately leading to hyperproliferation via the Ras/MAPK pathway (62). The hyperactivation of the Ras/MAPK pathway is a known mechanism of carcinogenesis in CRC.

The difference in tumor immunology in colon cancer between AA and CA patients remains an area of active research. Several studies have suggested variations in the TME, immune cell infiltration, and immune checkpoint expression between these patient populations. A previous study done in our lab by Dr. Jenny Paredes (63) to examine differences in tumor immunology in tumor and non-tumor tissue samples from AA (15 tumors and 15 adjacent non-tumor tissues) and
CA (18 tumors and 17 adjacent non-tumor tissues) patients, showed that AA colon tumors had higher levels of gene expression for specific genes involved in inflammation and immune suppression.

When examining the cell subsets with potential cytotoxic activity, AA tumors had significantly higher scores of CD8+ T cells ($p = 0.05$) and natural killer cells ($p = 0.04$). AA tumors also had a significant increase in scores of exhausted CD8+ cells ($p = 0.01$), natural killer CD56dim cells ($p = 0.04$), and Th1 cells ($p = 0.05$). The latter set of data suggests that tumors from AAs present significant numbers of cytotoxic natural killer cells and their reduced cytotoxic activity is specific for CD8+ T cells. The presence of exhausted CD8+ T cells suggest that T cells are recruited to the tumor site but lack activation as indicated by the gene expression of biomarkers indicative of T cell exhaustion such as CD8, CD244, EOMES, and LAG3. These differences in tumor immunology may contribute to variations in treatment response and prognosis between AA and EA patients with colon cancer. Further research is needed to elucidate the underlying mechanisms driving these disparities and to develop targeted immunotherapeutic strategies that account for the unique tumor immunology profiles observed in different patient populations.

Additionally, potential differences between circulating levels of cytokines between AA and CA colon cancer patients were evaluated by multi-analyte ELISA assays (63). AA and CA colon cancer patients secreted significantly different types and levels of cytokines. For the correlation of the expression of cytokines to T cell subsets, they used the following markers: Th1 (IL-2, IL-12, INF-γ, TNF-α), Th2 (IL-4, IL-5, IL-10), Th17 (IL-17A, IL-6), and Tregs (T-Reg: IL-17A and TGF-β1). CAs had significantly higher level of cytokines correlated to Th1 and Th2 activation, differentiation, and anti-tumor activity. However, they did not present differential expression of the pro-inflammatory cytokines; interleukin 17-A (IL-17A) and TNF-α when compared to AA patients. Together, these results suggest that AA patients may fail to promote the anti-tumor activity of the Th1 and Th2 subsets of T cells and cytokines associated with the presence of Th17 and inflammation. All these results indicate that AA colon tumors gene expression is associated with
an impaired immune response and diminished antitumor activity when compared to the tumors from CA colon cancer patients. This impairment was suggested by the downregulation of several markers for cytotoxic T cells as well as upregulation of multiple markers associated with inflammation and the presence of exhausted CD8+ and Tregs (63).

Systemic inflammation, a common factor in both periodontitis and colon cancer, may contribute to the development and progression of tumors in the colon. Periodontitis has been linked to an increased risk of various systemic diseases, including colon cancer (64) and is among the 10 most prevalent chronic diseases affecting the world’s population (65). The CDC included a discussion of health disparities and inequalities within periodontal disease prevalence in the US. The report shows significant disparities exist in the prevalence of periodontitis by race/ethnicity, education, and poverty level (66).

The prevalence of periodontitis in Non-Hispanic Blacks is 59.1% as compared to 40.8% in Non-Hispanic whites (67). Periodontal pathogens, *F. nucleatum* and *P. gingivalis* have been found to be significantly higher in AA compared with CA patients (15,57,58,60,68). In AA patients, disparities in access to dental care and higher rates of periodontitis may exacerbate the risk of colon cancer. The majority of studies examining the link between CRC and periodontitis lack information on the racial background of the participants. Among those studies specifying the racial demographic (40,43), over 90% of the patients are identified as White. Notably, there is a significant gap in research investigating the connection between periodontitis and colon cancer specifically in AA patients.

There is a predominance of right sided tumors in AAs of up to 30% followed by sigmoid colon 25%. Moreover, studies (43)(44) showed an increased risk for CRC adenomas among individuals with periodontitis, especially proximal advanced adenomas. Understanding the shared characteristics between right-sided colon cancers and periodontitis in this demographic could shed light on this cancer location and lead to improved outcomes. Molecular profiling and
alternative diagnostic methods present promising avenues for personalized treatment and early detection in AA patients.
III. Microbiome and colorectal carcinogenesis

Microbiome and colorectal carcinogenesis: Linked mechanisms and racial differences

Sofia C Tortora, Vimal M Bodiwala, Andrew Quinn, Laura A Martello, Shivakumar Vignesh

ORCID number: Sofia C Tortora 0000-0001-4028-7463; Vimal M Bodiwala 0000-0002-0156-3126; Andrew Quinn 0000-0001-8252-2004; Laura A Martello 0000-0002-6259-5757; Shivakumar Vignesh 0000-0001-5190-8222.

Author contributions: Tortora SC, Bodiwala VM and Quinn A wrote the manuscript; Tortora SC wrote Table; Martello LA and Vignesh S edited and added to manuscript; all authors have read and approve the final manuscript.

Conflict-of-interest statement: Authors declare no conflict of interests for this article.

Country/Territory of origin: United States

Specialty type: Gastroenterology and hepatology

Provenance and peer review: Unselected article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): 0
Grade D (Fair): 0
Grade E (Poor): 0

Abstract

Various studies have shown the interplay between the intestinal microbiome, environmental factors, and genetic changes in colorectal cancer (CRC) development. In this review, we highlight the various gut and oral microbiota associated with CRC and colorectal adenomas, and their proposed molecular mechanisms in relation to the processes of “the hallmarks of cancer”, and differences in microbial diversity and abundance between race/ethnicity. Patients with CRC showed increased levels of Bacteroides, Prevotella, Escherichia coli, enterotoxigenic Bacteroides fragilis, Streptococcus gallolyticus, Enterococcus faecalis, Fusobacterium nucleatum (F. nucleatum) and Clostridium difficile. Higher levels of Bacteroides have been found in African American (AA) compared to Caucasian American (CA) patients. Pro-inflammatory bacteria such as F. nucleatum and Enterobacter species were significantly higher in AAs. Also, AA patients have been shown to have decreased microbial diversity compared to CA patients. Some studies have shown that using microbiome profiles in conjunction with certain risk factors such as age, race and body mass index may help predict healthy colon vs one with adenomas or carcinomas. Periodontitis is one of the most common bacterial infections in humans and is more prevalent in Non-Hispanic Blacks as compared to Non-Hispanic Whites. This condition causes increased systemic inflammation, immune dysregulation, gut microbiota dysbiosis and thereby possibly influencing colorectal carcinogenesis. Periodontal-associated bacteria such as Fusobacterium, Prevotella, Bacteroides and Porphyromonas have been found in CRC tissues and in feces of CRC patients. Therefore, a deeper understanding of the association between oral and gastrointestinal bacterial profile, in addition to identifying prevalent bacteria in patients with CRC and the differences observed in ethnicity/race, may play a pivotal role in predicting incidence, prognosis, and lead to the development of new treatments.
INTRODUCTION

Colorectal cancer (CRC) is the third most commonly diagnosed cancer and the third leading cause of cancer death[1] with higher incidence and mortality rates for African Americans (AA) among general population[2]. The incidence rates are 24% higher in Non-Hispanic (NH) black males and 19% higher in NH black females, while the death rates are 47% higher in NH black men and 34% higher in NH black women compared to NH white men and women[3]. This significant impact underscores the importance of further understanding the mechanisms and factors that influence the progression of CRC so as to alter the disease process before it has progressed to cancer or improve outcomes in those with a CRC diagnosis. The microbiome and its potential role in colorectal carcinogenesis is a newly emerging yet important field of study.

Current evidence indicates a complex interplay between the gut microbiome, genetic alterations and environmental factors. The gut microbiome is highly diverse and compared to the human genome it contains approximately 100 times the number of genes[4]. There are vast differences found in the microbiota of apparently healthy individuals. Variables such as differences in host genetics, antibiotics usage, environmental and lifestyle factors, including ethnicity, geographic location, and urbanization may modify the gut microbiome[5]. The epithelial surface and mucus layer are enriched with Clostridium, Lactobacillus and Enterococcus and the intestinal lumen is enriched with Bacteroides, Bifidobacterium, Streptococcus, Enterobacteriaceae, Enterococcus, Clostridium and Lactobacillus[6]. Certain microorganisms could protect against pathogens that promote carcinogenesis by competing for attachment sites and thus preventing the development of CRC[7]. Probiotic strains of Lactobacillus and Bifidobacterium are thought to play a protective role by competing for adhesion sites, secreting antibacterial peptides and displacing the enteropathogens, Salmonella typhimurium and Escherichia coli (E. coli), as demonstrated in an enterocyte-like Caco-2 cell layer[8]. In CRC patients, a significantly reduced level of Bifidobacterium has been found[7]. In contrast, several studies have reported increased levels of Bacteroides, Prevotella, E. coli, enterotoxigenic Bacteroides fragilis (B. fragilis), Streptococcus galaloyticus (S. galaloyticus), Enterococcus faecalis (E. faecalis), Fusobacterium nucleatum (F. nucleatum) and Clostridium difficile in CRC subjects[9-15]. Additional human studies also have confirmed that F. nucleatum is associated with other Gram-negative bacteria, such as Streptococcus, Campylobacter and Leptotrichia, and synergistically promotes CRC[14].
Recent epidemiological studies have examined the association between periodontal diseases and CRC risk. Periodontitis is one of the most common bacterial infections in humans. In periodontitis, pathogenic opportunistic microorganisms in the oral cavity damage the integrity of the tooth-supporting tissues causing increased systemic inflammation, immune dysregulation, gut microbiota dysbiosis and thereby possibly influencing colorectal carcinogenesis[16,17]. *F. nucleatum* is one of the most prevalent species found in extra-oral sites. This bacterium regulates biofilm organization and interacts with the host cells by producing various adhesins and associates with other bacteria through cross-feeding and metabolic interactions[18]. As such, *F. nucleatum* has been suggested to be a “driver bacterium” with pro-carcinogenic characteristics that contribute to tumor development by facilitating “passenger bacteria” to continue the progression of CRC[19]. Periodontal-associated bacteria such as *Fusobacterium, Prevotella*, and *Bacteroides* have been found in CRC tissues[9] and *Fusobacterium* and *Porphyromonas* in feces of CRC patients[19,20]. Half of Americans age 30 or older have periodontitis and this increases to 70% for adults aged 65 years and older[7]. Moreover, the prevalence in NH blacks is 59.1% as compared to 40.8% in NH whites[7].

Race/ethnicity also has been associated with variations in microbial abundance. Analysis of the gut microbiota by 16S in 1673 participants in the United States reported 12 microbial genera and families that vary by race/ethnicity. This suggests that the gut microbiota could be inherited and associated with human genetic variation[21]. Farhana et al[22] analyzed microbial communities in colonic effluents using 16S rRNA profiling from AA and Caucasian American (CA) patients scheduled for an outpatient screening colonoscopy. The results showed higher levels of *Bacteroides* in AAs compared to CAs. Pro-inflammatory bacteria such as *F. nucleatum* and *Enterobacter* species were significantly higher in AAs. Also, AA patients had decreased microbial diversity compared to CA patients[22]. A study conducted in Malaysia reported that *Porvimonas microa*, *F. nucleatum*, *Peptostreptococcus stomatis* and *Akkermansia muciniphila* were enriched in colon tissue of CRC patients[23]. Another study described four periodontal pathogens, *Porphyromonas gingivalis*, *Prevotella intermedia*, *Treponema denticola*, and *F. nucleatum*, to be more prevalent among AAs than among CAS[24]. Significant differences in oral and gut microbial diversity and abundance between AA and CA may play a role in the CRC disparity observed between the populations.

The “Hallmarks of Cancer,” proposed by Hanahan and Weinburg[25] organize the fundamental processes involved in the complex progression from normal cell to a tumorigenic state within a neoplastic environment. During carcinogenesis, various changes affect the host environment processes such as cellular metabolism and immunological function. Several intestinal microbes may influence the initiation and progression of tumorigenesis by modulating host factors that comprise the hallmarks of cancer[26]. This influence may occur locally and distantly through infection and microbial products, by changing the metabolism of the products produced by host and microbes or by modulating tumor immunosurveillance, which in turn alters the balance between the rate of cell proliferation and apoptosis, triggering chronic inflammation and immunosuppression[27].

Here, we summarize the current knowledge on how specific members of the microbiota and their bacterial machinery influence the hallmarks of cancer in the initiation and progression of CRC (Table 1). Each listed bacterial species is systematically presented according to the following scheme: (1) Tumor-promoting inflammation; (2) Avoiding immune destruction; (3) Deregulating cellular energetics; (4) Sustaining proliferative signaling; (5) Inducing angiogenesis; (6) Resisting cell death; and (7) Genome instability and mutations.

**TUMOR-PROMOTING INFLAMMATION**

A progressive interplay between tissue cells, microbiota and immune cells has been described. T and B cells present in the intestinal mucosa have location-specific phenotypes and functions that can be modified by the microbiota[28]. Commensal microbiota can modulate innate immune cells to release pro-inflammatory cytokines such as interleukin 6 (IL-6), IL-23 and IL-1β, which stimulate the expansion of T-helper-17 (Th17) cells. Th17 cells are a unique CD4+ T-helper subset that secrete the cytokine IL-17. IL-17 increases Paneth cell production of anti-microbial peptides and promotes inflammation by the recruitment of polymorphonuclear neutrophils from the bloodstream[29]. Th17 have pro-tumorigenic effects and it has been associated with worse prognosis in CRC, as shown in Figure 1[30]. While Th17 immune
<table>
<thead>
<tr>
<th>Gut bacteria</th>
<th>Bacterial machinery</th>
<th>Hallmarks of cancer including enabling characteristics affected</th>
<th>Outcome</th>
<th>Methods</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Fusobacterium nucleatum</em></td>
<td>FadA</td>
<td>Tumor-promoting inflammation</td>
<td>Expression of NF-κB and pro-inflammatory cytokines IL-6, 8, and 18</td>
<td>HCT116 cells (expressing E-cadherin)</td>
<td>[56, 57]</td>
</tr>
<tr>
<td>Unknown</td>
<td>Unknown</td>
<td>Tumor-promoting inflammation</td>
<td>Infiltration of specific myeloid cell subsets and an NF-κB proinflammatory signature (shared with human CRC tissue with a high Fusobacterium abundance)</td>
<td>Apc(Min/+) mice fed <em>F. nucleatum</em></td>
<td>[39]</td>
</tr>
<tr>
<td>Unknown</td>
<td>Unknown</td>
<td>Tumor-promoting inflammation</td>
<td>TNF-α and IL-10 abundance</td>
<td>Rectal biopsies of adenoma cases compared to controls</td>
<td>[59]</td>
</tr>
<tr>
<td>Unknown</td>
<td>Unknown</td>
<td>Tumor-promoting inflammation</td>
<td>Cytokine production, which is mediated by the p38 MAPK signaling but independent of TLRs, NOD-1, NOD-2 and NF-κB signaling</td>
<td>HEK293T cells, (which lack endogenous TLRs)</td>
<td>[58]</td>
</tr>
<tr>
<td>Fap2</td>
<td>Avoiding immune destruction</td>
<td>Fap2 interacted with TIGIT, leading to the inhibition of NK cell cytotoxicity</td>
<td></td>
<td>Various BW cells</td>
<td>[31]</td>
</tr>
<tr>
<td>Unknown</td>
<td>Unknown</td>
<td>Avoiding immune destruction</td>
<td>MicroRNA-21 increases the levels of IL-10 and prostaglandin E2, which suppress antitumor T-cell-mediated adaptive immunity</td>
<td>Colorectal carcinoma tissues (stages I-IV) from Japanese patients</td>
<td>[66]</td>
</tr>
<tr>
<td>Generation of formyl-methionyl-leucyl-phenylalanine and SCFAs from dietary amino acids</td>
<td>Deregulating cellular energetics</td>
<td>Chemoattract myeloid cells</td>
<td></td>
<td>ApcMin/+ mouse model of intestinal tumorigenisation</td>
<td>[39]</td>
</tr>
<tr>
<td>Adhesin FadA</td>
<td>Sustaining proliferative signaling</td>
<td>FadA binds to E-cadherin and activates β-catenin signaling</td>
<td>HCT116 cells (expressing E-cadherin)</td>
<td></td>
<td>[56]</td>
</tr>
<tr>
<td>Unknown</td>
<td>Genome instability and mutations</td>
<td>CpG island methylating phenotype (CIMP), microsatellite instability (MSI), and MLH1 hypermethylation</td>
<td></td>
<td>Colorectal carcinoma tissue</td>
<td>[82]</td>
</tr>
<tr>
<td><em>Streptococcus galloyticus</em> (<em>S. bovis</em>)</td>
<td>Unknown</td>
<td>Tumor-promoting inflammation</td>
<td>Increase in the production of IL-8 in the colonic mucosa. Study suggests that bacteria act as a promoter of early preneoplastic lesions in the colon of rats</td>
<td>Male rats pre-treated with the carcinogen azosymethane (AOM)</td>
<td>[64]</td>
</tr>
<tr>
<td>Unknown</td>
<td>Tumor-promoting inflammation</td>
<td>Induce mRNA expression of proinflammatory cytokines, IL-1</td>
<td>Colorectal mucosa and tumors of CRC patients with and without a history of <em>S. galloyticus</em> bacteremia in the last 2 years</td>
<td></td>
<td>[87]</td>
</tr>
<tr>
<td>Wall extracted antigens (WEA) and whole bacteria</td>
<td>Sustaining proliferative signaling</td>
<td>MAPKs activation which up-regulate the expression of COX-2</td>
<td>Human colonic epithelial Caco-2-cells</td>
<td></td>
<td>[36]</td>
</tr>
<tr>
<td>Unknown</td>
<td>Inducing angiogenesis</td>
<td>Induce mRNA expression of angiogenic chemokine, IL-8</td>
<td>Feaces and colorectal tissue of CRC patients with and without a history of <em>S. galloyticus</em> bacteremia in the last 2 years</td>
<td></td>
<td>[87]</td>
</tr>
<tr>
<td>WEA and whole bacteria</td>
<td>Inducing angiogenesis</td>
<td>Over-expression of COX-2</td>
<td>Human colonic epithelial Caco-2-cells</td>
<td></td>
<td>[36]</td>
</tr>
<tr>
<td>Unknown</td>
<td>Resisting cell death</td>
<td>mRNA expression of proinflammatory cytokines, IL-1 and COX-2, as well as angiogenic chemokine, IL-8</td>
<td>Feaces and colorectal tissue of CRC patients with and without a history of <em>S. galloyticus</em> bacteremia in the last 2 years</td>
<td></td>
<td>[87]</td>
</tr>
<tr>
<td>Unknown</td>
<td>Resisting cell death</td>
<td>Higher IL-8 mRNA and NF-κB mRNA in tumorous than non-</td>
<td>Serum and tissue of CRC, CRA and healthy volunteers</td>
<td></td>
<td>[88]</td>
</tr>
<tr>
<td>Enterotoxigenic B. fragilis (BFT)</td>
<td>WEA and whole bacteria</td>
<td>Tumor-promoting inflammation</td>
<td>Over-expression of COX-2</td>
<td>Human colonic epithelial Caco-2-cells</td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------------------------</td>
<td>----------------------------</td>
<td>-------------------------</td>
<td>-------------------------------------</td>
<td></td>
</tr>
<tr>
<td>BFT</td>
<td>B. fragilis toxin (BFT)</td>
<td>Sustaining proliferative signaling</td>
<td>E-cadherin cleavage then β-catenin nuclear signaling is expressed and induces c-Myc translation</td>
<td>HT29/C1 cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sustaining proliferative signaling</td>
<td>Induces E-cadherin cleavage, interleukin-8 secretion, and epithelial cell proliferation</td>
<td>Specific pathogen-free (SPF) C57BL/6j or germfree mice</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genome instability and mutations</td>
<td>NF-kB and mitogen-activated protein kinases (MAPKs)</td>
<td>HT29/C1 cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genome instability and mutations</td>
<td>SMO-dependent generation of ROS and induction of γ-H2A.x, a marker of DNA damage</td>
<td>HT29/c1 and T84 colonic epithelial cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genome instability and mutations</td>
<td>Increases in proliferating cell nuclear antigen (PCNA) mRNA levels</td>
<td>CEAM-expressing mice</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genome instability and mutations</td>
<td>Accumulation of SUMO-conjugated p53 and production of hepatocyte growth factor (HGF) by targeting targets SENP1 (senescence-associated secretory phenotype)</td>
<td>AOM/IL-10-/- (azoxymethane/interleukin) mouse model</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genome instability and mutations</td>
<td>genotoxin-encoding genes in mucosa</td>
<td>Analysis of mucosa of patients with CRC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genome instability and mutations</td>
<td>Macrophage COX-2 is induced by superoxide and propagate genomic instability</td>
<td>Hybrid hamster cells [A(I)L]N containing human chromosome 11 and a dual-chamber tissue culture model</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genome instability and mutations</td>
<td>Double-stranded DNA breaks, tetraploidy and chromosomal instability (CIN)</td>
<td>Primary murine colon epithelial cells growth as allografts in immunodeficient mice</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genome instability and mutations</td>
<td>Epithelial cell DNA damage through the production of extracellular O₂</td>
<td>HT-29 intestinal epithelial cells and a rat intestinal colonization model</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genome instability and mutations</td>
<td>Uregulation of COX-2 and prostaglandin E2</td>
<td>Analysis of serum IgG antibodies against H. pylori (ELESA) and cagA protein (Western blot assay) in patients with colon cancer</td>
<td></td>
</tr>
</tbody>
</table>

COX-2: Cyclooxygenase-2; H. pylori: Helicobacter pylori; CIN: Chromosomal instability; CRC: Colorectal cancer; HGF: Hepatocyte growth factor; BFT: Bacteroides fragilis toxin; CM: Cyclo-modulins; SATA3: Signal transducer and activator of transcription 3; MSI: Microsatellite instability; PCNA: Proliferating cell nuclear antigen; TLRs: Toll-like receptors; cagA: Cytotoxin-associated gene; S. gallolyticus: Streptococcus gallolyticus; S. bovis: Streptococcus bovis.

Responses promote tumor development, other cytotoxic immune cells, such as natural killer (NK) and CD8+ T cells, are essential for recognizing and eliminating cancer cells[31].

One commonality across many diseases in which microbiota contribute to progression is the disruption of the mucosal/epithelial layers of organs, allowing bacteria and their metabolites to enter compartments that are not normally in close proximity to microbes[32]. This can trigger a local chronic inflammatory response, due to perpetually injured tissues and thus a constant stream of infiltrating microbes/microbial products. Resident commensal bacteria may trigger exaggerated immune responses (colitis) when key components of immune tolerance are broken and/or modify the general immune response upon entering systemic circulation[33-35]. Such disruption may allow some non-native bacterial species to colonize the gut and adapt to the new environment, including oral bacterium. For example, Fusobacterium, Peptostreptococcus, and Lactococcus access and adhere to the basement membrane of the...
**Figure 1 Tumor-promoting inflammation.** TLRs: Toll-like receptors; BFT: Bacteroides fragilis toxin; MAPK: Mitogen-activated protein kinase; ETBF: Enterotoxigenic Bacteroides fragilis; IL: Interleukin; TNF-α: Tumor necrosis factor α; COX-2: Cyclooxygenase-2; SATA3: Signal transducer and activator of transcription 3; Th17: T-helper-17.

*F. nucleatum* is an obligate anaerobic gram-negative bacteria commonly present in the mouth and typically a poor colonizer of healthy and intact intestinal mucosa[9]. Elevated *F. nucleatum* colonization in the normal tissues may predispose to the development of colorectal adenoma (CRA). Its virulence factor FadA adhesion gene (*fadA*) has been found in the colon tissue from patients with CRA and CRC in > 10-100 times higher compared to healthy individuals[36]. Other studies using metagenomic and transcriptomic analyses also have shown an enrichment of *Fusobacterium* species in CRA compared with adjacent normal tissue[7,37-39]. Infections of *Streptococcus galloyticus* (*S. galloyticus*) have been associated with CRC and CRA. *S. galloyticus*, formerly known as *Streptococcus bovis* (*S. bovis*), bacteremia has been associated with colon cancer in 25% to 80% of cases and *S. galloyticus* endocarditis has been associated with colon cancer in 18% to 62% of cases[11,40-45]. *S. galloyticus* has shown a specific association with CRC and CRA when compared with the more dominant intestinal bacteria, *B. fragilis*[46].

Enterotoxigenic *Bacteroides fragilis* (ETBF) is an anaerobic gram-negative rod that is known to have an affinity for colonizing the colonic mucosa[47]. Asymptomatic carriage can be seen in up to 40% of healthy children and adult fecal samples[48]. The pathogenicity of ETBF is a result of the *B. fragilis* toxin (BFT), which has 3 isotypes (BFT-1, BFT-2, BFT-3). BFT-1 is believed to more commonly colonize stool and BFT-2 more commonly colonizes the mucosa[49]. An association between ETBF and CRC is emerging. ETBF enhances Th17-driven inflammation and colonic tumor development. *B. fragilis* preferentially colonize the epithelial crypts of the colonic mucosa and thus evade the host immune response leading to a more stable colonization in CRC[10,50]. In general, ETBF in CRC patients has been more commonly found on the colonic mucosa as opposed to stool samples[51]. It has been hypothesized that the human colon’s exposure to BFT may lead to a chronic, possibly focal, inflammation of colonic
mucosa thus creating sites susceptible to carcinogenesis. BFT induces both acute and chronic colitis and carcinogenesis mediated by IL-17[52]. In multiple intestinal neoplasia (Min) mice, Wu et al[53] described a signal transducer and activator of transcription 3 that led to a pro-carcinogenic Th17-dependent pathway for inflammation-induced cancer by ETBF.

Some bacterial species, such as Proteobacteria, *E. coli* and *Bacteroides thetaiotaomicron*, have evolved to survive amidst harsh conditions of immune activation[54,55]. *E. coli*, which is benign under homeostatic conditions, possesses a significant growth advantage as it utilizes inflammatory nitric oxides as an energy source[54]. In the absence of IgA, the commensal bacterium *Bacteroides thetaiotaomicron* expresses high levels of gene products that are involved in the metabolism of nitric oxide and generates pro-inflammatory signals in the gut[55].

*F. nucleatum* is capable of affecting both innate and adaptive immune responses[7]. *F. nucleatum* enhances inflammation through engagement of its adhesin FadA. The host endothelial receptor for FadA is the vascular endothelial cadherin (CDH5), which is a member of the cadherin family[15]. FadA adheres to and invades epithelial and endothelial cells and activates inflammatory cytokines (IL-6, IL-8, IL-10, IL-18, TNF-α and NF-kB levels) that create a pro-inflammatory environment which accelerates the progression of CRC (Table 1)[15,56,57].

Within the host cytoplasm, *F. nucleatum* may release its RNA, leading to detection by cytosolic retinoic acid-inducible gene I (RIG-I), a cytosolic pattern recognition receptor (PRR) responsible for the type-1 interferon (IFN1) response, thereby stimulating cytosolic NF-kB and promoting inflammation. It generates a pro-inflammatory microenvironment outside the tumor cell through recruitment of tumor-infiltrating immune cells as the primary mechanism[15]. The ability of *F. nucleatum* to invade HEK293T cells, human embryonic kidney cells, which lack endogenous Toll-like receptors, allows the bacteria to activate a pro-inflammatory response through cytosolic pattern recognition receptors, NOD-1, NOD-2 and NF-kB signaling. This pro-inflammatory response is mediated by the p38 MAPK signaling pathway[58]. In a mouse model of intestinal cancer, introduction of *F. nucleatum* to ApcMin+/− mice resulted in accelerated small intestinal and colonic tumorigenesis, infiltration of specific myeloid cell subsets into tumors, and an NF-kB proinflammatory signature. This proinflammatory signature is shared with human CRC tissues with a high *Fusobacterium* abundance[39]. Additionally, a positive correlation has been found between the bacterial concentration in human tissues, for adenoma and non-adenoma controls, and TNF-α and IL-10 abundance[59]. Unlike other bacteria associated with CRC, however, *F. nucleatum* does not exacerbate colitis, enteritis, or inflammation-associated intestinal carcinogenesis. This suggests that *F. nucleatum* may drive non-colitis-associated intestinal tumorigenesis (Table 1).

*Helicobacter pylori* (H. pylori) is a small, spiral, gram-negative bacillus that has a well-established association with the development of gastric cancer and is considered a carcinogen by the World Health Organization[60]. There are conflicting data on the correlation of *H. pylori* as an etiological factor of CRC. The production of oxidative stress involves alteration of the intragastric environment through bacterial and neutrophilic production of ROS, pro-inflammatory cytokines, and upregulation of cyclooxygenase-2 (COX-2)[61]. This includes excessive production of ROS by neutrophils in an effort to eradicate the bacteria. The bacterial infection causes inflammation, leading to increased production and activity of COX-2 and prostaglandin E2, a biomarker associated with inflammation and CRC risk. Some evidence has been reported to support a potential association between *H. pylori* and CRC. Shmuely et al[62] described that cag-positive *H. pylori* strains were associated with a 10.6-fold increased risk of CRC compared to cagA negative strains. A systematic review with a meta-analysis found a moderate correlation between *H. pylori* infection and the risk of CRC[63]. Further investigations should be conducted to determine the role of cagA, and the mechanism by which *H. pylori* induces gastric carcinogenesis and potentially CRC.

*S. gallolyticus* produce a similar effect as *H. pylori*. In vitro experiments have shown that the binding of *S. gallolyticus* to intestinal cells leads to production of cytokines[36, 46,64]. It has been described that treatment with *S. bovis* or wall-extracted antigens in adult rats promoted the progression of preneoplastic lesions through the increased formation of hyperproliferative aberrant colonic crypts, and increased the production of IL-8 in the colonic mucosa[64].

https://www.wjgnet.com
AVOIDING IMMUNE DESTRUCTION

*F. nucleatum* directly interacts with the host immune system. Fap2, an autotransporter domain found in the bacterial outer membrane protein, facilitates *F. nucleatum* to adapt to different body habitats. In the oral cavity, Fap2 attaches to neighboring bacteria by co-adhering to different microorganisms, increasing the diversity and the stability of the developing dental biofilm[65]. Fap2 protein adhesion directly interacts with immunoglobulin and immunoreceptor tyrosine-based inhibitory motif domains (TIGIT) expressed in tumor-infiltrating lymphocytes, leading to the inhibition of NK cell cytotoxicity (Figure 2)[31]. TIGIT is an inhibitory receptor present on human NK cells and on various T cells. The interaction between *F. Nucleatum*, Fap2 and TIGIT induces lymphocytes apoptosis and generates an immuno-suppressive microenvironment that promotes the progression of colorectal tumors[15]. Fap2 protein is believed to directly interact with TIGIT, leading to the inhibition of NK cell cytotoxicity[31]. These results identify a bacterium-dependent, tumor-immune evasion mechanism in which tumors exploit the Fap2 protein of *F. nucleatum* to inhibit immune cell activity via TIGIT[31]. Several clinical and animal studies have correlated the abundance of *F. nucleatum* with suppression of antitumor T cell response by affecting the enhancement of myeloid-derived suppressor cells and tumor-associated macrophages (Table 1).

Furthermore, various microRNAs are induced during the macrophage inflammatory response and modulate host-cell responses to pathogens. MicroRNA-21 increases the levels of IL-10 and prostaglandin E2, which suppress antitumor T-cell-mediated adaptive immunity through the inhibition of the antigen-presenting capacities of dendritic cells and T-cell proliferation in CRC cells (Figure 2). *F. nucleatum* can expand myeloid-derived immune cells, which inhibits T cell proliferation and activation resulting in tumor cell growth by blocking the middle (G1) phase of the cell cycle and attracting myeloid-derived suppressor cells to the tumor site. *F. nucleatum* also induces T-cell apoptosis in CRC (Table 1)[66].

DEREGULATING CELLULAR ENERGETICS

Fermentation of dietary fibers leads to the production of short-chain fatty acids (SCFAs), such as butyrate, which serve as the primary energy source for intestinal cells (Figure 3). The presence of butyrate has an anticancer effect by starving cancer cells. Additionally, it produces epigenetic modifications by inhibiting cell proliferation and promoting apoptosis by inhibiting histone deacetylase[67]. Non-cancerous colonocytes utilize butyrate as their primary energy source in contrast with CRC cells which primarily use glucose and undergo increased glycolysis, a phenomenon known as the Warburg effect or aerobic glycolysis, decreasing mitochondrial oxidative metabolism [55,68].

*Fusobacteria* species also participate in the metabolism of amino acids ingested in the diet, generating formyl-methionyl-leucyl-phenylalanine and SCFAs that chemoattract myeloid cells. This explains the intratumoral expansion of myeloid cells that interconnect tumor, bacterial and immune cells metabolism (Table 1)[39]. In addition, *F. nucleatum* is an asaccharolytic bacterium, a competitive advantage in the tumor milieu; therefore, it will not compete for glucose, a preferred substrate for tumor metabolism[69]. Some bacterial communities could protect intestinal cells from inflammation and tumorigenesis. SCFAs function as signaling molecules between gut epithelia and immune cells modulating changes in the gene expression and providing nutrition to colonocytes. The three major SCFAs produced by bacterial fermentation of carbohydrates are acetate, propionate, and butyrate[70]. Published evidence supports the idea that butyrate is tumor-suppressive by inhibiting the proliferation of CRC cell lines while stimulating their apoptosis and/or differentiation (Figure 3)[40,54,55]. A Western diet slows mucus growth rate and increases penetrability of the colonic mucus barrier, and this effect co-occurs with the shifts in microbial community characterized by gradual decrease of SCFA-producing bacteria, such as *Bifidobacterium* and *Bacteroidales* family and increases in *Firmicutes*[71].

A higher prevalence of *Fusobacterium* and *Porphyromonas* in feces of CRC patients has been reported[19,20]. An increase in glutamate levels by 76% in fecal samples from colon cancer patient samples was reported but there was no increase in glutamine levels. The authors hypothesized that tumor cells may exhibit an increase of glutaminase activity, which results in the conversion of glutamine to glutamate. This
Figure 2 Avoiding immune destruction. TIGIT: T-cell immunoglobulin and ITIM domain; NK: Natural killer.

supports the theory of the role of certain bacteria as "driver bacteria" with pro-
carcinogenic characteristics that contribute to tumor development and then a
transition to "passenger bacteria" that contribute to an environment conducive to
cancer[19].

Hester et al[72] compared bacteria and SCFA in stool samples of AA and CA in a
small pilot study. They found lower acetate, butyrate, total SCFA content and a higher
pH in AA compared to the other racial groups. Similar results reported in another
study where AA had increased levels of SCFAs in stool than other racial/ethnic
groups and significantly lower intake of non-starchy vegetables[73]. Wnt/B-catenin
signaling plays a fundamental role in several biological processes such as development
and cell proliferation related to tumorigenesis. Butyrate has beneficial effects in
reducing colon cancer risk with anti-inflammatory, immunomodulatory effects and
down regulating Wnt signaling, which inhibits cell proliferation and migration. In
addition, AA had higher levels of Firmicutes bacteria compared to CA and Hispanics.
Moreover, the ratio of Firmicutes compared to Bacteroides, which has been associated
with obesity, was higher in AA. These results continue to suggest that AA have higher
risk of developing colon cancer[72].

SUSTAINING PROLIFERATIVE SIGNALING

E-cadherin is a type of cell adhesion molecule and usually is targeted by various
intestinal bacteria promoting epithelial proliferation by activating the Wnt/\(\beta\)-catenin
pathway. ETBF, an enterotoxin-producing bacterium, is involved in the initiation and
progression of CRC by not only modulating the mucosal immune response, but also
inducing epithelial cell changes. BFT promotes cleavage of E-cadherin (Table 1). This
produces nuclear translocation of \(\beta\)-catenin and subsequent transcription of the c-Myc
proto-oncogene causing hyperplasia due to proliferation of colonocytes[74]. After
treatment of HT29/C1 human colon cancer cells with BFT, cleavage of membrane-
associated E-cadherin and loss of intercellular adhesion occurs. This in turn leads to
subsequent expression of \(\beta\)-catenin nuclear signaling and induction of c-Myc
translation resulting in persistent cell proliferation[75]. The presence of ETRF may
contribute to chronic colon diseases, including oncogenic transformation, intestinal
inflammation, chronic colonic dysfunctions, and colorectal precancerous and
cancerous lesions[75].
The mucosa associated/internalized \textit{E. coli} have been shown to occur more frequently in patients with CRC than healthy controls[76-78]. Pathogenic \textit{E. coli} produces virulence factors called cyclomodulins (CM). These CMs can modulate cell cycle progression, apoptosis, cell differentiation, and proliferation (Table 1)[79,80].

Senescent cells secrete growth factors that increase cell proliferation resulting in tumor growth. Colibactin-producing (pks+) \textit{E. coli} promote CRC cancer in a murine AOM/IL-10−/− (azoxymethane/IL) mouse model by expression of SENP1, microRNA-20a-5p, hepatocyte growth factor (HGF) and phosphorylation of HGF receptor. In addition, senescence-associated secretory phenotype induces epithelial cell proliferation \textit{via} the production of growth factors by senescent cells. SENP1 downregulation and p53 SUMOylation are key features of pks+ \textit{E. coli}-induced senescence as result of modifying p53 function[81].

\textit{F. nucleatum} has shown a propensity to adhere to mucosa enabling it to invade human epithelial and endothelial cells[7]. The percentage of \textit{F. nucleatum}-enriched CRC gradually increases from rectum to cecum, suggesting that the rate of bacteria proliferation differs among the intestinal sites[82]. \textit{F. nucleatum} enhances epithelial proliferation through engagement of its adhesin FadA. FadA modulates E-cadherin, a tumor suppressor gene, and activates \( \beta \)-catenin signaling, leading to increased expression of oncopgenes, Wnt genes, and inflammatory genes as well as growth stimulation of CRC cells (Figure 4 and Table 1)[56]. The FadA binding site on E-cadherin has been mapped to an 11-amino-acid region. The experimental use of an 11-amino-acid inhibitory synthetic peptide has been shown to inhibit \textit{F. nucleatum} from binding and invading the epithelial cells and abolishing all subsequent host responses, including tumor growth and inflammatory responses (Table 1)[56].

\textit{S. gallolyticus}-induced mucosal inflammation may lead to enhanced mucosal permeability and increased entry for \textit{S. gallolyticus} into colonic cells. \textit{S. gallolyticus} \textit{in vitro} also has shown strong adherence to the proteins of the extracellular matrix, collagen I, collagen II and collagen IV, enabling it to have easy entry into cells and successfully colonize both colonic and vascular tissues[46,83]. \textit{S. gallolyticus} is able to grow in bile and can easily bypass the hepatic reticulo-endothelial system and access the systemic circulation[84]. \textit{S. gallolyticus} endocarditis infections also have been associated with increased hepatic dysfunction[85]. It is hypothesized that underlying colonic disease or \textit{S. gallolyticus}'s effects on the liver’s production of immunoglobulins
and bile acids may promote *S. gallolyticus* overgrowth and thus altering the colonic microbiome[85]. Although the mechanism is unclear, *S. gallolyticus* whole bacteria and wall extracted antigens have shown greater propensity towards colonizing colonic tumor cells compared to normal mucosa (Table 1)[36]. *S. gallolyticus* also may induce uncontrolled cell proliferation by triggering proteins known as mitogen activated protein kinases that promote cellular transformation and genetic mutations[36].

**INDUCING ANGIOGENESIS**

*S. gallolyticus* is believed to promote the advancement of preneoplastic lesions to neoplastic lesions through the increased formation of aberrant colonic crypts that show increased expression of cytokines such as IL-8 (Table 1)[64]. IL-8 is a cytokine that stimulates angiogenesis which in turn also may promote carcinogenesis[86]. The pattern of IL-8 mRNA expression in the tumor microenvironment may function as a significant regulatory factor rather than a promoter for the adenoma progression and the adenoma–canceroma transition. This might be attributed to the angiogenic role of IL-8 by which new blood vessels are formed to meet the increasing demands of cancer growth[46]. Moreover, *S. gallolyticus*-induced overexpression of COX-2 via prostaglandins acts as a promoter of carcinogenesis by inducing angiogenesis[36].

**RESISTING CELL DEATH**

*S. gallolyticus* plays an essential role in the oncogenic progression through different factors that cause an anti-apoptotic effect in colorectal mucosa as shown in Figure 5. In Abdulamir *et al*[46], *S. gallolyticus* appeared to induce mRNA expression of proinflammatory cytokines, IL-1 and COX-2 which induce transformation of normal or premalignant colorectal tissues into malignant status. After analysis of mRNA expression of the oncogene c-Myc and antiapoptotic Bcl-2 were not linked to
Figure 5 Resisting cell death. WEA: Wall extracted antigens; PGE2: Prostaglandin E2; IL: Interleukin; COX-2: Cyclooxygenase-2.

Colonization by these bacteria, but were associated with CRC transformation. These results may suggest that *S. gallolyticus* does not induce oncogenic changes or suppress cellular apoptosis, and might instead have a role as a propagator for premalignant or oncogene-positive tissues to enter the transformation cycle through inflammatory and angiogenic microclimates (Table 1)[87]. The release of PGE2-mediated by *S. gallolyticus* is correlated with the overexpression of COX-2, which is seen in about 85% of colon cancers, and through its association with enhanced angiogenesis and inhibition of apoptosis, is favorable to the development and progression of CRC[36]. In addition, this group also found *S. gallolyticus*-seropositive CRC patients were significantly associated with higher mRNA expression of both NF-kB and IL-8 that play an integrated role in a series of steps to escape cell death signals[88].

**GENOME INSTABILITY AND MUTATIONS**

A study comparing microbiota from more than 1000 fecal samples including 416 pairs of twins identified numerous microbial taxa whose abundance was influenced by host genetics. In the case of monozygotic twins, a more similar microbiota was observed than in dizygotic twins. However, it is unclear whether the host's genetic variation shapes and interacts with the gut microbiome to affect the host's phenotype[89]. The analysis of gut microbiota in stool of 2084 participants in the Healthy Living in an Urban Environment Study described that people who live in the same city tend to show similar gut microbiota with other people of their ethnic origin[90]. Ethnic differences in alpha diversity and inter-individual differences were independent of metabolic health and were only partially explained by ethnic characteristics, including sociodemographic, lifestyle, or dietary factors. Therefore, the ethnicity of individuals may be an important factor to consider in the research of microbiome and cancer CRC[90].

Sulfidogenic bacteria, such as *Fusobacterium*, *Desulfovibrio* and *Bilophila wadsworthia*, have been implicated in CRC development through the production of hydrogen sulfide (Figure 6)[91]. Hydrogen sulfide is a genotoxic compound that has been shown to damage DNA leading to genomic or chromosomal instability (CIN), effecting DNA repair in a multistep carcinogenic process. One hypothesis is that hydrogen sulfide diffuses into intestinal epithelial cells and interferes with mitochondrial function, ultimately leading to hyperproliferation via the Ras/MAPK pathway[92]. The hyperactivation of the Ras/MAPK pathway is a known mechanism of carcinogenesis in CRC. A study described the concentration of specific bacterial DNA in colonic tissue biopsies, and showed that AA with CRC had higher concentrations of sulfidogenic
bacteria compared to NH whites. At the same time, AA with CRC had increased levels of sulfidogenic bacteria compared to AA without CRC[93].

Some bacterial products, such as colibactin or BFT, may damage the genetic information inside the nucleus, and actively induce DNA damage in organs that are in direct contact with the microbiome (Figure 6 and Table 1)[94]. These genotoxins may directly promote the release from macrophages and other inflammatory cells of reactive oxygen species (ROS), reactive nitrogen species (RNS) and hydrogen sulfide (H$_2$S) from the bacterial microbiota products[95]. ROS and RNS inhibit the activity of T cells, including antitumor cytolytic CD8$^+$ T cells. An abundance of myeloid-derived suppressor cells also leads to increased production of ROS and RNS and subsequent tumor-supporting inflammation and neoangiogenesis[70]. Hydrogen sulfide has been shown to damage DNA leading to genomic or CIN, affecting DNA repair in a multistep carcinogenic process. One hypothesis is that hydrogen sulfide diffuses into intestinal epithelial cells and interferes with mitochondrial function, ultimately leading to hyperproliferation via the Ras/MAPK pathway[92]. Furthermore, metabolic actions of the microbiome may promote the development of CRC by activation of other genotoxins such as acetaldehyde, dietary nitrosamine and other carcinogens[95].

As shown in Figure 6, various gut microbes and their bacterial products can cause DNA mutations. *E. coli* contribute to the accumulation of mutations resulting from DNA damage induced by genotoxins, or by downregulating host DNA mismatch repair proteins (Table 1)[96]. One cyclomodulin is a hybrid polyketidinonribosomal peptide called colibactin, encoded by the polyketide synthase (pks) genomic island. Colibactin possesses genotoxic properties that result in CIN and double-strand breaks in the DNA of human eukaryotic cells. *E. coli* strains harboring the pks genotoxic island, which are found in a significantly high percentage of inflammatory bowel disease and CRC patients[79]. In *vitro* studies using different mammalian cells, including normal intestinal cells, showed that pks+ *E. coli* produce DNA double-strand breaks affecting the normal cellular division with the consequent cell cycle arrest and aneuploidy[97]. Conversely, different commensal bacteria could harm DNA strands by stimulating host inflammation and producing a pro-oxidant microenvironment. *ETBF* causes DNA damage by peroxide that comes from the colonocyte expression of enzyme spermine oxidase[98]. *E. faecalis* may have a microbiome driven bystander effect that leads to increased COX-2 expression in macrophages leading to ROS formation, which in turn promotes CIN in intestinal epithelial cells. Primary colonic epithelial cells can have induction of CIN or malignant tumor aneuploidy via macrophages that have been polarized by *E. faecalis*[99]. These results validate a novel mechanism for CRC that involves endogenous CIN and cellular transformation arising through a microbiome-driven bystander effect (Figure 6)[100].
**Enterococci** have increasingly been shown to cause infections in the elderly. They are gram positive, facultative anaerobe, diplococci that grow as short chains. *E. faecalis* has been found to aggregate at higher levels in stool samples in CRC patients than in healthier controls\[101,102\]. It also has been shown to be in greater abundance in the adjacent tissues of CRC when compared to healthy mucosa in controls\[103\]. It has been postulated that *E. faecalis* can damage colonic DNA and cause genomic instability via its ability to generate ROS that predisposes to mutations leading to CIN and subsequent carcinogenesis (Figure 6)\[12\].

A large abundance of *F. nucleatum* has been shown to induce a series of tumor-specific molecular events, including the CpG island methylating phenotype (CIMP), microsatellite instability (MSI), and genetic mutations in BRAF, CHD7, CHD8 and TP53. A higher abundance of *F. nucleatum* DNA in CRC tissues has been associated with an increased CRC-specific mortality, evidence that suggests *F. nucleatum* may potentially serve as a prognostic biomarker (Figure 6)\[82\].

Several etiological mechanisms of *H. pylori* in CRC pathology have been hypothesized. One alternative is that chronic *H. pylori* infection can lead to hypergastrinemia, which is considered to be a nutrient factor in the colorectal mucosa and may lead to the promotion of mutagenesis. In addition, this bacterial infection and bacterial CagA protein lead to chronic gastritis with an increase in gastrin production (Figure 6)\[62\]. Ammonia produced by *H. pylori* might also act as an endogenous carcinogen. More studies are needed to investigate the potential role of gastrin in the mechanisms of metastasis of the tumor cells\[63\].

**CONCLUSION**

In this review, we highlight the various gut and oral microbiota associated with CRC and CRA, and their proposed molecular mechanisms in relation to the processes of “the hallmarks of cancer.” Common oncogenic mechanisms include disruption of mucosal/epithelial layer of organs, promotion of inflammation and neoangiogenesis, use of genotoxins by bacteria, altering of genetic expression, and modification or weakening of the existing immune system.

To understand whether host genetics play a role in the associations between race/ethnicity, and oral and gut microbiota, it is essential to perform genome-wide association studies with microbial data among diverse populations. In 2016, 81% of existing genome-wide association study data was generated from individuals of European descent where the proportion of samples from individuals of African descent only increased by 2.5% and that of ancestry Hispanic or Latin American by about 0.5%\[104\]. There are huge gaps in knowledge related to understanding the underlying causes of racial/ethnic differences in the gastrointestinal microbiota and their possible role in colon cancer. Therefore, it is essential to be able to generate new knowledge by generating large and representative cross-sectional studies with gut, oral and fecal samples in populations that include individuals with various social determinants of health (for example, NHANES). Variables that could cause these racial/ethnic differences, such as diet, habits, socioeconomic status and oral health status should also be further studied. Genetic factors may contribute to biological reasons for CRC disparity. Diverse studies described somatic alterations in well-known CRC genes (APC, BRAF, KRAS, and PIK3CA) and lower frequency of MSI, a good prognostic biomarker, among AA patients\[105,106\]. *F. nucleatum* levels have been found to be significantly higher in AAs\[22\] and has been shown to induce a series of tumor-specific molecular events, including CIMP, MSI, and genetic mutations in oncogenes and tumor suppressors\[82\].

Differences in the immunological profiles of colon tumors from AA compared to CA suggested a deficiency of appropriate immune defense mechanisms in terms of gene expression, recruitment of immune cells and systemic secretion of cytokines. As such, these immune differences could be mitigated through population-specific therapeutic approaches\[106\]. Studies of associations between specific taxa in microbiome and race/ethnicity could provide an insight for examining specific bacterial members as mediators of health disparities. Defining the composition of a “healthy microbiota” is one of the challenges in the field of microbiome research. Our review reveals that unique opportunities exist in targeting racial/ethnic differences in oral and gut microbiome for a greater understanding of the complexity of CRC and CRA etiology and carcinogenesis.

As knowledge is gained regarding the microbiome as it pertains to CRC, the clinical implications will continue to grow and hopefully come to the forefront in the
prevention, detection, and treatment of CRC in clinical practice. Current non-invasive screening methods for CRC include FIT and Cologuard. However, their ability to detect precancerous lesions is not entirely reliable, thus creating a void for superior noninvasive screening methods that microbiome studies could likely one day fill[107]. When combining testing for oral bacteria \textit{F. nucleatum} with FIT, the combination showed superior sensitivity than FIT alone in detecting CRC, and additionally increased the performance of adenoma detection, suggesting the potential of bacterial biomarkers as more useful diagnostic tools for certain patients with familial history[108]. In addition, the ratio \textit{F. nucleatum}/\textit{Bifidobacterium} showed superior sensitivity of 84.6% and specificity of 92.3% for diagnosing CRC in comparison with the use of a single fecal bacterial biomarker candidate[109].

The deficit in the number of butyrate-producing bacteria can have detrimental consequences in the progression of the disease, hence the screening of SCFA and microbial-derived metabolites have potential as biomarkers and diagnostic tools for CRC. Some studies already have shown that using microbiome profiles in conjunction with certain risk factors such as age, race and body mass index can help predict high-risk colon \textit{vs} one with adenomas or carcinomas[110]. While detection of CRC and precancerous lesions is our current goal in preventing CRC mortality, the microbiome also provides promise into potentially preventing CRC by inhibiting colorectal tumorigenesis. A recent study by Li \textit{et al}[111] aimed to look at the role of depleted bacteria, specifically \textit{Streptococcus thermophilus} (\textit{S. thermophilus}), in CRC to see if when used as a probiotic that it could prevent CRC. They showed in a mouse model that tumor formation could be reduced using \textit{S. thermophilus} by oral gavage, and it was specifically the \textit{β}-Galactosidase secreted by \textit{S. thermophilus} that was critical in retarding the growth of CRC cells. \textit{S. thermophilus} also was able to increase other known probiotics, including \textit{Bifidobacterium} and \textit{Lactobacillus}, \textit{via} \textit{β}-Galactosidase. Albeit, just in a mouse model, it highlights the potential for possible prevention and reversal of CRC by use of the microbiome.

Additional prospective human studies must be undertaken to determine the role of the microbiome in CRC therapy and in the reliability of certain profiles for screening and prevention. Furthermore, when certain microbiome signatures that predispose risk for CRC do arise, the age at which one is to start looking for these predispositions also needs to be established to enable clinical use. The heterogeneity of CRC could be related to different microbiota communities that either predispose or provide resistance to the disease, and the profile analysis of the oral microbiome may offer an alternative screen as a biomarker for detecting CRC[112]. Only single studies detected associations with \textit{ Fusobacterium} or \textit{Porphyromonadaceae} and CRC[110,113]. Further studies with a larger sample size are needed to confirm the identified associations and estimate the potential utilization of the oral microbiota and periodontal diagnosis and treatment for use in CRC early detection or prevention.

The gut microbiome and its metabolites have therapeutic implications for CRC and other cancers. It has been postulated that therapeutic response to immune checkpoint inhibitors (ICI) may be influenced by presence of the gut microbiome. In 100 patients diagnosed with non-small cell lung cancer and on ICI therapy, the stool of patients responding to ICI therapy was rich in \textit{ Akkermansia muciniphila} compared to nonresponders[114]. The baseline microbiome or its modulation using antibiotics, probiotics or FMT have influenced treatment efficacy in numerous cancers[115]. For example, patients with non-small cell lung cancer, renal cell carcinoma or urotheilial cancer treated with antibiotics for routine indications shortly before, during, or shortly after treatment with anti-PD1/PD-L1 mAB had significantly lower progression-free survival and overall survival rates compared to patients who had not received antibiotics. This suggests that disrupting the gut microbiota (\textit{via} antibiotic use) could potentially impair anti-tumor immune responses as well as response to immune checkpoint blockade[114,115].

The gut microbiome is a dynamic mediator of immune and cellular response to cancer and influences the efficacy of cancer therapies. Manipulating the gut microbiome will likely emerge as a viable option of modulating the responsiveness of cancers to immune mediated and other therapies. In addition, other decisions regarding the use of antibiotics with cancer and other therapies should be weighed carefully considering their impact on the gut microbiome. Large prospective studies on the impact of dietary interventions (prebiotics), antibiotic use and the influence of environmental pollutants are needed to clarify many unanswered questions on the factors that impact the gut microbiome and the durability of that effect. Lastly, the price and availability of microbiome analysis will have to come to a point where it is widely available and accessible to make it mainstay in the clinical realm.
ACKNOWLEDGEMENTS

Dr. Lim S for reviewing our manuscript and offering good feedback.

REFERENCES


Leport C, Bare A, Leport J, Vilde JL. Incidence of colonic lesions in Streptococcus bovis and


64 Ellmerich S, Djouder N, Schöller M, Klein JP. Production of cytokines by monocytes, epithelial and


Tortora SC et al. Microbiome and colorectal cancer

[PMID: 20233397 DOI: 10.1186/1471-2180-10-78]


null
IV. The oral-gut-circulatory axis in colorectal cancer

The oral-gut-circulatory axis: from homeostasis to colon cancer

Sofia C. Tortora¹, Maria Gonzalez Agurto² and Laura A. Martello¹

¹Department of Medicine and Division of Gastroenterology & Hepatology, SUNY Downstate Health Sciences University, Brooklyn, NY, United States; ²Departamento de Rehabilitación Craneofacial Integral, Universidad de Los Andes, Santiago, Chile

The human microbiota is widely recognized as providing crucial health benefits to its host, specifically by modulating immune homeostasis. Microbial imbalance, known as dysbiosis, is linked to several conditions in the body. The oral cavity and gut host the two largest microbial communities playing a major role in microbial-associated diseases. While the oral-gut axis has been previously explored, our review uniquely highlights the significance of incorporating the circulatory system into this axis. The interaction between immune cells, inflammatory factors, circulating bacteria, and microbial metabolites influences the homeostasis of both the oral and gut microbiota in a bidirectional manner. In this comprehensive review, we aim to describe the bacterial components of the oral-gut-circulatory axis in both health and disease, with a specific focus on colon cancer.

KEYWORDS
microbiota, immune response, dysbiosis, oral-gut axis, circulation, colon cancer

Introduction

Disturbance of the microbiota homeostasis has emerged as a significant factor correlated with a multitude of diseases within the human body. The study of the microbiome and its association with cancer has primarily focused on an organ-specific relationship; however, increasing evidence underscores the pivotal role played by the microbial and inflammatory milieu at distant anatomical sites as an important regulator in both healthy and pathological processes.

The oral cavity and gut harbor the two largest microbial habitats in the body and play a significant role in infectious diseases and in different types of cancer. The mouth is the entry point of the digestive tract and is continuously exposed to various exogenous elements, including microorganisms, nutrients, and xenobiotics. Bacterial translocation can occur through different routes, such as swallowing, aspiration, and through circulation or transported inside immune cells. Certain oral pathogens, such as Porphyromonas gingivalis (P. gingivalis) and Fusobacterium nucleatum (F. nucleatum), have been found to invade and colonize the oral epithelial cells and periodontal tissues where the bacteria
can release virulence factors and toxins that disrupt the integrity of the oral mucosa. Additionally, chronic oral infections, such as periodontitis, can weaken the barrier function facilitating the translocation of oral pathogens into the bloodstream, ultimately reaching distant sites such as the liver, spleen, and gastrointestinal tract. In the liver, oral microbes can trigger a low-grade inflammatory reaction and impact various body sites through the release of cytokines. Once in the gut, these oral bacteria can interact with the existing gut microbiota, influencing its diversity and functionality. Changes in the microbial communities may trigger immune responses, inflammation, and increased intestinal permeability, allowing opportunistic bacteria to invade, and toxins and metabolites to leak into the bloodstream. Disturbances in either the oral or gut microbiota can cascade into the other, potentially affecting systemic health and underscoring their interconnectedness. This mutual relationship highlights the importance of considering both systems together when studying and addressing various health conditions.

In this review, we summarize the current knowledge of the oral-gut axis, including the role of the circulatory system in bacterial translocation and systemic inflammation while highlighting the importance of the interconnectedness of these systems. Both oral pathogens, *P. gingivalis* and *F. nucleatum* have been implicated in the pathogenesis of colon cancer by directly affecting the host cell signaling pathways involved in cell proliferation, apoptosis, and DNA damage repair. These microbial interactions can further exacerbate inflammation and promote the growth and survival of colorectal cancer cells. The presence of these oral pathogens in the gut underscores the connection between of the oral-gut axis in the context of carcinogenesis.

**The immune response in oral and intestinal mucosa**

The mucosal immune system constitutes a specialized and region-specific defense network safeguarding a substantial portion of the inner surface of human anatomy, encompassing the mucosal linings of the respiratory system, urogenital tract, oropharyngeal region, GI tract, and also the exocrine glands. Specialized functions in the oral cavity and GI tract lead to variations in immune responses. These include differences in mucosal surfaces, microbial composition, and immune cell populations. The main roles of the oral cavity, digestion and speech articulation, require different immune adaptations compared to the GI tract, which primarily handles nutrient absorption and hosts a larger and more diverse microbial community. These distinctions contribute to unique immune profiles and responses in each region (Moutsopoulos and Konkel, 2018).

The intricate mucosal surfaces of the oral cavity, including the gingiva, tongue, and buccal mucosa, constitute the first line of defense against pathogens. The oral mucosa components ensure the integrity of oral tissues and form a complex defense network that protect against potential threats while maintaining tolerance to harmless entities. Oral epithelia are multilayered barriers with highly diverse antigenic responses and different expression patterns of cytokeratin. Salivary glands secrete saliva that contains enzymes and proteins, like lysozyme and lactoferrin, which possess antimicrobial properties (Moutsopoulos and Konkel, 2018). The GI immune system has a unique characteristic known as oral tolerance allowing the immune system to be unresponsive or tolerant to ingested protein antigens. Although the immune system becomes more mature and less flexible as we age, the oral cavity and gut immune systems still actively maintain tolerance to dietary antigens in adulthood.

Both oral and intestinal mucosa have specialized lymphoid tissues known as mucosa-associated lymphoid tissue (MALT). In the oral cavity, MALT includes tonsils and adenoids, while in the gut, it encompasses Peyer’s patches, isolated lymphoid follicles. These lymphoid tissues contain immune cells, including T cells, B cells, innate lymphoid cells (ILCs), and antigen-presenting cells (APC), which play a crucial role in immune responses (Moutsopoulos and Konkel, 2018; Suarez et al., 2021). In the gut, innate immune tactics involve employing a combination of defenses such as a protective mucus layer, antimicrobial peptides (AMPs), and the coordinated action of ILCs. These mechanisms work together to contain a significant portion of the microbial community within the interior space of the intestinal tract. ILCs are innate equivalents of T cells without antigen-specific receptors, playing pivotal roles in immune responses by producing effector cytokines, maintenance of mucosal barriers, such as the epithelial lining of the oral and intestinal mucosa, and regulating other immune cells including T cells, B cells, and dendritic cells. They are present in lymphoid and non-lymphoid organs, along with mucosal barriers exposed to allergens, commensal microbes, and pathogens whereas these cells are uncommon in the bloodstream. Different ILC subsets, including ILC1s, ILC2s, and ILC3s, each have their own specific functions and cytokine profiles. ILC3s are abundant in the intestinal mucosa and have been implicated in inflammatory bowel disease (IBD). They play a dual role: promoting tissue repair and maintaining mucosal homeostasis, while also contributing to inflammation when dysregulated. Additionally, ILCs have been suggested to have roles in tumor immune surveillance, contributing to the recognition and control of tumor cells. However, their specific functions and contributions to antitumor immunity are still being studied (Panda and Colonna, 2019).

Regulatory T cells (Tregs) and other regulatory immune cells play an essential role in suppressing immune responses and maintaining tolerance (Upadhyay et al., 2013). Microbial-associated molecular patterns (MAMPs) are conserved structural components of microorganisms that are recognized by the innate immune system. Biofilm and bacterial metabolic products like lipopolysaccharides (LPS) are the main MAMPs that can stimulate the expression and production of pro-inflammatory cytokines through activation of toll-like receptors (TLRs). TLRs serve as crucial mediators in the inflammatory pathways, significantly contributing to the orchestration of immune responses against a diverse range of ligands originating from pathogens. They play a vital link connecting adaptive immunity with innate immunity. TLRs recognize MAMPs and activate signaling pathways that lead to the production of pro-
inflammatory cytokines and chemokines, as well as the upregulation of co-stimulatory molecules and antigen-presenting molecules (Mackey and McFall, 2006).

Moreover, nucleotide-binding oligomerization domain-like receptors (NLRs) also communicate with several innate immune sensors and receptors in the oral and intestinal epithelial cells. NLRs binding and downstream signaling by cytokines, chemokines, and antimicrobial peptides production provoke an inflammatory reaction (Franchi et al., 2009; Cekici et al., 2014). One major function of NLR proteins is to regulate and modulate inflammatory signaling pathways, including NFκ-B and MAPK. Once activated, NFκ-B translocates to the nucleus, where it binds to target genes and induces the transcription of pro-inflammatory cytokines and other immune response genes. Similarly, the activation of MAPK signaling pathway leads to the production of pro-inflammatory cytokines, including tumor necrosis factor alpha (TNF-α), interleukin-6 (IL-6) and interleukin-1 beta (IL-1β) (Franchi et al., 2009). Therefore, NLR proteins play an important role in maintaining the balance between immune defense and inflammatory damage in the oral and gut microenvironment.

Neutrophils play a vital role in homeostatic immunity by serving as the front-line defenders against microbial threats. Rapidly responding to infection signals, neutrophils migrate to affected areas, where they engulf and neutralize pathogens through phagocytosis and releasing antimicrobial proteins and cytokines modulating the behavior of other immune cells. Additionally, neutrophils support the epithelial barrier’s function by participating in the tissue repair process (Moutsopoulos and Konkel, 2018).

Dendritic cells (DCs) are another important antigen-presenting cell that play a critical role in linking innate and adaptive immunity. They derive from hematopoietic stem and progenitor cells (HSPCs) in the bone marrow and are found in various locations throughout the body. Being the most potent cells responsible for activating and directing naïve T cells, DCs in the oral region hold significant importance in orchestrating both immune responses and tolerance within the oral mucosa (Hovav, 2014). Unlike in murine models, the comprehensive examination of dendritic cells in human oral tissues remains limited, with the primary focus typically centered on Langerhans cells (LCs) (Hovav, 2014). Nonkeratinized mucosal regions, including the soft palate, ventral tongue, lip, and floor of the mouth, show the highest LC concentration, while the keratinized mucosa of the hard palate displays the lowest LC density (Daniels, 1984). LCs within the oral mucosa effectively collect oral fluids and bacteria, with their dendritic extensions reaching towards the surface, often constituting a diverse group. LCs become mobile and mature in response to inflammatory cytokines and pathogen-associated molecular patterns (PAMPs) released by oral mucosal pathogens. LCs are primarily responsible for presenting exogenous antigen-derived peptides through major histocompatibility complex (MHC) class II presentation to CD4+ helper T-cells and MHC class I-restricted cytotoxic (CD8+) T-cell responses. Furthermore, LCs stimulate Natural killer (NK) cells by producing cytokines, including IL-12 (Hovav, 2014).

Activated DCs migrate to the GALT, where they interact with naïve T cells, thereby initiating the adaptive immune response. This migration and subsequent T cell activation are crucial for immune surveillance and responses in diverse body regions (Hovav, 2014). A deeper exploration is necessary to comprehend how DC activation in oral epithelia and their migration across different tissues contribute to coordinating immune responses and maintaining immune tolerance throughout the body.

Th17 cells, a subset of CD4+ T cells, specialize in orchestrating immune responses at mucosal surfaces by producing IL-17 and other cytokines that recruit neutrophils and enhance antimicrobial defenses. The coordinated efforts of neutrophils and Th17 cells illustrate the intricate collaboration between innate and adaptive immunity in safeguarding mucosal homeostasis and combating infections. This dynamic interaction plays a crucial role in maintaining the delicate balance between protective immunity and immune tolerance within mucosal environments (Suárez et al., 2021).

The intestinal epithelium is acknowledged as the principal axis of mucosal immunity, given that approximately 70% of the total lymphocyte population resides within the gastrointestinal tract (Suárez et al., 2021). The GI mucosal immune system comprises three principal components: the epithelial layer, lamina propria, and the MALT, known as gut-associated lymphoid tissue (GALT). The epithelium and lamina propria serve as the frontline defenses, while the GALT functions as the central hub where adaptive immune responses are instigated, specifically within the context of the GI tract (Wu et al., 2014). The epithelium layer is primarily composed of intestinal epithelial cells (IECs), organized into distinct structures known as villi. Other specialized cell types coexist, including Goblet cells, tuft cells, enteroendocrine cells, and M cells. Goblet cells lubricate and shield the intestinal epithelial surface by secreting mucins (Suárez et al., 2021). Control of paracellular permeability is crucial for preventing microbial invasion, and this is achieved through the regulation of various types of intercellular junctions (Suárez et al., 2021). The gut-associated lymphoid tissue consists of Peyer’s patches and isolated lymphoid follicles (Wu et al., 2014). Goblet cells also play a role as luminal antigen-presenting cells to CD103+ DCs, facilitating the differentiation of Tregs. DCs extend their dendrites into the epithelium to capture antigens and subsequently migrate to the lamina propria, subsequently draining to secondary lymphoid tissues. Meanwhile, specialized M cells, present in the epithelium of Peyer’s patches, facilitate the transfer of antigens to DCs, macrophages, and other APCs. In secondary lymphoid tissues, naïve T cells undergo activation upon interaction with APCs and subsequently migrate to the lamina propria (Wu et al., 2014).

Overall, the oral and gut immune responses exhibit both similarities and differences, sharing common features in terms of lymphoid tissues and immune cell populations. On the other hand, microbial compositions, exposure to antigens, immune tolerance mechanisms, and antibody production are different. It has been reported that immune cells present in the oral draining lymph nodes can transmigrate to various other lymphoid organs, including the gut (Morton et al., 2014). Consequently, oral inflammation could lead to the emergence of T cells reactive to oral pathogens. These T cells have the potential to migrate from the oral mucosa to the intestine, where they could become activated by
specific microbes, potentially leading to the onset of intestinal inflammation.

Furthermore, the gut mucosa employs alternative immune mechanisms to regulate the production of microbiota-responsive effector T cells. Within the gut, a specific group of ILCs expressing MHC class II molecules hinders the expansion of T cells specific to the microbiota. It is plausible that the oral mucosa and gingiva do not possess these intricate tolerogenic mechanisms observed in the gut. Therefore, acute inflammation in the oral mucosa might lead to the development of microbiota-responsive T cells with potential pathogenic characteristics. Th17 cells normally present in the gut do not induce disease, unlike their significant pathogenic role in promoting inflammation driven by commensal bacteria in the oral cavity in both mice and humans. Upon migration to the gut mucosa, these cells may undergo functional conversion into Th17/Th1 mixed phenotype cells capable of producing interferon-γ (IFN-γ). This conversion might contribute to the development of colitis triggered by oral-origin T cells. In summary, migratory Th17 cells represent a double microbiota and immune mechanism linking the oral and gut environments (Kitamoto et al., 2020).

A comprehensive understanding of the immune dynamics operating in these sites is crucial for unraveling the complex interplay between the oral and gut immune responses. Such insights could pave the way for novel strategies aimed at modulating immune reactions and designing interventions for diseases that involve both oral and gut components.

**Dysbiosis in the oral microbiota and host response**

The oralome comprises the dynamic interactions coordinated between the ecological community of oral microorganisms and the host within the oral cavity (Radaic and Kapila, 2021). The mouth harbors over 700 species of bacteria, as well as fungi, viruses, and protozoa, making it the second-largest and diverse microbiota after the gut (Peterson et al., 2009). Interindividual differences also exist, but the principal function of the microbiome is the same in every person. Oral commensal microorganisms help maintain microbiota balance, inhibiting pathogen attachment and invasion. They support oral health by competing for resources, producing antimicrobials, and modulating the host immune response (Wade, 2013). The oral microbiome has been extensively characterized using both cultivation and culture-independent molecular techniques, such as 16S rRNA cloning. The six most abundant phyla identified in the oral microbiome are: Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria, Spirochaetes, and Fusobacteria, which together comprise 96% of the taxa (Dewhirst et al., 2010).

Biofilms are organized communities of similar or different aggregated bacterial cells encased in a self-produced polymeric matrix adhered to a surface (Donlan, 2002). Oral biofilm formation begins with the initial attachment of planktonic “early colonizer” bacteria, primarily saccharolytic aerobes and facultative anaerobes that utilize glycopolymers and salivary mucins as nutrients. Streptococcus species, making up around 80%, dominate this early colonization phase (Kreth et al., 2009). Following the attachment, surface-bound bacteria undergo shifts in their metabolic and gene expression profiles resulting in the generation of extracellular polymeric substances, including polysaccharides, proteins, lipids, and extracellular DNA (Radaic and Kapila, 2021). *F. nucleatum* plays a key role as a “bridge” bacterium within the biofilm by adhering to a diverse array of co-aggregated species through surface adhesins (Coppenagen-Glazer et al., 2015).

The oral biofilm presents varying levels of oxygenation across its structural composition, providing the necessary environmental conditions for the attachment and proliferation of proteolytic obligate anaerobes, commonly referred to as “late colonizers” (Wade, 2013). Ahn et al. (Ahn et al., 2016) examined how *F. nucleatum* and *P. gingivalis* infections affect reactive oxygen species (ROS) production and bacterial complex formation in human oral cells. Despite both being anaerobic bacteria, *F. nucleatum* and *P. gingivalis* have different oxygen tolerances and are commonly found together at infection sites. ROS production primarily relies on NADPH oxidases (NOX) enzymes. Notably, *F. nucleatum* exhibits a 20-fold higher NOX activity compared to *P. gingivalis*, giving it a much greater ability to utilize oxygen molecules (Diaz et al., 2002). Similarly, the presence of *F. nucleatum* increases *P. gingivalis* attachment to human gingival fibroblasts by a factor of ten. Ultimately, the study (Ahn et al., 2016) revealed that *F. nucleatum* contributes to the co-aggregation of *P. gingivalis*, promoting the formation of bacterial complexes.

The most prevalent oral diseases that are dental caries and periodontitis are both associated with a disruption in the balance of oral microbiota. Dental caries is a biofilm disease caused by multiple microorganisms, influenced by dietary habits leading to the demineralization of tooth enamel through acid production by oral pathogens fueled by dietary sugars (Bowen et al., 2018). Periodontitis is a chronic inflammatory disease initiated by a dysbiotic biofilm in the gingival pocket. Its polymicrobial nature is driven by variations in subgingival microbiota composition and interactions, rather than new bacterial colonization (Aas et al., 2005; Bowen et al., 2018). The subgingival microbiota is typically dominated by *Firmicutes*, with fewer *Actinobacteria* and *Bacteroidetes*. In contrast, periodontitis is characterized by the presence of *P. gingivalis*, *Tannerella forsythia* (*T. forsythia*), *Treponema denticola* (*T. denticola*), and *F. nucleatum* (Paster et al., 2001; Donlan, 2002). Dominant species in the subgingival biofilm play a crucial role in driving dysbiosis and inflammatory response, in periodontitis development (Aas et al., 2005).

The aggregation of bacteria triggers local inflammation, increasing crevicular gingival fluid flow, leading to bleeding, and providing protein-rich nutrients. This promotes the growth of the Gram-negative anaerobes (Hoare et al., 2019). Within the periodontal pocket, the initial host response to the dysbiotic subgingival community involves NK cells, neutrophils, and granulocytes, initiating early inflammation. Subsequently, lymphocytes infiltrate, facilitating antigen presentation to dendritic cells. T cells, including CD8+ and CD4+ cells, generate a proinflammatory environment rich in cytokines like TNF-α, IL-1, IL-4, IL-10, IFN-γ, and transforming growth factor β (TGF-β).
(Pihlstrom et al., 2005). This inflammatory cascade leads to changes in the subgingival environment, contributing to shifts in the subgingival biofilm composition that drive the progression of periodontitis. Interestingly, certain periodontal pathogens directly play a role in promoting chronic inflammation by activating specific intracellular pathways (Hoare et al., 2019).

Several studies have linked periodontitis to systemic diseases, such as cardiovascular, diabetes mellitus, respiratory disease, adverse pregnancy outcomes (Geisinger et al., 2016), and increased risk of pancreatic and colon cancers (Pihlstrom et al., 2005; Momen-heravi et al., 2018; Kim et al., 2019; Koliarakis et al., 2019). Nevertheless, the available evidence about remains limited, necessitating further research to elucidate the nature of this connection and the underlying mechanisms. Potential mechanisms linking oral infections to systemic conditions encompass the spread of oral pathogens through transient bacteremia and intracellular infection in immune cells, the release of oral microbial toxins into the circulation at the site of injury, and interactions between oral pathogens and the host’s immune response, resulting in systemic inflammation (Xiaojing et al., 2000). Further prospective investigations are imperative to gain a comprehensive understanding of the causal mechanisms linking oral dysbiosis and its potential impact on the development of colon cancer (Figure 1).

**The circulation as a key component in the oral-gut axis**

The mechanisms underlying bacterial leakage into the bloodstream remain unknown, with hypotheses ranging from dendritic cell processes, and oral pathogen-immune cells transport to dysfunctional epithelial junctions. Some studies reported pathogen *P. gingivalis* in the bloodstream of healthy individuals and those with periodontitis after activities like tooth brushing, flossing, or chewing food (Horliana et al., 2014). Tsukasaki et al. observed the formation of oral bacterial colonies in liver and spleen cells following prolonged ligature placement around a tooth in a murine model of periodontitis. The bacterial species were also present in the oral cavity but were notably absent in fecal samples. This suggests that the systemic spread of oral bacteria occurs when the oral barrier is compromised (Tsukasaki et al., 2018).

Severe periodontitis is associated with elevated levels of pro-inflammatory mediators and recruitment of immune cells to the affected site, including increased neutrophil numbers in the blood. The inflammatory response caused by this inflammatory chronic disease can disrupt the tight junctions between endothelial cells, leading to increased permeability of blood vessels. This increased permeability allows inflammatory mediators, bacterial components, and immune cells to enter the vessel walls and surrounding tissues, promoting systemic inflammation (Nakajima et al., 2010). Periodontal pockets surface area presents a large bacterial biofilm accumulation, ranging from 50 cm² to 200 cm², allowing bacterial products such as lipopolysaccharides or proteases to diffuse into the blood stream (Hujoel et al., 2001).

Via transient bacteremia, bacterial infection is cleared by the immune system. However, some can evade the immune response and survive in the bloodstream. It is known that leukocytes are not effective in recognizing and engulfing bacteria in high-velocity liquids, although erythrocytes can attract bacteria by electrical charges on their surface and kill them by oxidative attack. Despite

![FIGURE 1](image-url)
This, some bacteria can survive in the liver or spleen promoting low-grade inflammatory response and inducing cytokine secretion by other tissues/organs (Minasyan, 2014). Bacterial reservoir inside erythrocytes provides a long-term bacterial survival causing antibiotics ineffectiveness and immune reactions (Minasyan, 2017).

Individuals with periodontitis have elevated circulatory levels of pro-inflammatory cytokines, such as IL-1, IL-6, and TNF-alpha. Once in the circulation, these cytokines can then exert their effects on other tissues and organs, contributing to systemic inflammation and potentially leading to the development of comorbidities at distant sites (Nakajima et al., 2010). In addition, these cytokines are released in response to bacterial pathogens in the gums, acting as immune cells recruiters at the infection site.

Recent research suggests a potential link between immune cells activated in the oral cavity and the development of gut inflammation. It proposes that these immune cells can migrate from the oral cavity to the gut, where they can play a role in promoting inflammation. This is referred to as an indirect pathway, as the immune cells themselves do not directly trigger gut inflammation. Instead, they contribute to the inflammatory process by priming other cells or interacting with various components of the immune system. An alternative idea, called the Trojan Horse hypothesis, suggests that immune cells could transport pathogens throughout the body.

In periodontitis, P. gingivalis found intracellularly in DCs, which affects the differentiation of these immune cells while using them as vehicles to enter the circulation and disseminate to distant organs (Carrion et al., 2012). On the other hand, recent research has shown that human neutrophils can carry and spread viable F. nucleatum after phagocytosis by using both an in vitro microfluidic device and a zebrafish model. The researchers previously found that F. nucleatum can suppress immune responses in neutrophils and survive within them, suggesting that intracellular bacterium can avoid the host’s immune defenses and spread from the oral cavity (Ellett et al., 2023). Additionally, macrophages have been proposed as a communication linking the oral cavity to distant sites. P. gingivalis has been shown to invade and persist within resident macrophages. Moreover, studies in animal models indicate that P. gingivalis-infected macrophages can produce cytotoxic extracellular vesicles, damaging distant organs. P. gingivalis has evolved intricate strategies to evade macrophage antimicrobial defenses. Understanding P. gingivalis-macrophage interactions may yield insights into disease mechanisms and potential therapies (Lin et al., 2022). Moreover, F. nucleatum can also survive within the cytoplasm of infected macrophages by the expression of indoleamine-2,3-dioxygenase, which can amplify the impairment of peripheral blood lymphocyte function. This dual effect allows F. nucleatum-infected macrophages to evade cellular death (Xue et al., 2018). These mechanisms suggest that common oral bacteria could connect oral and systemic diseases by utilizing the body’s immune cells for transport. Further investigation should be done to study if this mechanism for immune cells-mediated bacterial dissemination could apply to other oral bacterial species and potentially contribute to the link between oral and systemic diseases.

From oral bacterial translocation to gut microbiota dysbiosis

The study of the oral and gut microbiomes has mostly been conducted in an organ-specific manner, which overlooks the fact that the mouth and gut are anatomically continuous regions; moreover, both regions are chemically connected through the passage of salivary fluids and digested food through the GI tract (Figure 2A). Saliva is estimated to contain about 10^8 bacteria per ml, which would result in people orally ingesting as many as 10^{12}-10^{13} bacteria day (Von Troil-Linden et al., 1995). The oral cavity is the entry point for the digestive tract and is continuously exposed to various external factors, including microorganisms, nutrients, and xenobiotics. The Human Microbiome Project (Peterson et al., 2009) showed that more than half of the total bacteria in the human body are present in the GI tract (29%) and the oral cavity (26%). Interestingly, oral bacteria have also been found in distant sites such as the pancreas and gut, indicating direct cross-talk between microbiota at different locations (Chung et al., 2021).

The transmission and invasion of oral pathogens in the gut require the presence of at least two essential factors, one within the oral microbiota and the other within the gut microbiota. The first prerequisite is an elevated number of oral bacteria in the mouth caused by oral dysbiosis which enhances the likelihood of gut translocation. The second prerequisite requires the disruption of the intestinal mechanisms to resist colonization by opportunistic bacteria, which is typically conferred by gut dysbiosis. This disruption may be a necessary step to allow oral pathobionts, which have successfully traversed the gastric barrier, to establish colonization within the gut (Kitamoto and Kamada, 2022). Kageyama et al. (Kageyama et al., 2023) investigated both saliva and stool samples using 16S rRNA gene amplicon analysis and the amplicon sequence variant (ASV) approach to explore the translocation of oral bacteria to the gut. The study found that ASVs shared with an individual’s salivary microbiota were present in the gut microbiota of 72.9% of subjects. Notably, the sharing and similarity of oral-gut microbes were more pronounced within an individual than across different individuals (Schmidt et al., 2019).

Overall, the acidic environment of the stomach plays an important role in protecting the host against potential infections from ingested microbes, as well as in regulating the composition of the gut microbiota. Research findings indicate that over 99% of oral microbes that are ingested are rendered inactive or eliminated as they traverse the stomach. The high acidity of the stomach environment makes it inhospitable for the survival of most bacteria. However, some microorganisms, such as Helicobacter pylori (H. pylori), are adapted to survive in the acidic environment of the stomach and can colonize the gastric mucosa (Hunt et al., 2015). Long-term colonization by H. pylori can cause achlorhydria and decreased acid secretion, leading to changes in the gastric microbiota (Hunt et al., 2015). Additionally, gut colonization by oral-origin bacteria, Streptococcus and Veillonella, is observed in patients with gastric achlorhydria caused by long-term use of proton pump inhibitors (Schmidt et al., 2019). Furthermore,
aging displayed a significant correlation with an increased abundance of oral bacteria in the gut microbiota. This association could be attributed to the diminished barrier function of the gastrointestinal tract due to aging-related alterations (Kageyama et al., 2023).

The gut microbiota contains 1000-1150 bacterial species, some of the most abundant groups are Bacteroidetes, Dorea/Eubacterium/Ruminococcus, Bifidobacteria, Proteobacteria, and Streptococci/Lactobacilli (Qin et al., 2010). Maintaining a balanced relationship between the host and commensal microorganisms is crucial for proper functioning. The host provides a fit environment and nutrients for the microorganisms to thrive, while the commensal bacteria contribute to various physiological functions and offer protection against pathogens invasion. This intricate balance is achieved through complex interactions between the host and microbiota (Lee et al., 2013). Oral administration of P. gingivalis in C57BL/6 mice alters gut microbiota with an increase in Bacteroides and Staphylococcus and a decrease in commensal species such as Firmicutes and Lactobacillus when compared to sham-inoculated mice (Nakajima et al., 2015). Lactobacillus species are recognized as probiotics, known for their ability to promote digestive health and support immune function. Decreased levels of

FIGURE 2
(A) Microbiome research has traditionally focused on specific organs, neglecting the continuous anatomical connection between the mouth and gut. (B) The development of an F. nucleatum-P. gingivalis bacterial complex is central to the formation of oral biofilm. F. nucleatum acts as a bridge attaching early colonizers like Streptococcus spp. Actinomyces spp. P. gingivalis and others. (C) In the periodontal pocket, initial host response is characterized by the infiltration of NK cells, neutrophils, granulocytes, and lymphocytes to present antigens to dendritic cells. T cells secrete cytokines such as TNF-α, IL-1, IL-4, IL-10, IFN-γ, TGF-β, RANK-L. (D) Bacteria, metabolites and inflammatory factors can travel from the oral cavity to other parts of the body into the blood vessels by increased endothelial permeability. (E) Oral bacteria, particularly F. nucleatum and P. gingivalis, can infiltrate intestinal cells, provoke pro-inflammatory cytokine production, and induce a pro-inflammatory environment. (F) They activate various pathways linked to inflammation and cancer development and can disrupt the colon’s epithelial barrier, increasing permeability to opportunistic bacteria and promoting chronic inflammation, angiogenesis, and cancer progression. (G) A diverse array of immune cells are recruited to the tumor site, including myeloid-derived suppressor cells (MDSCs). MDSC block T cells and NK cells activation, and recruitment and activation host suppressor cells such as FoxP3+ regulatory T cells.
probiootic bacteria can disrupt the gut balance, leading to dysbiosis and compromised intestinal barrier. This may increase intestinal permeability, allowing toxins and bacteria to enter the bloodstream and potentially trigger inflammation and health problems (Arimatsu et al., 2014; Gao et al., 2015; Kobayashi et al., 2020).

In mice administered with P. gingivalis, the expression of genes responsible for intestinal alkaline phosphatase (Akp3) and the tight junction protein (Tjp1) in the small intestine was both decreased. (Arimatsu et al., 2014). Additionally, Kobayashi et al. (Kobayashi et al., 2020) found that the effects on the gut microbiota were not solely due to P. gingivalis, but also due to the other oral bacteria such as Streptococcus mitis, Streptococcus salivarius, Porphyromonas nigrescens. Interestingly, the study found that dysbiosis led to a decrease in the abundance of Th17 cells, a type of immune cell involved in protecting against pathogens, and a decrease in the production of immunoglobulin A (IgA), an antibody important for mucosal immunity in the small intestine. IgA and Th17 cells play important roles in maintaining gut homeostasis and protecting against pathogen invasion. With the growing body of evidence highlighting the significance of interactions between gut microbiota and immune cells in maintaining the immune system balance, Morton et al, utilizing Kaede transgenic mice, observed a bidirectional immune cell trafficking between the gut and other organs (Morton et al., 2014). The migration of oral immune cells to the gut is a pivotal aspect of the mouth-gut axis in intestinal inflammatory conditions, such as colitis. During periodontal inflammation, Th17 cells are generated in the oral draining lymph nodes, specifically recognizing oral bacteria. Upon reaching the gut, these orally primed Th17 cells can be activated by translocated oral pathobionts, contributing to colitis development (Kitamoto and Kamada, 2022).

Interestingly, Nagao et al. (Nagao et al., 2022) hypothesized that oral pathogens in the oral cavity, known as pathobionts, could migrate to the intestinal tract and trigger an immune response that exacerbates periodontitis. They infected mice with P. gingivalis showing that the infection provoked pathobiont migration from the oral cavity to the intestine where they were recognized by DCs triggering the Th17-type immune response, leading to the production of IL-17 and other pro-inflammatory cytokines. Interestingly, the study revealed that Th17 cells can migrate from the intestine to the oral cavity upon oral infection. This suggests that the migration of Th17 cells may be one of the mechanisms by which the gut microbiota can affect periodontitis. These findings suggest that the oral-gut axis is bidirectional, and that the gut microbiota can affect the homeostasis of the mouth. It is crucial to further explore and understand the specific mechanisms involved in the communication between the oral and gut microbiota to gain insights into their collective impact on oral, gut, and systemic health. The examination of host-microbe interactions in murine models has played a crucial role in unraveling the gut-axis relationship. However, given the variations between the microbiotas of mice and humans, it is imperative to invest additional effort into comprehending how these findings apply to humans (Kitamoto and Kamada, 2022).

An imbalance in the oral microbiome can not only lead to oral disorders such as dental caries and periodontitis, but also systemic diseases such as irritable bowel syndrome, IBD (Tszunou et al., 2021; Kitamoto and Kamada, 2022), and colorectal cancer (CRC) (Flemer et al., 2018; Tortora et al., 2022). Metagenomic sequencing analysis revealed the presence of four species enriched in tumor samples in CRC patients with oral cavity origin: Porphyromonas asaccharolytica, F. nucleatum, Prevotella intermedia, and Parvimonas micra. Remarkably, these bacteria were found to form mutually beneficial networks within the microbial community where F. nucleatum occupied a central position within this network, suggesting a pivotal role for oral bacteria in shaping these interactions (Dai et al., 2018) similar to the role that F. nucleatum plays in oral biofilm as a “bridge” bacteria between early and late colonizers. Bacterial biofilms that invade the mucus layer can be observed on the colon mucosa in around 50% of CRC patients and roughly 13% of individuals without the disease (Tomkovich et al., 2019). These findings suggest that the composition and arrangement of the microbiota, rather than the overall health status of the human donor, are connected to the development of tumors. Additionally, given the association between bacterial biofilms and CRC, it is crucial to gain insight into the functional significance of bacterial organization within the colon mucosa in relation to CRC. The next phases of research should involve the retrieval of oral and colonic mucosal biofilms for comprehensive analyses, focusing on bacterial community composition and functional profiling to identify commonalities shared among these microbiota, providing valuable insights into the mechanisms underlying their role in disease development.

Oral pathogens in the carcinogenesis of colon cancer

The development and progression of colon cancer is associated with several factors including genetics, environmental, diet, lifestyle, and microbiota (Tortora et al., 2022). Several studies have highlighted the relationship between the gut microbiota and colon cancer (Cho et al., 2014; Fulbright et al., 2017; Koliarakis et al., 2019; Tortora et al., 2022). Oral diseases, such as periodontitis, involve chronic inflammation triggered by a multispecies bacterial community in the subgingival region in the mouth. While inflammation mainly occurs in the oral cavity, studies reveal that inflammatory agents, subgingival bacteria, and their components can disseminate, contributing to extraoral diseases like cancer (Hoare et al., 2019; Koliarakis et al., 2019). The association between periodontal disease and CRC was established in the Nurses’ Health Study, in which women with moderate or severe periodontitis were at moderately increased risk of developing CRC (Momen-heravi et al., 2018; Nwizu et al., 2020).

As previously mentioned, the formation of an F. nucleatum-P. gingivalis bacterial complex is central to the pathogenesis of periodontitis (Ahn et al., 2016). F. nucleatum acts as a bridge attaching other early colonizers like Streptococcus spp, Actinomyces spp, P. gingivalis and others (Figure 2B). In the periodontal pocket (Figure 2C), the initial host response is marked by the infiltration of NK cells, neutrophils, and DCs,
followed by the subsequent influx of lymphocytes after antigen presentation. Recruited T cells subsequently contribute to the response by secreting cytokines like TNF-α, IL-1, IL-4, IL-10, IFN-γ, TGF-β, and receptor activator of nuclear factor kappa-B ligand (RANK-L) (Hoare et al., 2019). One of the complications of periodontal disease is the migration of bacteria from the oral cavity to other parts of the body. This occurs because there is an increase in the number of oral bacteria in the sub-gingival biofilm in intimate contact with the ulcerated gingiva creating an entry point for oral bacteria into the bloodstream, facilitating their dissemination to remote locations such as the colon (Figure 2D) (Nwizu et al., 2020). These bacteria are mostly multispecies with high numbers of Clostridium, Peptostreptococcus and Fusobacterium. (Hoare et al., 2019) In the gut (Figures 2E-G), oral bacteria, can invade and adhere to intestinal epithelial cells, increase the production of pro-inflammatory cytokines, and contribute to a pro-inflammatory microenvironment (Bashir et al., 2016). These findings imply that F. nucleatum may exert a significant indirect influence on the development of periodontal diseases by promoting the proliferation of P. gingivalis, potentially contributing to the etiological factors involved. The ability of F. nucleatum to adhere to both host cells and other bacterial species enhances its capacity to form complex microbial communities and promote the colonization of harmful bacteria in the gut. These findings highlight the interconnected nature of the oral-gut axis and suggest that F. nucleatum and P. gingivalis may have significant influence not only in the oral cavity but also in gut (Figure 2F) (Castellarin et al., 2012). Interestingly, F. nucleatum and P. gingivalis are known to synergistically promote oral cancer progression (Binder Gallimidi et al., 2015). Similar studies should be proposed to study potential synergy between both oral pathogens and gut dysbiosis in colon cancer.

In the study by (Tsuzuno et al., 2021) P. gingivalis amplified gastrointestinal inflammation by directly engaging with the intestinal epithelial barrier in a susceptible host and showed a higher colitogenic potential compared with other periodontal pathogens such as Prevotella intermedia and F. nucleatum in mice. In summary, the potential mechanisms through which P. gingivalis reduces the protein level of ZO-1 in vivo, also known as tight junction protein-1, encompass several interconnected events: detachment of intestinal mucus bacterial invasion into intestinal epithelial cells, and cytosolic degradation of ZO-1 (Tsuzuno et al., 2021). The study noted that following the intravenous administration of P. gingivalis, CD4+ T cells stimulation heightened the inflammatory response among colon and lamina propria lymphocytes, resulting in an elevated Th17/Treg ratio. These results suggest that the increase in Treg cells could potentially counterbalance the escalation of intestinal tissue inflammation induced by P. gingivalis. However, further research is imperative to gain a comprehensive understanding of this phenomenon (Li et al., 2022). There is growing evidence suggesting a potential link between chronic inflammation, such as colitis, and an increased risk of developing colon cancer. This connection underscores the intricate interplay between inflammation caused by dysbiosis and the development of colon cancer.

Numerous investigations have shown dysbiosis-related gut microbiota associated with F. nucleatum infection in the tumor tissues of CRC patients (Gao et al., 2015; Dai et al., 2018; Koliarakis et al., 2019; Huh and Roh, 2020; Yu et al., 2022). The argumentative aspect revolves around determining whether F. nucleatum contribution to CRC is correlational or causational. Compelling evidence substantiates both hypotheses. Numerous studies on the gut microbiota have consistently demonstrated a significant prevalence of F. nucleatum within the tumor tissue and fecal samples of individuals with CRC (Wang et al., 2021). Etiological investigations have further elucidated the role of F. nucleatum as a bacterium that promotes carcinogenesis at various stages of CRC development. Two different groups were the first to show an increased abundance of this oral pathogen in CRC tissues when compared to normal tissues (Castellarin et al., 2012; Kostic et al., 2013). A recent study highlighted that identical strains of F. nucleatum were identified in both the saliva and colorectal tumors of patients with CRC (Komiya et al., 2019), in addition to the findings that oral F. nucleatum translocate to the colon through the hematogenous route (Abed et al., 2020). This suggests that CRC-associated F. nucleatum likely originates from the oral cavity, underscoring a potential link between oral bacteria and the development of CRC.

Fusobacterium is numerous in human adenomas, proposing an early role in colon carcinogenesis (McCoy et al., 2013). Kostic et al. (Kostic et al., 2013) investigated the association of F. nucleatum in stool and colon samples from human colorectal adenomas and adenocarcinomas. It assessed the impact of F. nucleatum on cancer progression and tumor-related inflammation in ApcMin/+ mouse model of intestinal tumorigenesis. The bacterium amplified tumor numbers and led to accelerated tumorigenesis in both the small intestine and colon, infiltration of distinct myeloid cell subsets into the tumors, and an NF-κB-driven proinflammatory profile, like F. nucleatum-positive colorectal carcinomas in humans. APC gene mutations typically manifest as early molecular alterations during the transition of epithelial cells into adenomas (Kostic et al., 2013). Hence, probably the early somatic mutations are responsible for tumor initiation, which occur before F. nucleatum accumulation in the tissue. This concept is further substantiated by the mechanisms described by which this bacterium contributes to CRC in the study by Rubinstein et al, where they showed that F. nucleatum through virulence factor FadA adhesin can adhere and invade tumor cells, and promote oncogenic and inflammatory responses to promote progression of CRC (Rubinstein et al., 2013).

Another study using the same ApcMin/+ mouse model showed that F. nucleatum induced DNA damage and cell growth in CRC by activating the E-cadherin/β-catenin pathway (Guo et al., 2020). These genetic alterations play a role in disruption of the epithelial barrier and the mucous layer facilitating the infiltration of F. nucleatum and other opportunistic bacteria, allowing them to establish themselves within the tumor microenvironment. However, it is essential to acknowledge that the murine model utilized lacks full characterization, potentially limiting its ability to entirely recapitulate all aspects of F. nucleatum-associated CRC observed in humans. These collective findings suggest that F. nucleatum infection has been implicated as an added
environmental risk factor for CRC by promoting a proinflammatory microenvironment conducive to the progression of colorectal neoplasia. Moreover, the high abundance of *F. nucleatum* in CRC is associated with poorer survival (Flanagan et al., 2014; Mima et al., 2016; Komiya et al., 2019) and recurrence after chemotherapy (Yu et al., 2017).

There are three major virulence factors that contribute to the promotion of CRC: LPS, the adhesin FadA, and autotransporter protein Fap2 (Kostic et al., 2013; Abed et al., 2016; Lee et al., 2019). Fap2 operates as an inhibitor, dampening the tumor-killing effectiveness of T cells and NK cells. It achieves this by directly triggering the inhibitory receptor known as T-cell immunoglobulin and immunoreceptor tyrosine-based inhibitory motif domain (TIGIT), thereby contributing to a mechanism of tumor-immune evasion (Gur et al., 2015). Furthermore, Fap2 exhibits binding affinity towards a carbohydrate structure, specifically D-galactose-β(1→3)-N-acetyl-D-galactosamine (Gal-Gal/NAC), which is prominently present in CRC. Moreover, metabolic specialization may underlie the distinctive impact of *F. nucleatum* within the tumor milieu. As an asaccharolytic bacterium, it refrains from competing for glucose, a favored substrate in cancer metabolism, thereby conferring upon it a competitive advantage distinct from other microbes residing in the tumor microenvironment (Flynn et al., 2016).

Additionally, Fap2 triggers the release of proinflammatory cytokines, namely IL-8 and C-X-C motif chemokine ligand 1 (CXCL1), which enhances the migration of cancer cells (Casasanta et al., 2020). This effect contributes to the promotion of CRC cell invasiveness and potentially impacts disease progression. Synergistically, LPS induces the secretion of a range of inflammatory cytokines, including IL-2, IL-6, IL-8, IL-10, IL-17, TNF-α, and upregulates NF-κB levels. On the other hand, FadA facilitates the attachment to epithelial and endothelial cells while also triggering inflammation. This interaction facilitates the attachment of *F. nucleatum* to CRC cells through the hematogenous route, contributing to the association between the oral pathogens and colon carcinogenesis (Abed et al., 2016). Furthermore, this adhesin activates pro-carcinogenic pathways directly within colon cancer cells, specifically by initiating E-cadherin-β-catenin signaling (Ito et al., 2015; Nosho et al., 2016; Rubinstein et al., 2019). In summary, *F. nucleatum* plays a role in promoting the initiation and progression of CRC through various mechanisms, including its localization within the colorectal environment, its ability to proliferate, promote immune suppression, facilitation of metastasis, and contribution to chemoresistance. These multifaceted effects collectively contribute to the complex relationship between *F. nucleatum* and CRC development.

*F. nucleatum* and *P. gingivalis* are involved in the direct activation of inflammation and carcinogenesis through multiple pathways. Although the precise role of these oral pathogens in the carcinogenesis of colon cancer is still unclear (Castellarin et al., 2012). Some studies suggest that *P. gingivalis* and *F. nucleatum* may directly interact and attach with intestinal cells and can disrupt the integrity of the epithelial barrier lining the colon. This disruption can lead to increased permeability of the barrier, allowing other opportunistic bacteria and their byproducts to enter the underlying tissues and alter the microbial communities in the gut (Figure 2F). This promotes chronic inflammation creating a favorable environment for cancer development and progression. These bacteria also stimulate angiogenesis, the formation of new blood vessels, which is essential for tumor growth as it supplies nutrients and oxygen to cancer cells, contributing to the growth and spread of cancer cells, bacterial metabolites, and inflammatory factors. Furthermore, they interact directly with colon cells, promoting signaling pathways that encourage cell survival, proliferation, and resistance to cell death – all hallmarks of cancer cells (Fulbright et al., 2017).

In response to the abnormal growth of colonic cells and the oral pathogens infection, a diverse array of immune cells are recruited to the tumor site (Figure 2G). Among them are tumor-infiltrating lymphocytes (TILs), including CD4+ helper T cells and CD8+ cytotoxic T cells, pivotal for recognizing and combating cancer cells. Macrophages, exhibiting both pro-inflammatory (M1) and anti-inflammatory (M2) functions, influence tumor progression and immune responses. DCs play a crucial role by capturing antigens and initiating immune reactions. NK cells directly target and eliminate cancer cells, restricting tumor growth. Conversely, myeloid-derived suppressor cells (MDSCs) can hinder immune responses by inhibiting T cell activity, allowing tumors to evade immunity. The intricate interplay and functions of these immune cells within the tumor context remain a focal point of research, offering potential avenues for innovative cancer treatments (Yin et al., 2020).

*P. gingivalis* is often recognized as highly proficient in evading and undermining the immune system, employing various tactics to elude and undermine immune defenses. While possessing an array of virulence factors, including gingipains, LPS, and fimbriae, *P. gingivalis* is known for invading epithelial, fibroblast, endothelial and specific immune cells. Leukocytes in circulation may function as a "Trojan horse" facilitating the oral bacterial dissemination through the bloodstream and spreading to the gut (Carrion et al., 2012). A recent review described the potential mechanisms of translocation of *P. gingivalis* from the oral mucosa by circulatory dissemination, like the Trojan Horse mechanism (de Jongh et al., 2023). A study showed that a *P. gingivalis* strain was able to significantly avoid phagocytosis and escape from macrophages (Werheim et al., 2020). Aside from macrophages, DCs are the other phagocytic cells that may play a role in blood dissemination of *P. gingivalis* (de Jongh et al., 2023). One mechanism is by engaging its fimbrial proteins with complement receptor 3 (CR3). CR3 is a major receptor for the phagocytosis of opsonized particles (Carrion et al., 2012). The clinical significance of this study (Carrion et al., 2012) lies in the remark that *P. gingivalis* strains expressing Mfa1+ infect DCs within oral mucosal tissues and the bloodstream of individuals with periodontitis. Subsequently, these infected DCs can spread to distant sites where angiogenesis occurs. El-Awady et al. investigated the mechanism of autophagy evasion by *P. gingivalis* promoting the survival within human monocyte-derived dendritic cells (MoDCs). This process is facilitated by its glycoprotein fimbriae, specifically Mfa-1, which interacts with the C-type lectin DC-SIGN found on DCs. The other primary fimbriae, known as
FimA, a TLR2 agonist targets thereby inhibit the autophagic degradation of *P. gingivalis*, which may also facilitate its dissemination to distant sites within DCs (El-Awady et al., 2015). Furthermore, another investigation revealed that macrophages subjected to differentiation in the presence of IL-34, a predominant cytokine in the oral gingival environment, exhibit significantly diminished capacity to eliminate engulfed *P. gingivalis* (Almarghlanl et al., 2022). In conclusion, *P. gingivalis* exhibits a remarkable capacity to evade immune defenses and establish persistence within diverse cell types, notably DCs and macrophages. Nevertheless, the specific role of *P. gingivalis* on intestinal inflammation has yet to be fully investigated. Additionally, the precise mechanisms of translocation from the oral mucosa to the gut remain unclear.

Despite a continuous influx of numerous oral bacteria into the gastrointestinal tract, the impact of periodontal pathogens on intestinal inflammation remains as an uncharted territory (Tszun et al., 2021). In Figure 3 we depicted a comprehensive landscape of *P. gingivalis* and *F. nucleatum* virulence factors in the development and progression of CRC. Gingipain proteases produced by *P. gingivalis* can activate NF-κB and MMP-9, both significant for tumor invasion and metastasis. Furthermore, proteins associated fimbriae aid in the formation of biofilms, as well as the invasion and dissemination of the bacteria through blood DCs. Notably, Mfa1 fimbriae have been demonstrated to promote oncogenic signaling, recruiting tumor-infiltrating myeloid cells and promoting the expansion of immune-suppressive MDSC. Additionally, FimA can interact with Streptococci, Actinomyces, and Treponema spp and promote the release of the inflammatory cytokines TNF-α, IL-6, metalloproteinase-8 (MMP–8) and MMP–9 via the TLR4/NF-κB signaling pathway (Hoare et al., 2019). *P. gingivalis* binds to protease activated receptor and cleaves the MMP-9 active form, which subsequently facilitates tumor cell invasion and migration (Hoare et al., 2019; Yu et al., 2022). *P. gingivalis* displays antiapoptotic properties within epithelial cells through various mechanisms. These include the enhancement of both the PI3K/AKT and JAK/STAT3 signaling pathways, alongside the inhibition of caspase 3 and caspase 9 activity which are the final stage of the pathways of apoptosis. Additionally, *P. gingivalis*-triggered Akt–STAT3 signaling facilitates the expression of programmed cell death ligand 1 (PD-L1) while dampening CD8+ T-cell functionality, presenting another potential immune mechanism employed by the bacterium to further cancer progression (Yu et al., 2022). Also, *P. gingivalis* increases cell proliferation by modifying the activity of p53, cyclins and the WNT/β-catenin and MAPK/ERK pathways (Mu et al., 2020).
Therefore, these results offer a valuable understanding of the multiple mechanisms employed by *P. gingivalis* to suppress the immune response. By elucidating the complex components and identify specific molecular targets of *P. gingivalis*-mediated immunosuppression, we can develop innovative strategies to augment the immune response that can be leveraged to enhance the effectiveness of immunotherapeutic approaches in colon cancer, ultimately benefiting patient outcomes.

Other oral bacteria, such as *T. denticola* and *T. forsythia* can also trigger the death of epithelial cells. Damaged cells release chemical signals in the form of chemokines, cytokines, and pro-inflammatory molecules that attract immune cells to the site, creating a specific environment for both innate and adaptive immune responses including MDSCs recruitment. If the cause of this process is not resolved, it can lead to chronic inflammation and persistent tissue damage (Jun et al., 2017). Both periodontitis and colon cancer have been associated with elevated MDSC levels, and increased levels of TNF-α, IL-1β, IL-6 and MMP-9 are also common factors in both conditions. In individuals without underlying health issues, hematopoietic stem cells undergo a developmental process, progressing into immature myeloid cells (IMCs). These IMCs can subsequently differentiate into granulocytes, monocytes, or mature into macrophages or DCs. In contrast, cancer patients experience a disruption in the maturation of IMCs, resulting in an elevated population of MDSCs. Importantly, both *F. nucleatum* and *P. gingivalis* have been demonstrated to induce the recruitment and proliferation of MDSCs (Kostic et al., 2013; Nosho et al., 2016; Sakamoto et al., 2021). These cells can suppress the activity of T cells, leading to T cell exhaustion and impaired anti-tumor immune responses (Cui et al., 2021). A previous report by Chen et al. (Chen et al., 2020) found that the expression of NOD2, a gene coding for proteins associated with MDSCs, was upregulated in intestinal epithelial cells infected with *F. nucleatum*. *F. nucleatum* directs its actions towards caspase activation via NOD2, subsequently triggering the IL-17F/NF-kB pathway both in vivo and in vitro models. This cascade of events leads to damage to the intestinal epithelium and the upregulation of IL-1β, IL-6, IL-17F, and TNF-α (Chen et al., 2020). The activation of the TLR4/NF-kB signaling pathway by gingipain proteases and FimA, bacterial proteins produced by *P. gingivalis*, leads to an increased release of inflammatory cytokines which can have implications for the expansion and activation of MDSCs (Cai et al., 2019). Activated MDSCs within the tumor microenvironment can suppress both innate and adaptive immune responses in several ways. MDSCs inhibit the activity of T cells and NK cells and increase Treg numbers. MDSCs can also release reactive oxygen and nitrogen species, which can damage T cell receptors and interfere with T cell signaling (Figure 3) (Tang et al., 2021).

Based on the evidence, we propose that tumors exhibit elevated expression of genes related to MDSCs when co-infected with *F. nucleatum* and *P. gingivalis*. This co-infection synergistically enhances the presence of MDSCs within the tumor microenvironment, promoting the secretion of pro-inflammatory cytokines, producing sustained chronic inflammation, and leading to intestinal damage. These effects can lead to augmented gut permeability, enabling the infiltration of additional pathogens into the underlying tissues, thereby promoting the development of tumors.

MDSCs can also express PD-L1, which binds to PD1 on T cells, and causes secretion of IL-10 and TGF-β, which stimulate Treg activation and expansion (Tang et al., 2021). T cell exhaustion is a state of dysfunction characterized by the progressive loss of effector functions and sustained expression of inhibitory receptors, such as PD-1. Exhausted T cells are less capable of recognizing and eliminating cancerous cells. Targeting the recruitment and activity of MDSCs holds promise as a potential therapeutic strategy to alleviate T cell exhaustion and enhance anti-tumor immune responses. Through the suppression of MDSC recruitment or function, there exists a potential opportunity to recover T cell activity and support the role of the immune system in identifying and eradicating cancerous cells. This could involve targeting specific molecules or signaling pathways involved in the recruitment process. *F. nucleatum* and *P. gingivalis* have been shown to recruit MDSCs, leading to T cell exhaustion and impaired anti-tumor immune responses. Further research is needed to identify specific molecular targets and develop effective therapeutic interventions aimed at disrupting the immunosuppressive effects of MDSCs in the context of CRC and other cancers influenced by these bacteria.

Studying the link between oral bacteria and CRC also provides opportunities for prevention strategies. Promoting good oral hygiene practices and maintaining oral health may reduce the colonization and translocation of oral bacteria to the gut. Additionally, strategies that target the gut microbiota composition, such as dietary modifications or probiotic supplementation, can help create an environment that is less favorable for the growth and activity of these bacteria. As periodontitis is polymicrobial, mixed species potentially promote carcinogenesis both locally and in extraneous tissues, likely through complex microbial interactions. While mono-species effects on tumorigenesis pathways are elucidated, further research is required to understand the combined impact of oral species on carcinogenesis. Nevertheless, further research is needed to establish a clearer temporal relationship between *F. nucleatum* and CRC, which may be elucidated through prospective studies. By investigating the specific mechanisms by which *F. nucleatum* and *P. gingivalis* promote carcinogenesis, the field can not only develop sensitive and specific diagnostic tests that can aid in early detection and timely intervention but also identify potential therapies that specifically inhibit the activity of these bacteria or disrupt their interactions with host cells.

**Conclusions and future perspectives**

In conclusion, the oral-gut-circulation axis emerges as a dynamic model for host-microbial interactions, extending its influence well beyond the oral and gut systems including the circulatory system as an essential component. The oral cavity and gastrointestinal tract house the largest and most diverse microbial communities in the human
body. Recent research has revealed that these systems are not isolated but rather interact dynamically, shaping each other’s composition and function. The circulatory system, as a conduit for pathogens, bacterial products, immune cell trafficking, and inflammatory factors, adds a layer of complexity to the oral-gut axis. Strategies aimed at mitigating systemic inflammation and circulatory factors may hold therapeutic promise in this context.

Moreover, there is an emerging link between oral microbiota and cancer susceptibility. Individuals diagnosed with periodontitis exhibit a significantly elevated colon cancer risk. Thus, maintaining oral health and treating oral infections such as periodontitis may hold the key to reducing the incidence of several conditions, including colon cancer. Based on the literature that shows the tumor-promoting potential of F. nucleatum in CRC, treatment targeting the reduction of Fusobacterium populations, particularly in the oral cavity where they are notably abundant and play an essential role in biofilm formation and structure, could potentially serve as a strategy to delay or prevent tumor progression in individuals at elevated CRC risk. Additionally, P. gingivalis has emerged as another player in colon cancer pathogenesis. Together their ability to stimulate a pro-inflammatory milieu within the tumor microenvironment, modulate gut microbiota, and promote angiogenesis highlights their significance. Targeting these pathogens or their virulence factors could provide novel therapeutic avenues.

The development of non-invasive screening and diagnostic tools based on microbial signatures and inflammatory markers could significantly enhance current CRC screening practices. Traditional methods, such as colonoscopy or fecal occult blood tests, although effective, may face challenges related to patient compliance, invasiveness, and cost. Salivary and blood biomarkers, on the other hand, could provide a convenient, patient-friendly, and cost-effective alternative. Additional research is needed in prevention approaches such as personalized medicine based on microbiota profiles, probiotics for gut health restoration, and lifestyle modifications as non-invasive tools for the benefit of public health to decrease the incidence of conditions such as cancer.

Fostering multidisciplinary collaboration among experts from diverse fields will be instrumental in advancing our understanding of colon cancer and the development of immunotherapeutic interventions that target, for example, specific oral pathogens. In summary, the oral-gut-circulatory axis represents a promising frontier in colon cancer research, potentially reshaping our conception of this disease and opening innovative opportunities for prevention and treatment.

Author contributions

ST: Writing – original draft, Writing – review & editing, Conceptualization, Visualization, MA: Writing – review & editing, LM: Writing – review & editing, Supervision.

Funding

The author(s) declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References


Frontiers in Cellular and Infection Microbiology 57
Porphyromonas gingivalis evasion of autophagy and intracellular killing by Fap2 protein of fusobacterium nucleatum to human inhibitory receptor Tigit protects tumors from immune cell attack. Immunity 42, 344–355. doi: 10.1016/j.immuni.2015.01.010


Tortora et al. 10.3389/fonc.2020.610104


V. Significance

a) Colon cancer disparities in AA: exploring the genetic and microbial landscape

There is substantial evidence indicating racial disparities in both the incidence and mortality of colon cancer, with AAs experiencing higher rates compared to other racial/ethnic groups. Furthermore, AA individuals are more likely to receive diagnoses at later stages, leading to worse clinical outcomes. Compounding this issue, studies highlight a greater mistrust among AA individuals towards the medical system and biomedical research compared to CA/EA counterparts.

Despite the limited number of studies examining disparities between AA and CA colon cancer patients, there remains a notable gap in our understanding of the genetic factors influencing inflammation and immune evasion, particularly in the context of AA colon tumors. Additionally, there is a scarcity of research exploring genetic variations related to colon location and cancer stage. This study aimed to fill these knowledge gaps by investigating whether tumors from AA colon cancer patients exhibit distinctive genetic landscapes based on both colon location and tumor stage.

Mounting evidence underscores the significance of the gut microbiota in CRC, with race-associated variations observed in microbial abundance. However, few studies have specifically collected and analyzed colon tumors and adjacent non-tumor tissues from AA colon cancer patients. Consequently, our study sought to characterize the gut microbiome in this population, establishing connections between the microbial landscape and the tumor immune microenvironment.

Numerous association studies have shown an elevated risk of proximal advanced adenomas and proximal colon cancer in individuals with periodontitis. Notably, AAs exhibit a higher prevalence of both colon cancer and periodontitis. In our comprehensive study, we aimed
to understand the shared factors between these two conditions that could potentially underlie health disparities. Our focus included investigating the presence of periodontal pathogens in the colon and levels of inflammatory and bone metabolism mediators as potential contributors to the observed health inequities.

b) \textit{F. nucleatum} in the pathogenesis of colon cancer in AA patients

While previous studies have explored how \textit{F. nucleatum} promotes carcinogenesis \textit{in vitro} and \textit{in vivo} models, there are limited studies specifically conducted in human patients and no previous studies in colon cancer in health disparities. Our research aimed to address this gap by pioneering an examination into the role of oral bacteria, with a particular emphasis on \textit{F. nucleatum}, in AA patients diagnosed with colon cancer (Figure 5). Our goal was to comprehend the connection between \textit{F. nucleatum} and colon cancer, investigating its potential to co-aggregate with other oral bacteria, biofilm- and CRC-associated bacteria. Moreover, we sought to analyze the host immune system response in these patients by examining antibodies against bacterial proteins and the levels of inflammatory mediators in circulation. Additionally, our investigation extended to a comprehensive analysis of the tumor microenvironment in immune cell recruitment patterns and the genetic pathways linked to the progression of colon tumorigenesis.

AA patients exhibit the highest incidence of CRC and periodontitis, a phenomenon with multifaceted contributing factors. In periodontitis, the persistence of dysbiotic oral microbiota triggers the host’s immune response, leading to tissue destruction through the synthesis and release of cytokines, pro-inflammatory mediators, and metalloproteinases. This continuous process, coupled with oral bacterial translocation, could eventually result in complications at distant sites such as in the colon. \textit{F. nucleatum} can colonize intestinal epithelium and acts as a bridge connecting early and late colonizers, promoting CRC progression, accelerating tumorigenesis post-mutation, and inhibiting antitumor immunity.
Figure 5. The relationship of *F. nucleatum* presence in AA colon cancer patients with alterations to the gut microbiota, immune response, and tumorigenesis.

In this study, we employed DNA and RNA sequencing of tumor vs. non-tumor tissues, along with the analysis of inflammatory mediators in blood and antibody responses to bacterial proteins, to investigate the roles of inflammation, immune response, and the microbiome in colon cancer disparities. Additionally, T cell subsets within the tumor was analyzed through immunohistochemistry using the presence/absence of oral pathogens for comparison (Figure 6). To our knowledge, we are the first group to explore the interrelationship between periodontal bacteria, inflammation, and colon cancer in this demographic. Moreover, no previous study has investigated the use of antibodies against oral bacteria and serological markers associated with periodontitis to analyze the subsite tumor locations and stage of CRC in African American patients.

Exploring these interconnections not only enables researchers to devise sensitive diagnostic tests for timely intervention but also facilitates the identification of targeted therapies that can effectively inhibit the activity of these bacteria or disrupt their interactions with host cells. It is crucial to promote interdisciplinary perspectives on overall patient health by advancing our understanding of the body as a unified system and exploring the interconnectedness between
various pathologies and comorbidities. Implementing the use of validated and standardized methods, such as questionnaires, to assess oral health status in all patients, with a special emphasis on high-risk groups, can significantly contribute to comprehensive healthcare.

Figure 6. Schematic overview of the primary analyses conducted in the study, the laboratory techniques employed, the range of samples included, and the collaborative efforts involved in the research.
Specific Aims

**Aim 1:** To characterize the oral/gut microbiome and circulating antibody levels against bacteria in African American (AA) colon cancer patients.

To assess the difference in microbial communities between tumor and non-tumor colon tissues, examine the impact of oral bacteria in the composition of the gut microbial composition, and correlate these changes with antibody levels against these bacteria in serum from these patients.

**Sub Aims:**

1a) To characterize gut microbiome in tumor tissue versus matched non-tumor tissues.

1b) To explore gut microbiota changes associated with oral bacterial infection in tumors of colon cancer patients.

1c) To assess the serum antibody levels against oral and gut bacteria in colon cancer patients.

**Research design and methods – Aim 1**

The authors’ contributions to this study are as follows:

- Sofia Tortora at SUNY Downstate, Division of Gastroenterology:

  - Project design, hypothesis, and research strategy.
  - IRB oral sample-related protocols and revisions
  - Processing of samples and storage
  - Recording of clinical data
  - RNA and DNA isolation from tissue samples
  - Serum, plasma, and buffy coat isolation from blood
  - Quantitative polymerase chain reaction (qPCR) assays from DNA isolated from tissue and oral samples
  - Data analysis of qPCR, 16S rRNA and Nucleic Acid Programmable Protein Array (NAPPA) array
• Data interpretation and generation of graphs (R studio, Prism, and Excel)
• Manuscript writing and editing
• Manuscript submission and revision
• Provision of funding (2021 Downstate Health Sciences University Seed Grant)

  - Emily Vogtmann, PhD at NIH:
    • Data interpretation of 16S rRNA
    • Manuscript revision

  - Constanza Camargo, PhD at NIH:
    • Data analysis of NAPPA
    • Data interpretation of NAPPA
    • Manuscript revision

  - Ji Qiu, PhD at Arizona State University:
    • NAPPA array
    • Data interpretation of NAPPA
    • Data of colon cancer cases and healthy controls (Mayo Clinic)
    • Manuscript revision

Patient’s criteria for participation

Male and female patients with a diagnosis of colon adenocarcinoma from SUNY Downstate Medical Center and Kings County Hospital aged 45 years and older who self-identified as African American or Caribbean American consented for participation in an IRB-approved prospective
study (#312509). Samples were acquired in coordination with the Departments of Surgery and Pathology. To evaluate the immunological profile of tumors and blood of these patients without potential confounders, we excluded patients with infectious diseases such as HIV and Hepatitis B/C, patients currently treated with immunosuppressive drugs or antibiotics, and patients diagnosed with Crohn’s disease or ulcerative colitis. To study naive tumors, only patients where colon cancer was their primary cancer were included and patients treated with neoadjuvant chemotherapies were excluded. In addition, since treatment for rectal cancer differs from colon cancer (i.e. neoadjuvant therapy is given regardless of cancer stage), we excluded rectal cancers from our research design. Lastly, as we aim to study sporadic colon cancer, we excluded patients diagnosed with Lynch syndrome or patients with familial adenomatous polyposis.

**Clinical data**

We collected clinicopathological variables such as age, sex, race/ethnicity, BMI, diabetes, smoking status, alcohol intake. Post-surgery, we collected pathology reports that confirmed cancer stage, tumor volume, histological differentiation, TNM classification, MSI status, and histological results of MMR protein examinations. In addition, we follow up for information regarding treatment, cancer recurrence, and survival. In addition, we recently started to collect data for periodontal status and oral hygiene habits, before surgery, by using a standardized and validated self-report questionnaire with 8 questions used to assess oral hygiene and predict moderate/severe periodontitis (69).

**Sample collection**

After the patient had consented to participate in the study, approximately 10 ml blood was collected on the day of the surgery. Tumor and adjacent non-tumor colon tissues were collected from fresh specimens. Colon surgical specimens were taken from the operating room and
transported to the pathology department for assessment of the specimen to provide the research samples. Blood was spun to separate serum. Serum was stored at -80°C until use. Oral samples were collected by rubbing a soft swab tip against gums and teeth surface for 30 seconds. After collection, tip was placed into the tube prefilled with reagent of DNA/RNA Shield SafeCollect collection kit (Zymo).

**Processing of tumor and non-tumor tissues:** When collecting colon tissues, specific precautions must be taken to prevent contamination and ensure the accuracy of results. Alongside standard sterilization procedures for equipment and the use of personal protective equipment (PPE) during surgery, it's essential to implement strict aseptic techniques throughout the collection process. Once collected, to preserve the DNA and RNA from host and microbial communities, samples should be frozen at ultra-low temperatures at -80°C and preserved using stabilization solutions designed for nucleic acids (DNA/RNA shield, Zymo). In addition to tissue samples, appropriate negative controls were included in the sequencing workflow to monitor for potential contamination introduced during processing or sequencing. These controls undergo the same steps as tissue samples but without tissue present, aiding in the identification and mitigation of contamination sources. During data analysis, rigorous quality control measures should be implemented to identify and remove sequences originating from contaminants rather than tissue samples. This may involve filtering sequences based on abundance, quality scores, or known contaminant sequences.

**RNA and DNA isolation:** Tumor and normal (non-tumor) tissue samples were used to obtain RNA, DNA, and protein samples. We used the AllPrep DNA/RNA/Protein Mini Kit from Qiagen and followed the manufacturer's protocol to obtain host DNA. DNeasy Power Soil Pro Kit (Qiagen) was used to extract bacterial genomic DNA. For monitoring bacterial DNA after extraction, we ran qPCR by amplification of DNA from universal 16S rRNA locus (Table 1).
16S rRNA Gene Amplification: Considering the presumed high host DNA content, all samples were treated as having low bacterial DNA content. Consequently, we employed the Zymo Quick-16S/ V3V4 Plus NGS Library Prep Kit, a qPCR-based library preparation technique, to optimize bacterial 16S rRNA gene amplification and identify the most suitable region for this study. PCR amplifications for the target region were performed using the CFX Opus 96 Real-Time PCR System (Bio-rad). After amplification, the resulting 16S libraries were pooled by adding an equal volume to each, and the final library concentration was quantified using the Qubit 2.0 Fluorometer (Invitrogen). Following equimolar pooling based on these quantifications, sequencing was conducted on the Illumina MiSeq platform. A loading concentration of 8pM, supplemented with 20% PhiX, was used along with the paired-end 300 cycle MiSeq Reagent Kit V3 (Illumina).

16S rRNA Analysis: 16S rRNA sequences from targeted (V3V4) were processed and applied using the DADA2 (DADA2 1.12.1) pipeline (70) in conjunction with R v4.1.0. DADA2 was employed for quality filtering, trimming, error correction, exact sequence inference, chimera removal, and generation of the amplicon sequence variant table (ASV) with a minimum count cutoff set at 2000. Based on the quality score profiles of the sequencing reads, forward reads were truncated at 240 bp, and reverse reads were truncated at 240 bp before merging. Ambiguities in the overlap region were not permitted, and default parameters were utilized in the R dada2 package's filterAndTrim() function \( \text{truncLen} = c(240,240); \text{trimLeft} = c(5,5), \text{maxN} = 0, \text{maxEE} = c(2,2), \text{truncQ} = 2 \). After the dereplication and merging of reads, chimeric reads were identified by consensus across samples using the DADA2 function removeBimeraDenovo(). All samples passed the set threshold of 2000 reads post-quality filtering for inclusion in the analysis.

The MAFFT and FastTree modules in QIIME2 (71) facilitated the generation of a phylogenetic tree of all ASV sequences. Taxonomic classification was undertaken using a native naïve RDP Bayesian classifier aligned against the Silva version 138 database (72).
**16S rRNA sequencing analysis:** This analysis was performed in collaboration with the Microbiome Core at Columbia University. Robotic extraction was done manually with pre-treatment step using QIAamp® 96 DNA QIAcube® HT (catalog #51331), DNA extraction for positive control using ZymoBIOMICS Microbial Community Standard (catalog #D6300) and extraction kit: QIAamp® 96 Virus QIAcube® HT (catalog # 57731). Sequencing was done using Zymo Quick-16S/ V3V4 Plus NGS Library prep kit (catalog # D6420), which is a qPCR-based library preparation technique. Sequencing was performed in collaboration with Microbiome Core at Columbia University for library generation, and analysis of alpha-diversity and beta-diversity. The analysis of compositional change and relative abundance (at genus and phylum level) was done by ST.

**Table 1.** Primer and probe sequences for each oral bacterial, pan bacteria 16S rRNA gene, and housekeeping gene.

<table>
<thead>
<tr>
<th>Species/gene</th>
<th>5' label</th>
<th>Sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal bacterial 16S</td>
<td>SYBR Green</td>
<td>Forward primer: ACTCCTACGGGAGGCAGCAGT</td>
</tr>
<tr>
<td>rRNA</td>
<td></td>
<td>Reverse primer: ATTACCGCGGCTGCTGGC</td>
</tr>
<tr>
<td>PGT(25,73)</td>
<td>FAM</td>
<td>Forward primer: ATCCCCAAAGCACCCTGGTTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse primer: AGAGGCCCAAGATAGTCTGCTGTAA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FAM probe: CCATCCATGTCTCCTC</td>
</tr>
<tr>
<td><em>Fusobacteria nucleatum</em>(73)</td>
<td>FAM</td>
<td>Forward primer: CAACCATTACTTTAATCTACCATGTCCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse primer: GTTGACTTTACAGAAGGAGATTATGTAAAAAATC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FAM probe: TCAGCAACTTTGCTCTCTGCTTTAATGAACC</td>
</tr>
<tr>
<td><em>Porphyromonas gingivalis</em>(74)</td>
<td>FAM</td>
<td>Forward primer: TGGTTTTCATGCGAGCTTCTTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse primer: TCGGCACCTTCGTAATCTCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FAM probe: CGTACCTCATATCCCGGAGGGGCTG</td>
</tr>
<tr>
<td><em>Prevotella intermedia</em>(75)</td>
<td>FAM</td>
<td>Forward primer: CCACATATGCGCATCTGAGCTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse primer: CTGATCGTAGCCTTGGTGGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FAM probe: ACCAAGAGATTTCATCGGTGGAGGATGGG</td>
</tr>
</tbody>
</table>
**qPCR:** We performed qPCR of DNA extracted from colon tissues (surgical resections) of colon cancer patients to measure the relative abundance of oral bacteria by custom TaqMan Gene Expression Assays (FAM) designed to amplify *F. nucleatum*, *P. gingivalis*, and *P. intermedia* DNA. DNA was quantified using a Nanodrop spectrophotometer. Each reaction contained 5ng of DNA and was assayed in triplicate in 10 mL reactions containing 1X final concentration TaqMan Universal PCR Master Mix (Applied Biosystems), 18 mM each primer, and 5 mM probe in a 384-well optical PCR plate, protocol adapted from a paper by Castellarin et al (73,76,77). Cycling conditions were as follows: 2 min at 50°C, 10 min at 95°C, and 40 cycles of 15 sec at 95°C and 1 min at 60°C. Cycle thresholding was calculated using the automated settings for SDS 2.2 (Applied Biosystems). The Ct values were normalized to the amount of human gDNA in each reaction by using a primer/probe set for the reference gene, prostaglandin transporter (PGT). A specific gene of *F. nucleatum* and the reference human gene SLCO2A1 were amplified using custom-made TaqMan primer/probe sets (Applied Biosystems) as previously reported by Castellarin et al (77) and corrected *F. nucleatum* sequence in Registration and Replication study by Repass(73,76) (Gene Accession NC_003454). *P. gingivalis* primers and probe for the target gene were found in the paper by Hyvärinen et al (74) and *P. intermedia* primers and probe were found in the paper by Nonnenmacher et al (75) (Gene Accession L16468) (Table 1).

Genomic DNA from *F. nucleatum subsp. nucleatum* strain VPI 4355 (ATCC # 25586D-5), *P. gingivalis* strain 2561 (ATCC # 33277D-5), and *P. intermedia* strain VPI 4197 (ATCC # 25611D-5) were used to prepare the standard curve and as a positive control; the negative control was of sterile H₂O. We prepared a series of dilutions of each DNA sample that covered a range of concentrations, created 5 to 7 dilutions, spanning at least a 10-fold difference in concentration (initial concentration was 10 ng/μl). A standard curve using the cycle threshold (Ct) values of the standard dilutions was plotted on the x-axis and the logarithm of the initial DNA concentration of the standard dilutions on the y-axis. The Ct value represents the cycle number at which the
fluorescence signal crosses the threshold level. Then we determined the efficiency of the amplification reaction using the slope of the standard curve. The formula for efficiency (E) is given by: 

\[ E = (10^\frac{-1}{\text{slope}}) - 1 \times 100 \]

The slope is the negative of the standard curve’s slope. R values for *F. nucleatum*, *P. gingivalis*, and *P. intermedia* were 0.995, 0.981, and 0.991, respectively (Supplementary Figure 1). By measuring the efficiency of amplification reactions using the standard curve method, we ensured the accuracy and reliability of the qPCR results when working with DNA samples. Then, we calculated ΔCt values for each sample. 

\[ \Delta C_t = C_t(\text{Target gene}) - C_t(\text{Reference gene}) \]

ΔCt is the difference in threshold cycle number for each target gene and reference assay (PGT). The fold difference \(2^{-\Delta \Delta C_t}\) in Fn, Pg and Pi abundance in tumor versus non-tumor tissues was calculated by subtracting ΔCttumor from ΔCtnon-tumor. Not-detected values were set to 40. Technical triplicates were averaged for each sample.

**NAPPA array:** The assay was performed in collaboration with Arizona State University for the analysis of the IgG and IgA antibody response against 297 bacterial proteins from 10 bacteria associated with colon cancer. These are oral bacteria *P. gingivalis*, *Fusobacterium varium*, and *F. nucleatum*, gastric bacteria *Helicobacter pylori*, gut bacteria *Streptococcus gallolyticus*, *Bacteroides fragilis*, and *Enterococcus faecalis*, and probiotic bacteria *Bifidobacterium adolescentis*, and *Lactobacillus plantarum*. Utilizing Microsoft Excel, the median intensity values of IgG and IgA serology profiling for an entire slide were inputted. The median intensity value (IMedian) for all proteins was calculated. Each protein’s raw intensity was then divided by IMedian, resulting in a new set of median-normalized intensity (MNI) values. A seropositive cutoff of MNI ≥ 2.0 has been used in prior studies (78–80). Seropositive response proteins will be selected as possible candidate biomarkers (81).

**Data analysis:** we construct an amplicon sequence variant (ASV) table through the DADA2 pipeline, a bioinformatics tool that accurately identifies and delineates sequence variants
providing a more precise representation of microbial diversity. The taxonomic assignment of these ASVs was performed against the Silva database (99% identity), a comprehensive reference database for rRNA gene sequences, enabling accurate classification of microbial taxa based on their genetic signatures. To assess the microbial differences between groups, we analyzed α diversity using Observed ASVs and Shannon (richness and evenness) indexes, then the Wilcoxon rank-sum test to determine α diversity differences by tumor status, sex, colon location, and cancer stage. Permutational multivariate analysis of variance for β diversity to examine whether the overall bacterial community composition differed by groups and principal coordinate analysis (PCoA) to observe clustering between the different groups performed based on weighted and unweighted UniFrac distances. Differential abundance testing was done by DESeq2 by volcano plots and ANCOM-BC2 test plots by R tool. To assess the differences between relative abundance between tumor and non-tumor tissue samples, we employed the nonparametric Wilcoxon matched-pairs signed-rank test. To investigate the co-abundance relationship between \textit{F. nucleatum} and \textit{P. gingivalis/P. intermedia}, we utilized one-way ANOVA (Analysis of Variance). To account for multiple comparisons, we employed the Tukey correction. We categorized cancer stage into two groups: early (I-II) and late (III). Mann-Whitney U test for analysis for right/left location and early/late stage. Statistical analyses were conducted using GraphPad Prism 10 (GraphPad Inc., San Diego, CA, USA), R and Excel. P values of < 0.05 was considered statistically significant.

**Results – Aim 1**

Sub-aim 1a) **To characterize gut microbiome in tumor tissue versus matched non-tumor tissues.**

The study includes 47 patients, all of whom are of African American descent. The mean age of the patients is 66 years, with an age range spanning from 48 to 67 years. In terms of sex distribution, 22 patients were male (45.5%), and 25 patients were female (54.5%). Most patients were in the early stages of cancer (I-II), accounting for 59.6% of the cohort, while 40.4% were in
the late stage (III). Regarding the tumor location, 48.9% of the cases were on the right side (proximal), and 51.1% were on the left side (distal). Only tumors from 30 patients were analyzed for microsatellite instability (MSI), from this cohort, 26.7% of tumors exhibited MSI-H (microsatellite instability-high). The average BMI of the patients was $26.78 \pm 4.11$, with 36.1% categorized as overweight and 19.2% as obese. None of the patients had Type 1 diabetes, while 29.8% had Type 2 diabetes. In terms of lifestyle factors, 22.7% of patients had a history of smoking, while the majority (77.3%) have never smoked. Similarly, 31.8% of patients have a history of alcohol consumption, while the majority (68.2%) have never consumed alcohol (Table 2).

**Table 2. Patients’ characteristics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients, n</td>
<td>47</td>
<td>-</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>66</td>
<td>-</td>
</tr>
<tr>
<td>Range</td>
<td>48-67</td>
<td>-</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>47</td>
<td>100</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>22</td>
<td>45.5</td>
</tr>
<tr>
<td>Female</td>
<td>25</td>
<td>54.5</td>
</tr>
<tr>
<td><strong>Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early (I-II)</td>
<td>28</td>
<td>59.6</td>
</tr>
<tr>
<td>Late (III)</td>
<td>19</td>
<td>40.4</td>
</tr>
<tr>
<td><strong>Location</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right (Proximal)</td>
<td>23</td>
<td>48.9</td>
</tr>
<tr>
<td>Left (Distal)</td>
<td>24</td>
<td>51.1</td>
</tr>
<tr>
<td><strong>MSI-H tumor</strong></td>
<td>8/30</td>
<td>26.7</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>26.78</td>
<td>-</td>
</tr>
<tr>
<td>Range</td>
<td>17-35</td>
<td>-</td>
</tr>
<tr>
<td><strong>Overweight</strong></td>
<td>17</td>
<td>36.1</td>
</tr>
<tr>
<td><strong>Obese</strong></td>
<td>9</td>
<td>19.2</td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
* Information from colon cancer patients that self-identified as African American includes African descendant, African American, and Caribbean American. Tumors staged I and II were combined in the Early Stages group and tumors staged III were considered Late stage.

We extracted DNA from 66 matched tumor and non-tumor samples from 33 AA colon cancer patients. The rarefaction curve for taxon accumulation for each sample was plotted, with the number of taxa as a function of the number of samples among all samples (Figure 7). Rarefaction curves generally grow rapidly at first, as the most common taxa are found, but at the plateau of the curve, only the rarest species remain to be sampled. 2000-read cut-off point was selected after visualization of the Shannon index rarefaction curve. Three out of 66 samples did not reach the plateau after about 2000 reads and therefore are not suitable for further analyses. Thus, 63 samples are used for the downstream analysis.

<table>
<thead>
<tr>
<th></th>
<th>Type 2</th>
<th>Ever</th>
<th>14</th>
<th>29.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td></td>
<td>10</td>
<td></td>
<td>22.7</td>
</tr>
<tr>
<td>Never</td>
<td></td>
<td>34</td>
<td></td>
<td>77.3</td>
</tr>
<tr>
<td>Alcohol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td></td>
<td>14</td>
<td></td>
<td>31.8</td>
</tr>
<tr>
<td>Never</td>
<td></td>
<td>30</td>
<td></td>
<td>68.2</td>
</tr>
</tbody>
</table>

**Figure 7.** Rarefaction curve for Taxon accumulation for each sample using R software.
We investigated the proportions of different taxa within each group to highlight the percentage composition of the microbiome at various taxonomic levels. This analysis visually represents the distribution of organisms at the phylum and genus levels. Figure 8 illustrates the relative abundance of the most prevalent organisms found at these taxonomic levels specifically in tumor tissues in comparison to non-tumor tissues. In the legend, the category 'other' encompasses unclassified and less abundant taxa (below a cutoff of < 0.005). Notably, the microbial communities within tumor tissues were predominantly comprised of Firmicutes (67.95%), Bacteroidota (15.38%), Proteobacteria (12.82%), and Actinobacteria (2.56%). Additionally, Fusobacteriota (1.28%) was exclusively present in the tumor group and notably absent in the non-tumor group (Figure 8).

**Figure 8.** The relative abundance of the most frequent organisms at the phylum and genus levels in tumors when compared to non-tumor tissues. In the legends “other” is representative of unclassified and lower-abundance taxa (cut off < 0.005).
We analyzed alpha diversity to explore significant differences among distinct groups (Figure 9). Alpha diversity serves as a metric representing the average diversity of taxa within a sample. Using two diversity indexes, the number of observed species (ASV) and the Shannon index, we assessed richness (the number of different taxa) and evenness (the relative abundance of each taxon) within microbial communities. For example, a higher Shannon index indicates greater diversity. Our findings suggest that the microbial community in tumor tissue may exhibit less even distribution of abundances compared to non-tumor tissue across all groups, except for late-stage tumors. Specifically, we observed that tumor tissues displayed significantly lower microbial diversity than non-tumor tissues (Shannon index, \( p=0.039 \)). Interestingly, male patients exhibited approximately a 1.5-fold increase in the number of observed species compared to female patients (\( p=0.048 \)). However, when incorporating the Shannon index analysis, no difference in richness was observed between the sexes. Moreover, our analysis revealed that right-sided tumors exhibited reduced microbial diversity (\( p=0.0093 \)) and approximately half the number of observed species compared to left-sided tumors (\( p=0.0011 \)) (Figure 9).
We utilized permutational multivariate analysis of variance (PERMANOVA) to explore β diversity, aiming to understand potential variations in the overall bacterial community composition across different factors such as tumor status, sex, tumor location, and cancer stage. Using Principal Coordinate Analysis (PCoA) based on weighted and unweighted UniFrac distances, we visualized clustering patterns among diverse groups. Despite our comprehensive analysis using PCoA plots, we found no significant differences in beta-diversity between sex, left/right colon locations, and early/late cancer stages among colon tumor samples (p > 0.005).

Although we didn't detect significant differences in overall bacterial communities across the groups we examined (tumor vs. non-tumor, sex, tumor location, and cancer stage), there were
particular bacterial species or strains that did show noticeable differences in their abundance or presence among these groups (Figure 10).

**Figure 10.** Differential abundance testing (DESeq2; an R package) and ANCOM-BC2 testing (analysis of compositions of microbiomes with bias correction; an R package) were carried out to determine taxa differences between groups. Here, we identified a small number of bacterial ASVs that differed between the baseline and the other group(s). Significantly differential abundant bacteria were determined by $p < 0.05$ from both tests. In the following graphs, bacterial taxa in colors are significantly different at $p < 0.05$ and absolute fold change $> 0$. Only DESeq2 volcano plots were provided for the following four comparisons since they had no microorganism with significant ANCOMBC result.
Interestingly, all the ASV identified with significant differential abundance are oral-origin bacteria. *Treponema socranskii* was identified with significant differential abundance; 1) Tumor samples when compared to non-tumor ones, 2) Left colon when compared to right ones 3) Early colon cancer samples when compared to late cancer stage. In addition, among colon tumor samples, *Prevotella_7 NA* was significantly increased in male patients, in left-sided location when compared to right, and in early cancer stage compared to late. Moreover, among colon tumor samples, *Fusobacterium necrophorum* were identified in right colon samples compared to left side, as well as in late cancer stage compared to early stage. *F. nucleatum* was significantly present in early stage and right-sided colon tumors. *Veillonella spp* is a normal bacterium in the intestines and oral mucosa of mammals which was identified in non-tumor tissues (Figure 10). This suggests that while the broad bacterial community composition might be similar, there are certain bacterial species or strains that might play a role in distinguishing between these groups.

**Sub-aim 1b)** To explore gut microbiota changes associated with oral bacterial infection in tumors of colon cancer patients.

*F. nucleatum*, a key bacterium in periodontitis, has been consistently associated with an increased risk of colon cancer. Studies report increased levels of *F. nucleatum* in colon tumor tissues, observed across diverse racial/ethnic groups. This correlation is especially noteworthy among AA patients, known to have a higher prevalence of periodontitis, further accentuating the potential role of periodontal pathogens in colon cancer. Considering the established link between periodontitis and colon cancer risk, we hypothesize the presence of other periodontal pathogens, such as *P. gingivalis* and *P. intermedia*, in colon tumor tissues. These bacteria are integral components of the bacterial complex associated with periodontitis. This hypothesis aligns with the understanding that periodontal pathogens might contribute to systemic diseases like colon cancer, potentially through translocation.
Through 16S rRNA sequencing analysis, we initially identified oral pathogens at a general genus and higher hierarchy level. To deepen our understanding and target specific species, we conducted qPCR using genomic DNA extracted from colon tumors and matched non-tumor tissues obtained from 47 AA colon cancer patients. This method enabled us to precisely pinpoint and quantify the presence of oral pathogens, including *F. nucleatum*, *P. gingivalis*, and *P. intermedia*, at the species level. To calculate the fold difference (2-ΔΔCt) in bacterial abundance between tumor and non-tumor tissues, we computed the difference in threshold cycle numbers (ΔCt) for each sample, contrasting the test bacteria against the reference assay prostaglandin transporter (PGT)(25). The results obtained through qPCR were in alignment with our earlier findings from the 16S rRNA sequencing data enrichment analysis.

The Ct values is a measure of the concentration of the target sequence in the qPCR reaction (76). In our observations, the raw Ct values for PGT fell well within the accepted normal range (<30). Conversely, the raw Ct values for *F. nucleatum* were consistently high across all samples in both independent qPCR runs, a pattern consistent with findings from previous studies (73,76,77). For non-tumor samples, the Ct values for the *F. nucleatum* sequence ranged between 25.10 and 40, with a mean value of 36.45, while tumor samples exhibited Ct values ranging from 31.55 to 40, with a mean value of 37.40 (Figure 11A). This data illustrates that the overall abundance of this bacterium was notably higher in tumor tissues compared to adjacent non-tumor tissues, exhibiting a 62-fold increase (p=0.0003). The fold increase of *F. nucleatum* in tumor tissues is calculated by dividing its abundance in tumor tissues by its abundance in adjacent non-tumor tissues (Figure 11B). Furthermore, our analysis revealed associations between the relative abundance of *F. nucleatum* and specific characteristics of the cancer. In particular, its presence was significantly linked to the right-sided location of tumors (p= 0.0253), more advanced disease stages within tumors (p=0.0323), and notably higher levels in late-stage III tumor tissues compared to their paired non-tumor counterparts (p= 0.0258) (Figure 11C).
Figure 11. The relative abundance of *Fusobacterium nucleatum* DNA was determined using quantitative real-time PCR (qPCR). A. The cycle threshold (Ct) values for *F. nucleatum* gene in non-tumor and tumor samples. B. Mean fold expression change in tumor vs non-tumor colon tissue for all patients. C. *F. nucleatum* levels in tumor and adjacent non-tumor tissue and by cancer location and stage. Wilcoxon matched-pairs signed-rank test to assess the differences in relative abundance between tumor and non-tumor tissues and ANOVA test for comparison between location and cancer stage groups. Tumors n=47, non-tumors n=47; right n=23, left n=24; early n=28, late n=19; ns= not significant, * p < 0.05, ** p < 0.005, *** p < 0.001.

Our analysis extended to *P. gingivalis*, where we observed distinct Ct values among both tumor and non-tumor samples. For non-tumor tissues, the Ct values for the *P. gingivalis* sequence ranged between 29.60 and 40, with a mean value of 38.24. In comparison, tumor samples exhibited Ct values ranging from 30.94 to 40, with a mean value of 38.97 (as depicted in Figure 12A). Surprisingly, although the overall abundance of *P. gingivalis* showed a substantial 1700-fold increase in tumor tissues compared to adjacent non-tumor tissues, statistical analysis did not reveal any significant differences between tumor and non-tumor tissue (p=0.4586) (Figure
Additionally, there were no significant differences detected concerning the location of the tumor or the cancer stage (p= 0.8548, >0.999, respectively) (Figure 12C). Furthermore, when assessing the infection rates by *P. gingivalis*, our findings showed a 19.15% infection rate in tumor tissues (n=9) and a slightly higher rate of 25.53% in non-tumor tissues (n=12). Despite the substantial increase in overall abundance, these infection rates did not demonstrate significant differences between tumor and non-tumor tissues.

**Figure 12.** The relative abundance of *Porphyromonas gingivalis* DNA was determined using quantitative real-time PCR (qPCR). A. The Ct values for *P. gingivalis* gene in non-tumor and tumor colon tissues. B. Mean fold expression change for *P. gingivalis* in tumor vs non-tumor colon tissue for all patients. C. *P. gingivalis* levels in tumor and adjacent non-tumor tissue, cancer location, and stage. Wilcoxon matched-pairs signed-rank test to assess the differences between relative abundance between tumor and non-tumor tissues and ANOVA test for comparison between location and cancer stage groups. Tumors n=47, non-tumors n=47; right n=23, left n=24; early n=28, late n=19; ns= not significant.
Similar analysis was done for *Prevotella intermedia*. The Ct values for *P. intermedia* sequence in the non-tumor samples had a range of 26.91 to 40 with a mean value of 35.66, and the Ct range for the tumor samples was between 20.38 and 40 with a mean value of 34.91 (Figure 13A). The overall abundance of *P. intermedia* was 120 times greater in tumor versus adjacent non-tumor tissues (Figure 13B). No significant differences were found between tumor vs non-tumor tissue (p=0.2443) and cancer stage (p=0.9037). *P. intermedia* was associated with higher amounts bacteria in left-sided colon cancer in tumor tissue when compared to matched non-tumor tissues (p=0.0246) (Figure 13C). Rate of infection by *P. intermedia* was 55.32% in tumor (n=26) and 51.06% in non-tumor tissues (n=24).

**Figure 13.** The relative abundance of *Prevotella intermedia* DNA was determined using quantitative real-time PCR (qPCR). A. The Ct values for the non-tumor and tumor colon tissues. B. Mean fold expression change for *P. intermedia* in tumor vs non-tumor colon tissue for all patients. C. *P. intermedia* levels in tumor and adjacent non-tumor tissue, cancer location, and
stage. Wilcoxon matched-pairs signed-rank test to assess the differences between relative abundance between tumor and non-tumor tissues and ANOVA test for comparison between location and cancer stage groups. Tumors n=47, non-tumors n=47; right n=23, left n=24; early n=28, late n=19; ns= not significant, * p < 0.05.

Our findings showed that in tumor tissues, both *P. intermedia* and *F. nucleatum* were more abundant than in non-tumor tissues. Specifically, we detected *P. intermedia* in tumors of 25 patients (55.0%) and *F. nucleatum* in 21 patients (45.0%). As periodontal inflammation progresses, there is a notable increase in the quantities of *P. gingivalis, P. intermedia, and F. nucleatum* within the subgingival biofilm (82). Based on the study by Lo et al (83), we investigated whether the presence and amount of *F. nucleatum* in colon tissue corresponded to the concurrent presence of *P. intermedia* and *P. gingivalis* (Figure 15). To discern associations between the presence of *P. intermedia/P. gingivalis* and *F. nucleatum* in tumor tissues, Fisher’s exact test was employed. Interestingly, approximately half of the cases positive for *F. nucleatum* also exhibited positivity for *P. intermedia* in both tumor and non-tumor tissues (p=0.5279 and 0.5439, respectively). Whereas, roughly a third of *F. nucleatum*-positive cases displayed positivity for *P. gingivalis* in both tumor and non-tumor tissues (p= 0.3586 and 0.4496, respectively). In addition, we classified *F. nucleatum* infection as high or low based on the median cutoff value, to analyze if the elevated abundance of this bacterium could have a relationship in its coexistence with these two bacteria. Notably, there was no significant association found between the co-abundance of *P. gingivalis* and *F. nucleatum*. Intriguingly, the relative abundance of *P. intermedia* notably increased in Fn-high cases compared to Fn-low and Fn-negative cases in both tumor and non-tumor tissues (Figure 15).
As previously discussed, *F. nucleatum* has a crucial role within the biofilm structure, acting as a pivotal 'bridge' bacterium by adhering to a diverse range of co-aggregated species through surface adhesins (84). Using information collected from both qPCR and 16S rRNA sequencing, we aimed to uncover differences in the microbial communities associated with the presence or
absence of *F. nucleatum*. By comparing Fn-positive (n=16) and Fn-negative (n=15) tumors samples, our goal was to identify unique microbial signatures or alterations in the overall composition linked to the presence of *F. nucleatum* within these groups. In Figure 16A, we showed the dominant genus-level bacteria identified in our analysis. Notably, Fn-positive tumors exhibited a remarkable surge in several genus-level bacteria compared to Fn-negative tumors. Specifically, Fn-positive tumors displayed significant fold increases in several bacteria: There were significant increases in the relative abundance of several bacteria in Fn-positive tumors including: *Leptotrichia* (168-fold, p=0.201), *Campylobacter* (78-fold, p=0.024), *Pseudomonas* (51.5-fold, p=0.009), *Treponema* (50-fold, p=0.0081), *Fusobacterium* (10-fold, p>0.001), *Porphyromonas* (10-fold, p=0.003), *Diallister* (8.5-fold, p=0.003), *Prevotella* (4-fold, p=0.001). Conversely, Fn-negative tumors showed notable rises in *Haemophilus* (99-fold, p=0.016), *Escherichia-Shigella* (5-fold, p=0.404), *Lactobacillus* (4.8-fold, p=0.917), *Streptococcus* (3.7-fold, p=0.342), and *Gemella* (1.8-fold, p>0.001). These findings underscore distinctive bacterial shifts between Fn-positive and Fn-negative tumors, emphasizing the potential influence of *F. nucleatum* in changing the microbial community within colon tumors. The presence of *F. nucleatum* might be linked to heightened levels of oral pathogens and decreased beneficial bacteria, contributing to alterations in the gut microbiota within colon cancer.

To facilitate clear visualizations highlighting the differences between Fn-positive and Fn-negative groups, we grouped bacteria into five distinct clusters for comparative analysis between the two groups, allowing for a more comprehensive examination of microbial shifts (Figure 16B). Cluster 1 contained ASVs linked to genus bacteria associated with periodontitis, labeled as 'Oral pathogens'. Cluster 2 comprised bacteria commonly found in normal oral flora. Cluster 3, termed 'Probiotics', encompassed ubiquitous 'good bacteria' of the gut microbiota. Cluster 4 showed strong taxonomic connections to biofilm, as noted in the studies by Drewes et al (20) and Warren et al (23). Lastly, cluster 5, 'CRC pathogens', exhibited diverse taxonomy with high association
and prevalence in individuals with CRC (18). We observed that ‘Oral pathogens’ and ‘Biofilm’ were approximately 3-fold and 1.3-fold higher, respectively, in the Fn-positive group compared to the Fn-negative group. This trend aligns with the aggregative role suggested for *F. nucleatum* in the colon, reminiscent of its behavior in the oral epithelium. Conversely, Fn-positive cases displayed reduced abundance of ‘Probiotics’. A similar pattern was evident in ‘CRC pathogens’. This implies that *F. nucleatum* infection might contribute to increased oral pathogen abundance and reduced levels of beneficial bacteria, potentially exacerbating gut microbiota alterations in colon cancer.

**Figure 16.** Analysis of microbial relative abundance in genus-level between *F. nucleatum* presence (Fn-positive, n=16) and absence (Fn-negative, n=15) in tumors. A. Relative abundance comparison between Fn-positive and Fn-negative groups for most prevalent genus bacteria. B.
Relative abundance comparison between Fn-positive and Fn-negative for cluster groups of bacteria with similar characteristics (lower area of the panel).

**Figure 17** shows further analysis to determine the differences in the “Oral pathogens” cluster and specific taxa of oral pathogens. **Figure 17A** shows the total relative abundance of cluster “Oral pathogens” between Fn-positive and Fn-negative. Fn-positive has increased levels of relative abundance in “Oral pathogens” as a whole (p=0.0006) and without *Fusobacterium* species (p= 0.0012). All oral pathogens were significantly increased in association with the presence of *F. nucleatum* in tumors (*Fusobacterium*, p= 0.0002; *Porphyromonas*, p= 0.0171; *Prevotella*, p= 0.0106; *Treponema*, p= 0.0081) (**Figure 17B**).

**Figure 17.** Comparative analysis to determine the differences in the “Oral pathogens” cluster and specific taxa of oral pathogens. **A.** The relative abundance of cluster “Oral pathogens” between Fn-positive and Fn-negative groups. **B.** The relative abundance of specific oral pathogens between Fn-positive and Fn-negative groups. Mann Whitney test, Fn-positive n=16, Fn-negative n=15; *ns* = not significant, * p < 0.05, ** p < 0.005, *** p < 0.001.
Similarly, Figure 18 shows the relative abundance of bacteria for other clusters between Fn-positive and Fn-negative. Fn-positive group is associated with increased levels of relative abundance for taxa in “Oral bacteria” cluster, Dialister and Gemella (p=0.002 and p<0.001, respectively). Campylobacter (p=0.263), Haemophilus (p=0.016), and Pseudomonas (p=0.009) high levels were also associated with F. nucleatum infection.

**Figure 18.** The relative abundance of specific taxa between Fn-positive and Fn-negative groups. Mann Whitney test, Fn-positive n=16, Fn-negative n=15; ns= not significant, * p < 0.05, ** p < 0.005, *** p < 0.001.

**Sub-aim 1c** To assess the serum antibody levels against oral and gut bacteria in patients diagnosed with colon cancer.

We evaluated humoral responses to oral and gut immunoproteome using Nucleic Acid Programmable Protein Array (NAPPA) in collaboration with Arizona State University. This is a novel protein microarray technology that replaces the process of spotting purified proteins with the simple process of spotting plasmid DNA. Proteins corresponding to the complementary DNA
(cDNA) are produced in situ as needed at the time of the assay by human in vitro transcription and translation (IVTT)-coupled cell lysates. The cDNAs are configured to attach a common epitope tag to all the proteins on their C-termini so that they can be captured by a high-affinity capture reagent. This method minimizes direct manipulation of the proteins and produces them at the time of the experiment, avoiding problems with protein purification and stability (Figure 19). NAPPA has been successfully used to evaluate humoral responses against autoantigens and microbial antigens in various cancer types (79,80).

Figure 19. NAPPA processing steps are:
1. Spot cDNA plasmid and a capturing reagent in an array format onto a glass slide.
2. Express the protein, with a fusion tag, using IVTT reagent
3. In situ capture the expressed proteins through the binding of the fusion tag and the fusion tag antibody.

We analyzed 44 human serum samples using a bacterial array containing 297 genes coding for bacterial proteins sourced from 10 distinct bacteria linked with colon cancer (Table 3). This selection was curated following an extensive literature review aimed at identifying bacteria associated with CRC (Tortora et al., World J Gastrointest Oncol, 2022)(18). The array included oral bacteria \textit{P. gingivalis}, \textit{Fusobacterium varium}, and \textit{F. nucleatum}, gastric bacteria \textit{Helicobacter pylori}, gut bacteria \textit{Streptococcus galloyticus}, \textit{Bacteroides fragilis}, and \textit{Enterococcus faecalis}, as well as probiotic bacteria including \textit{Bifidobacterium adolescentis} and \textit{Lactobacillus plantarum}. 

<table>
<thead>
<tr>
<th>HD-NAPPA works...</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1:</strong> Production of HD-NAPPA by printing plasmid DNA and capture antibody into Silicon NanoWells (SiNWs).</td>
</tr>
<tr>
<td><strong>Step 2:</strong> Protein display on HD-NAPPA after isolated expression and in situ capture in individual wells.</td>
</tr>
<tr>
<td><strong>Step 3:</strong> Antibody profiling to detect distinct sero-reactivity against individual antigens.</td>
</tr>
</tbody>
</table>

| Capture-antibody | Plasmid | Proteins | Serum-antibody | Secondary antibody |

We analyzed 44 human serum samples using a bacterial array containing 297 genes coding for bacterial proteins sourced from 10 distinct bacteria linked with colon cancer (Table 3). This selection was curated following an extensive literature review aimed at identifying bacteria associated with CRC (Tortora et al., World J Gastrointest Oncol, 2022)(18). The array included oral bacteria \textit{P. gingivalis}, \textit{Fusobacterium varium}, and \textit{F. nucleatum}, gastric bacteria \textit{Helicobacter pylori}, gut bacteria \textit{Streptococcus galloyticus}, \textit{Bacteroides fragilis}, and \textit{Enterococcus faecalis}, as well as probiotic bacteria including \textit{Bifidobacterium adolescentis} and \textit{Lactobacillus plantarum}.
The arrays were produced following established protocols (85). To maintain consistency, a commercial normal serum pool underwent routine screening using the assay, ensuring daily quality control during runs. To evaluate array consistency, we examined duplicates of proteins displaying an anti-GST response. This assessment revealed a high correlation coefficient of 0.92 for GST, confirming the robust reproducibility of the bacteria-NAPPA arrays. Detecting protein expression involved the use of anti-GST monoclonal antibodies (Cell Signaling) diluted at a ratio of 1:200. Subsequently, slides underwent scanning with a Tecan PowerScanner, and Array-Pro software was employed for extracting raw fluorescence intensity data. The resulting median normalized intensity (MNI) data comprehensively depicted the intensity levels across all 297 proteins tested against the 44 samples. To interpret this data, an MNI threshold of ≥2.0 was established to denote seropositivity for a particular bacterial antigen. This threshold reliably indicated the presence of antibodies against a given bacterial antigen in the tested samples.

**Table 3.** The list of selected bacteria for the NAPPA array.

<table>
<thead>
<tr>
<th>Bacteria name</th>
<th>No of antigens</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bacteroides fragilis</em> NCTC 9343</td>
<td>25</td>
</tr>
<tr>
<td><em>Bifidobacterium adolescentis</em> ATCC 15703</td>
<td>5</td>
</tr>
<tr>
<td><em>Enterococcus faecalis</em> V583</td>
<td>18</td>
</tr>
<tr>
<td><em>Fusobacterium nucleatum subsp. nucleatum</em> ATCC 25586</td>
<td>42</td>
</tr>
<tr>
<td><em>Fusobacterium varium</em> ATCC 27725</td>
<td>10</td>
</tr>
<tr>
<td><em>Helicobacter pylori</em> 26695</td>
<td>113</td>
</tr>
<tr>
<td><em>Helicobacter pylori</em> J99</td>
<td>5</td>
</tr>
<tr>
<td><em>Lactobacillus plantarum</em> WCFS1</td>
<td>30</td>
</tr>
<tr>
<td><em>Porphyromonas gingivalis</em> ATCC 33277</td>
<td>31</td>
</tr>
<tr>
<td><em>Streptococcus galmolyticus subsp. galmolyticus</em> ATCC 43143</td>
<td>18</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>297</strong></td>
</tr>
</tbody>
</table>

Among the 297 bacterial proteins assessed on the NAPPA array, IgG antibodies targeting 42 of these proteins and IgA antibodies targeting 3 exhibited a seroprevalence of >10% in the samples (**Figure 20, Table 4**). The goal for this first stage was to limit the number of potential targets by excluding all the uninformative antigens. Notably, two-thirds (30/45) of these proteins
corresponded to bacterial proteins from *H. pylori*. The top five proteins with the highest seroprevalence were HP1341 (68% for IgG and 20% for IgA response), HP-1564 (61% for IgG), GroEL (50% for IgG) from *H. pylori*, a transglycosylase SLT domain-containing protein (64% for IgG and 11% IgA response), and an amino acid ABC transporter substrate-binding protein (57% for IgG and 11% for IgA response) from *S. gallolyticus*. Additionally, among these, four bacterial proteins were from *F. nucleatum*, including two from the OmpA family, one from the cell wall endopeptidase M23/M37 family, and another from the ABC transporter substrate-binding protein. Furthermore, only two *P. gingivalis* proteins elicited IgG antibody responses: FimA protein (18%) and Omp28 protein (11%).

Additionally, we categorized the subcellular localization of 45 immunodominant bacterial proteins based on their sequences using the Protein Subcellular Localization Prediction Tool (PSORTdb 4.0 [https://db.psort.org/]) (79)(86) (Table 4). The distribution revealed diverse subcellular locations: 27% were unknown, 7% periplasmic, 16% outer membrane, 13% extracellular, 7% cytoplasmic membrane, and 31% cytoplasmic. Notably, a significant number were identified in the bacterial cytoplasm. This varied distribution across subcellular locations may hold implications for their functional roles, host interactions, and potential involvement in disease processes.

We categorized these proteins into MSPs (Membrane or Secreted Proteins) and Non-MSPs, grouping them based on their subcellular locations. MSPs encompassed proteins from the outer membrane and extracellular space, while Non-MSPs included cytoplasmic, cytoplasmic membrane, and periplasmic proteins. Utilizing Fisher’s exact test, we assessed antigenic differences between MSPs and Non-MSPs. Results revealed that 29% of the proteins localized to membranes or secretion pathways, showing a 1.5-fold enrichment compared to the 20% found in Non-MSPs (p = 0.0395).
Table 4. List of oral, gut and probiotic bacteria-NAPPA array that showed >10% seroprevalence in the samples.

<table>
<thead>
<tr>
<th>Ab</th>
<th>Species</th>
<th>Protein</th>
<th>Gene symbol</th>
<th>Gene Description</th>
<th>Subcellular location</th>
<th>Samples (&gt;2)</th>
<th>Seropositivity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgA</td>
<td>Helicobacter pylori 26695</td>
<td>HpCD0781366</td>
<td>HP1341</td>
<td>energy transducer TonB</td>
<td>U</td>
<td>30</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781050</td>
<td>HP1564</td>
<td>ABC transporter substrate-binding protein</td>
<td>CM</td>
<td>27</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781014</td>
<td>grEL</td>
<td>molecular chaperone GroEL</td>
<td>C</td>
<td>22</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781066</td>
<td>fipE</td>
<td>flagellar hook protein FipE</td>
<td>EC</td>
<td>21</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781341</td>
<td>HP1125</td>
<td>outer membrane protein Omp18</td>
<td>OM</td>
<td>21</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0780975</td>
<td>fIA</td>
<td>flagellin A</td>
<td>EC</td>
<td>20</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781632</td>
<td>HP12871</td>
<td>biotin carboxyl carrier protein</td>
<td>C</td>
<td>18</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781632</td>
<td>HP1463</td>
<td>outer membrane beta-barrel protein</td>
<td>C</td>
<td>16</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781329</td>
<td>rplL</td>
<td>SOS ribosomal protein L7L12</td>
<td>C</td>
<td>14</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0780383</td>
<td>pOA</td>
<td>pyruvate flavodoxin oxidoreductase subunit alpha</td>
<td>C</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781308</td>
<td>HP0185</td>
<td>hypothetical protein</td>
<td>C</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781725</td>
<td>HP1488</td>
<td>efflux RND transporter</td>
<td>U</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781060</td>
<td>fglL</td>
<td>flagellar hook-associated protein FglL</td>
<td>OM</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781118</td>
<td>HP1293</td>
<td>DNA-directed RNA polymerase subunit alpha</td>
<td>C</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781213</td>
<td>HP0263</td>
<td>Hop family outer membrane protein HopU/HopK</td>
<td>U</td>
<td>10</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781516</td>
<td>HP1172</td>
<td>ABC transporter substrate-binding protein GinH</td>
<td>U</td>
<td>10</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0780459</td>
<td>recA</td>
<td>recombinase RecA</td>
<td>C</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781207</td>
<td>HP0264</td>
<td>chaperone protein ClpB</td>
<td>C</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0780792</td>
<td>hsiU</td>
<td>ATP-dependent protease ATP-binding subunit HsiU</td>
<td>C</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0780874</td>
<td>HP0231</td>
<td>protein-disulfide isomerase</td>
<td>U</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781010</td>
<td>HP1118</td>
<td>gamma-glutamy/transportase</td>
<td>P</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781127</td>
<td>HP1138</td>
<td>ParB/RepB/SpOj family partition protein</td>
<td>C</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781478</td>
<td>fglE</td>
<td>flagellar hook protein FglE</td>
<td>EC</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0780848</td>
<td>HP0875</td>
<td>catalase</td>
<td>P</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0780858</td>
<td>HP0407</td>
<td>biotin sulfoxide reductase</td>
<td>P</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0780923</td>
<td>HP1124</td>
<td>hypothetical protein</td>
<td>OM</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781387</td>
<td>fglL</td>
<td>flagellin B</td>
<td>EC</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781630</td>
<td>gyrB</td>
<td>DNA gyrase subunit B</td>
<td>C</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HpCD0781673</td>
<td>HP0385</td>
<td>hypothetical protein</td>
<td>U</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Streptococcus gallolyticus subsp. gallolyticus ATCC 43143</td>
<td>SgCD0783208</td>
<td>NA</td>
<td>transglycosylase SLT domain-containing protein</td>
<td>U</td>
<td>29</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SgCD0783208</td>
<td>NA</td>
<td>ABC transporter substrate-binding protein</td>
<td>U</td>
<td>25</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Enterococcus faecalis V583</td>
<td>EcCD0783252</td>
<td>NA</td>
<td>LPXTG cell wall anchor domain-containing protein</td>
<td>EC</td>
<td>15</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EcCD0783264</td>
<td>NA</td>
<td>LysM peptidoglycan-binding domain-containing protein</td>
<td>C</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EcCD0783265</td>
<td>NA</td>
<td>TPA serine protease</td>
<td>U</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Lactobacillus plantarum WCF51</td>
<td>LpCD0883821</td>
<td>ip_2978</td>
<td>WxL domain-containing protein</td>
<td>C</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LpCD0884135</td>
<td>NA</td>
<td>extracellular protein, NlpC/P60 family</td>
<td>EC</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Porphyromonas gingivalis ATCC 33277</td>
<td>PgCD0783526</td>
<td>fma</td>
<td>fimbral protein</td>
<td>OM</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PgCD0783414</td>
<td>omp28</td>
<td>Outer membrane lipoprotein</td>
<td>OM</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Fusobacterium nucleatum subsp. nucleatum ATCC 25586</td>
<td>FnCD0783099</td>
<td>NA</td>
<td>ABC transporter substrate-binding protein</td>
<td>CM</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FnCD0783088</td>
<td>NA</td>
<td>OmpA family protein</td>
<td>OM</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FnCD0783084</td>
<td>NA</td>
<td>OmpA family protein</td>
<td>OM</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FnCD0783898</td>
<td>NA</td>
<td>Cell wall endopeptidase family M23/M37</td>
<td>CM</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Helicobacter pylori 26695</td>
<td>HpCD0781366</td>
<td>HP1341</td>
<td>energy transducer TonB</td>
<td>U</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Streptococcus gallolyticus ATCC 43144</td>
<td>SgCD0783208</td>
<td>NA</td>
<td>ABC transporter substrate-binding protein</td>
<td>U</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SgCD0783206</td>
<td>NA</td>
<td>Transglycosylase SLT domain-containing protein</td>
<td>U</td>
<td>5</td>
<td>11</td>
</tr>
</tbody>
</table>

*EC extracellular, P Periplasmic, U unknown, CM cytoplasmic membrane, C cytoplasmic, OM Outer Membrane.

Upon assessing IgA responses, they predominantly demonstrated low levels, subsequently not meeting the criteria for further analysis. We conducted a heatmap analysis focusing on IgG antibody responses against 42 proteins with a seroprevalence exceeding 10% (Figure 20). The heatmap analysis in Figure 20 revealed an intriguing observation highlighted...
within the blue box: individuals in the late cancer stage, located on the left side, showed no IgG response against antigens from oral bacteria.

**Figure 20.** Heatmaps of IgG antibodies on bacteria-NAPPA (with >10% seroprevalence) by location and cancer stage in 44 colon cancer patients.

**Figure 21** graphs the seropositivity of IgG antibody response per bacterium by antigens consolidated between right- and left-sided location and colon cancer stage. There are differences in the seropositivity for IgG antibody against bacteria in location and cancer stage. Interestingly, most of the patients showed antibody response against *H. pylori*. More than 80% of the total of patients were positive for IgG response against *S. gallolyticus* in both early and late cancer stage, and was significantly increased in left-sided location (100% vs 91%, *p* = 0.003), similar to *B. fragilis*. 
Seropositivity for IgG antibody levels against oral bacteria, seropositivity for \( P. \) gingivalis was 46% and higher in the right-sided compared with 25% in the left-sided location (\( p=0.003 \)), like \( E. \) faecalis (73% vs 50%, \( p>0.001 \)). \( F. \) nucleatum and \( P. \) gingivalis, showed significantly increased seropositivity in early cancer stage when compared to late cancer stage (79% vs 33%, \( p<0.00001 \) and 42% vs 17%, \( p<0.001 \), respectively). Probiotic bacteria, \( B. \) adolescentis showed 25% seropositivity in late cancer stage compared to 0% in early stage (\( p<0.00001 \)) in contrast to \( L. \) plantarum which showed a decreased seropositivity in late stage (74% vs 42%, \( p<0.00001 \)).

**Figure 21.** Seropositivity of IgG antibody response per bacterium by antigens between right- and left-sided location and colon cancer stage. Fisher's exact test, right n=20, left n=24; early n=27, late n=17; ns= not significant, \* \( p<0.05 \), ** \( p<0.005 \), *** \( p<0.001 \), **** \( p<0.0001 \).

**Figure 22** presents data on the seropositivity of IgG antibody responses against specific antigens of bacteria associated with CRC. The Venn diagrams within the figure illustrate the presence of seropositivity for each individual target bacterium (\( E. \) faecalis, \( B. \) fragilis, \( S. \) gallolyticus, and \( H. \) pylori) as well as the combined targets across all patients. For example, when considering the prevalence of seropositivity for individual targets across all individuals such as
TonB, GroEL, and HP1564, the seroprevalence response for *H. pylori* is 68%, 50%, and 61%, respectively. However, when analyzing seropositivity in combination across all three targets, increases to 89%. Overall, the individual sensitivities ranged from 6.7% to 64.4% and the combined targets ranged from 20% to 89%. We found a seropositivity of more than 80% for IgG antibody response against *S. gallolyticus* and for *H. pylori*.

**Figure 22.** Seropositivity of IgG antibody response against individual and clustered antigens of CRC-associated bacteria. The Venn diagrams showcases the seropositivity observed for each individual target as well as the combined targets across all patients (n=44) for *E. faecalis*, *B. fragilis*, *S. gallolyticus*, and *H. pylori*.

Additionally, in **Figure 23**, we juxtapose the seropositivity of IgG antibodies against individual oral bacterial proteins with their combined targets. Individually, seroprevalence varied from 4.4% to 38%, while the combined targets showed a range of 29% to 38%. Building upon the absence of IgG antibody responses against oral bacteria proteins in the left location-late stage colon cancer, as evident in the highlighted area within **Figure 20**. We further investigated potential
differences in the seroprevalence of combined oral bacterial targets among different colon locations and cancer stages. This analysis was conducted using Fisher’s exact test. Interestingly, seropositivity for both, *P. gingivalis* and *F. nucleatum*, was significantly increased in right-sided colon and early cancer stage.

**Figure 23.** Seropositivity of IgG antibody response against individual and clustered antigens of oral bacteria *F. nucleatum* (A) and *P. gingivalis* (B) in the total patient’s population, and between right/left sided location and cancer stage groups. Fisher’s exact test, right n=20, left n=24; early n=27, late n=17; ns= not significant, * p < 0.05, ** p < 0.005, *** p < 0.001, **** p <0.0001.

**Figure 24** shows a diagram for positive/negative detection of *F. nucleatum* and *P. gingivalis* by qPCR in non-tumor and tumor tissues and by IgG and IgA antibody response by each patient. A Spearman’s rank-order correlation (p two-tailed, 95% CI) was run to determine the relationship between the relative abundance of *F. nucleatum* and *P. gingivalis* by qPCR in non-tumor and tumor tissues and IgG and IgA antibody response for these oral pathogens. There was a positive weak correlation found for *F. nucleatum* relative amount in colon tissue and IgA
antibody response in serum for antigens FnCD00783087 ($r=0.31, p<0.05$) and FnCD00783099 ($r=0.33, p<0.05$), and IgG against FnCD00783081 ($r=0.34, p<0.05$). Moreover, there was a negative weak correlation found for *P. gingivalis* relative amount in colon tissue and IgA antibody response in serum for antigens PgCD00783633 ($r=-0.39, p<0.01$) and PgCD00783418 ($r=-0.32, p<0.05$), and a positive weak correlation for IgG against PgCD00783522 ($r=-0.31, p<0.05$). Additionally, we conducted Fisher's Exact Test to analyze the presence/absence of antibody response by NAPPA with the presence/absence of each bacterium in tumors, as well as in non-tumors using 16S rRNA data. However, none of the obtained results reached statistical significance.

![Diagram](image)

**Figure 24.** Diagram for positive/negative detection of *F. nucleatum* and *P. gingivalis* by qPCR in non-tumor and tumor tissues and by IgG and IgA antibody response by each patient (n=44).

However, despite these compelling results, it's imperative to note that, up to this point, none of the proposed antibody markers have exhibited adequate discriminatory power to conclusively distinguish between colon cancer patients and those without the disease. This underscores the ongoing need for further research and refinement in identifying more precise and specific biomarkers for distinguishing individuals with colon cancer from healthy individuals.
Discussion- Aim 1

Heterogeneity between colon cancer tissues with *Fusobacteriota* only present in tumors

Several studies (18, 54, 87–89) indicate that the composition of the gut microbiome is a major factor in CRC risk with strong associations between intratumor microbial abundance and race, suggesting a significant influence of racial background on the microbiome (90). In our investigation, we aimed to bridge this gap by analyzing the gut microbial communities by 16S rRNA gene sequencing in 66 matched tumor and non-tumor samples from 33 AA patients.

We found that colon cancer tumors have a different microbiota profile than the non-tumors in AA patients. Tumor tissues exhibited reduced diversity, with *Firmicutes* (67.95%), *Bacteroidota* (15.38%), *Proteobacteria* (12.82%), and *Actinobacteria* (2.56%) being the predominant phyla. Notably, *Fusobacteriota* (1.28%) were exclusively found in the tumor group, consistent with findings from studies by Drewes et al. (20) and Roelands et al. (91), where patient race information was not disclosed. In contrast, Xu et al. (92) reported dissimilar results in 533 CRC patients. They observed *Proteobacteria* as the predominant phylum in CRC patient tissues, followed by *Firmicutes, Actinobacteria, and Bacteroidetes*, with relative abundances of 43.5%, 25.3%, 23.0%, and 5.1%, respectively. While the authors acknowledged a significant association between race and microbiota abundance, they did not conduct further analysis to compare microbial communities among racial groups. Although they did suggest the need for further analyses to differentiate patients of different races and identify biologically meaningful microbial markers, age, sex, and BMI exhibited minimal correlation with microbiota composition according to their findings (92).

AAs face disproportionate rates of both CRC incidence and mortality compared to CAs, yet the underlying factors driving this disparity remain elusive. Recently published research by Byrd and Wolf (55) underscores two key points: first, the significant role of the gut microbiome in CRC development, and second, the observed variations in the gut microbiome across diverse
individuals. This highlights the potential of the gut microbiome as a potent and actionable molecular mechanism underlying racial and ethnic disparities in CRC.

Farhana et al. (58) also proposed that the gut microbiome might play a role in shaping this difference. Their study revealed marked variations in species richness and diversity between AAs and CAs, with the AA cohort displaying an overall lower microbial diversity compared to their CA counterparts. Notably, Bacteroidetes exhibited higher abundance in AAs, with a significant elevation of pro-inflammatory bacteria such as *F. nucleatum*, and *Enterobacter* species, while probiotic bacteria like *Akkermansia muciniphila* and *Bifidobacterium* were more prevalent in CAs. These findings suggest that the distinct prevalence of pro-inflammatory bacteria, levels of beneficial bacteria, and variations in overall microbial diversity between AAs and CAs could provide valuable insights into the CRC disparity observed among these demographic groups. Probiotic bacteria play a protective role in the gut by suppressing the growth of opportunistic pathogens, thereby promoting gut health, and reducing the risk of infections. One plausible hypothesis stemming from these observations is that the diminished presence of probiotic bacteria in AAs may render them more vulnerable to colonization by opportunistic pathogens.

**Increased presence of oral-origin bacteria detected in right-sided colon cancer tumors**

Given the potential bacterial translocation from mouth to gut, we hypothesized increased oral bacteria in the most proximal section of the colon, the right-sided colon location. In our study, the differences between right-sided and left-sided colon tumors were striking when considering their microbial diversity and specific bacterial abundances. Right-sided tumors exhibited a notable decrease in alpha-diversity (Shannon index, \( p=0.011 \)), harboring nearly half the observed species compared to their left-sided counterparts (Observed ASVs, \( p=0.0093 \)).

Additionally, we observed that approximately 17% of the microbial composition in right-sided tumors consisted of bacteria associated with oral-origin bacteria, including *Fusobacterium*, *Prevotella*, *Treponema*, *Leptotrichia*, and *Porphyromonas*, compared to 13% in left-sided tumors.
These could be influenced by various factors related to the tumor microenvironment, host factors, and the local conditions in the colon. Right- and left-sided colon environments may differ in terms of pH, nutrient availability, oxygen levels, transit time, and other factors, where certain bacteria may thrive or struggle leading to variations in microbial diversity (2).

Few studies have investigated the microbiota across different colon locations, often relying on 16S rRNA gene sequencing of stool samples or Formalin-Fixed Paraffin-Embedded (FFPE) blocks, which offer only a partial representation of the microbiota in CRC. We are the first study to analyze the microbial difference between locations in fresh frozen tissues from colon cancer AA patients. Zhong et al. (21) compared the microbiota of left- and right-sided colon tumors using 16S rRNA sequencing analysis in Chinese patients with colon cancer. Colon cancer tissues were collected when the patient underwent colon surgery and were stored at -80°C until DNA extraction, similar to the methodology of our study. They found that right-sided colon microbial taxa were less abundant than those in the left-sided colon, and *Fusobacteria* was more abundant in right-sided colon tumors, which aligns with our findings.

Drewes et al. (20) highlighted a fundamental difference between CRC located on the right versus left side of the colon. They found that the presence of bacterial biofilm structures within the mucosal layer of the gut is a distinguishing characteristic found distinctly in right-sided CRC but noticeably absent in the left-sided variant. This finding introduces the notion that the arrangement and formation of microbial communities within the gut might profoundly impact the initiation and progression of the disease, presenting a pivotal advancement in our comprehension of CRC. Importantly, biofilm was predominantly composed of *Bacteroidetes*, *Lachnospiraceae*, *Fusobacterium* spp., and the *Proteobacteria* phylum. Our study found increased *Proteobacteria* (22.0% vs. 8.3%) and *Fusobacterium* spp. (10.4% vs. 2.6%) and decreased *Bacteroidota* (9.5% vs. 27.9%) in right-sided compared to left-sided tumors, suggesting that the microbial composition of right-sided tumors could correspond with the presence the presence of bacterial biofilm
structures. This association prompts intriguing insights into the possible colonization of oral pathogens related to periodontitis within the right-sided colon via biofilm structures.

Periodontitis, marked by an imbalanced oral microbiota featuring pathogens such as *P. gingivalis*, *Tannerella forsythia*, *Treponema denticola*, *P. intermedia*, and *F. nucleatum* (38,39), triggers chronic inflammation that can extend systemically, potentially influencing conditions like cancer (64,93). Research has associated moderate to severe periodontitis with an elevated risk of CRC. AA individuals have been found to exhibit a higher prevalence of periodontitis compared to other racial groups (67). However, studies examining the risk relationship between periodontitis and CRC specifically within this demographic are lacking.

The presence of predominantly oral-origin bacteria, previously unassociated with CRC, in early-stage tumors is a significant finding. *Fusobacterium necrophorum* and *F. nucleatum*, typically associated with periodontitis, were notably abundant in early-stage right-sided tumors, while *Treponema socranskii* and *Prevotella_7 NA* showed significant abundance in early-stage left-sided tumors. *Porphyromonas* abundance was increased in tumors. Ahn et al (54) showed with that the carriage of *Porphyromonas* (*P =0.05; OR = 1.44*) and of *Fusobacterium* (*P = 0.01; OR = 1.44*) in feces increased risks of CRC. *F. nucleatum* and *Porphyromonas spp.* have been observed to be significantly higher in colon tumors of AA compared with CA patients (15,57,58,60,68,94). The presence of these oral bacteria in early-stage colon tumors in our study raises intriguing questions about the potential role of periodontitis-associated bacteria in the pathogenesis of CRC. Further investigation is warranted to elucidate the mechanisms underlying this association and its implications for CRC development and progression, particularly in populations with a high prevalence of periodontitis like AA individuals.

Co-abundance of oral pathogens associated with *F. nucleatum* presence in colon tumors

*F. nucleatum*, a common oral microbiome component, is not commonly found in healthy colonic tissues(27), but it tends to colonize tumor tissues rather than non-tumor tissues in
individuals with CRC (25,29,83,95). This preference is linked to its lectin Fap2, which attaches to the prevalent polysaccharide Gal-GalNAc found on the surface of CRC cells (96). We identified a significant 62-fold increase in *F. nucleatum* levels within tumors compared to matching non-tumor tissues (*p*=0.0003) and it is associated with right-sided cancer (*p*= 0.0253) and advanced cancer stage (*p*=0.0323). Similar to our findings, Castellarin et al (25) reported that the mean overall abundance of *Fusobacterium* was found to be 415 times greater in the tumor samples, Mima et al (97) also showed that the amount of *F. nucleatum* was higher in colorectal carcinoma tissue than in paired adjacent non-tumor tissue (*P* < 0.0001).

Recent research suggests that *F. nucleatum* may play a more pivotal role as ‘driver’ in colon cancer than previously thought (28), potentially serving as an early colonizer in the epithelium and acting as a crucial "bridge" bacterium within biofilms by adhering to a diverse range of species through surface adhesins (84). Building on this, our hypothesis speculates that *F. nucleatum* could play a significant role by co-aggregating with other oral pathogens and biofilm-associated bacteria, thereby altering microbial communities within colonic tumors. To investigate this further, we integrated data from qPCR and 16S rRNA gene sequencing to compare microbial communities in samples with and without *F. nucleatum*, aiming to uncover any distinctive patterns associated with its presence.

We found that Fn-positive and -negative groups were heterogenous in their microbial abundance. Fn-positive tumors displayed significant increases in various genus-level bacteria such as *Pseudomonas, Fusobacterium, Leptotrichia, Prevotella, Diallister, Treponema, Campylobacter*, and *Porphyromonas*. In contrast, Fn-negative tumors showed rises in *Streptococcus, Escherichia-Shigella, Haemophilus, Lactobacillus, and Gemella*. These shifts highlight distinct bacterial differences between Fn-positive and Fn-negative tumors, showing an association of *F. nucleatum* presence with dysbiosis-related gut microbiota in colon cancer tumors, similar to other studies (93,98–101).
By clustering bacteria groups, we found a positive correlation between Fn-positive tumors and increased levels of “Oral pathogens” and “biofilm-associated bacteria”. The oral pathogens, *Fusobacterium*, *Porphyromonas*, *Prevotella*, and *Treponema* were significantly increased in association with the presence of *F. nucleatum* in tumors. This occurrence isn't surprising considering they are all Gram-negative anaerobic microbes that typically occupy the same space within the oral cavity in the subgingival plaque (23). Our findings are aligned with the study by Flynn et al. (94) that suggested that oral bacteria may work together in a symbiotic or cooperative manner to create an ecological niche within the gut microbiome, potentially contributing to the development or progression of certain conditions such as colon cancer. In this oxygen-deprived setting, oral bacteria such as *P. intermedia*, *F. nucleatum*, and *P. gingivalis* co-aggregate potentially leading to ongoing persistent inflammation and facilitate the progression of CRC (94).

We also noted that the Fn-positive group exhibited elevated levels of taxa associated with the "Oral bacteria" cluster. Consistent with previous human studies, there appears to be a correlation between *F. nucleatum* and other Gram-negative bacteria such as *Streptococcus*, *Campylobacter*, and *Leptotrichia* (23). Our investigation revealed that in the presence of *F. nucleatum* within tumor tissues, there was a positive correlation with increased relative levels of *Leptotrichia* and *Campylobacter*, while *Streptococcus* levels were decreased under similar circumstances.

Additionally, we found that Fn-positive tumors had lower levels of 'Probiotics' and 'CRC pathogens'. *F. nucleatum* presence was linked to shifts in the microbial community, elevating abundance of oral pathogens while reducing beneficial bacteria. Reductions in these beneficial bacteria can disrupt the homeostatic state in the gut, causing dysbiosis and compromising the intestinal barrier. This disruption might heighten intestinal permeability, allowing toxins and bacteria to enter the bloodstream, potentially instigating inflammation and health complications (99,102,103). These findings align with the study by Farhana et al (58) where AA cohort displayed an overall lower abundance of probiotic bacteria compared to their CA counterparts.
Microbiome and the modifiable factors related to colon cancer

Roughly more than half of newly diagnosed cases of CRC in the US are believed to stem from factors that can be modified, including smoking, excessive alcohol consumption, and being overweight or obese (4). Factors that increase the risk of CRC also affect the gut microbiome, include overweight and obesity. In our study, the mean BMI for the group was 26.78 ± 4.11, with 36.1% classified as overweight and 19.2% as obese. Obesity has been associated with a significant decrease in the diversity of the gut microbiota (17). A recent meta-analysis (104) found a weak connection to BMI, and the risk of CRC when considering certain gut bacteria. Additionally, microbial dysbiosis was not a significant factor in explaining the higher risk of colon cancer among individuals with obesity. To better understand the role of gut microbiota in obesity-related CRC risk, more studies are needed that control variables such as diet, exercise, healthy habits, and others, in addition to BMI.

Regarding alcohol use, 31.8% of patients reported ever consuming alcohol, while the majority (68.2%) had never consumed alcohol. Chronic consumption of excess alcohol increases the risk of CRC and is partially mediated by gut microbiota. A study (105) found a synergistic effect of habitual drinking and smoking on the structural alteration of gut microbiota. Additionally, bacteria residing in the intestinal environment of alcoholics likely face higher ROS exposure than those in non-alcoholic individuals possibly increasing gut dysbiosis (105). However, we don’t fully understand how alcohol affects the microbial diversity and bacterial community in the human gut and how these bacteria contribute to CRC.

The patient cohort's medical history indicated a absence of Type 1 diabetes, with 29.8% of patients having Type 2 diabetes. Regarding smoking history, 22.7% of patients reported a history of smoking, while the majority (77.3%) were non-smokers. A comprehensive review revealed correlations between dental conditions and chronic systemic diseases, with periodontitis
being among the most frequently observed. Smoking is known to heighten the risk and severity of periodontitis (66). Notably, periodontitis exhibited numerous correlations with Type 2 Diabetes, suggesting a bidirectional relationship between the two conditions. Individuals with Type 2 diabetes may be predisposed to periodontal disease, which, in turn, could impact blood sugar control in diabetic individuals.

**Increased IgG response compared to IgA against selected bacteria in colon cancer**

Understanding the immune response to bacteria is crucial in deciphering factors affecting susceptibility to colon cancer. Our study evaluated the IgG and IgA antibody responses against bacteria linked to CRC. Among the 297 bacterial proteins assessed on the NAPPA array, IgG antibodies targeting 42 of these proteins and IgA antibodies targeting 3 exhibited a seroprevalence of >10% in the samples. An increased IgG response compared to IgA may suggest a more systemic or long-term immune response to the presence of the specific bacteria. This could potentially be due to a chronic or persistent infection, or a prolonged exposure to the bacteria.

Additionally, we categorized the subcellular localization of 45 immunodominant bacterial proteins based on their sequences using the Protein Subcellular Localization Prediction Tool (PSORTdb 4.0 [https://db.psort.org/]) (79,86). Remarkably, a significant number were identified in the bacterial cytoplasm which may have varying degrees of exposure to the host immune system compared to those located on the bacterial surface or secreted outside the bacterial cell. This variation in localization could impact the way antibodies like IgA and IgG recognize and respond to these bacterial proteins, thereby affecting the host's immune defense mechanisms and potentially contributing to disease processes.
Further analysis is necessary to establish the distinction between bacterial core genome proteins and accessory proteins. Proteins constituting the core genome are often fundamental to the bacterial species' survival, functionality, and basic biological processes. On the other hand, accessory proteins, being strain-specific or condition-specific, might contribute to traits like virulence, antibiotic resistance, or niche adaptation. Additionally, understanding the distribution of core and accessory proteins across different subcellular locations can offer insights into the immune response triggered by specific bacterial strains or species. This knowledge might uncover potential targets for therapeutic interventions, vaccine development, or diagnostics aimed at specific bacterial strains or conditions.

Increased anti- \textit{H. pylori} and \textit{-S. gallolyticus} IgG response in colon cancer

Out of the 297 bacterial proteins examined on the NAPPA array, IgG antibodies against 42 of these proteins and IgA antibodies against 3 showed a seroprevalence exceeding 10% in the samples. Remarkably, the majority (30 out of 45) of these proteins belonged to \textit{H. pylori}. It's worth noting that a striking 89% of the patient cohort exhibited an immune response against \textit{H. pylori}. This highlights the significant prevalence of \textit{H. pylori}-associated antibodies among the study participants.

A study conducted by Shmuely et al. (106) tested serum IgG antibodies against \textit{H. pylori} by ELISA test in 67 patients with colorectal adenocarcinoma and 36 with gastric adenocarcinoma. In their investigation, they detected IgG antibodies against \textit{H. pylori} in 74.6% of colorectal adenocarcinoma patients, similar to our study. Moreover, a meta-analysis study (107) provided additional evidence supporting the positive association between \textit{H. pylori} infection and CRC risk. This bacterial infection is widespread globally, with almost half of the population carrying it, and prevalence rates soaring to 80–90% in certain developing nations. Even in the US, approximately
35–40% of the population is affected by *H. pylori* infection, which could account for the elevated seroprevalence of anti-*H. pylori* antibodies observed in our cohort population.

TonB is a crucial component of the energy-transferring complex within *H. pylori*, facilitating the transport of nutrients across the bacterial membrane. In our study, we found that TonB protein elicited the highest seroresponse among all the bacterial proteins analyzed within our patients' cohort, with a prevalence of 68%. This finding aligns with previous research conducted by our collaborators Camargo and Qiu (108), also using NAPPA technology to explore anti-*H. pylori* antibody responses in gastric intestinal metaplasia (IM), a precancerous lesion, and non-atrophic gastritis (NAG). Their study revealed an exceptionally high seropositivity of 98% for HP1341/TonB in IM and NAG combined. Moreover, in the same study (106), positive associations were found with other *H. pylori* proteins, such as flagellar antigens HP0870/FlgE and HP0601/FlaA, with overall prevalence ranging from 40% to 58% across all individuals. We also observed similar trends, with seroprevalence for HP0870/FlgE and HP0601/FlaA closely matching these findings, with seroprevalences of 48% and 45%, respectively.

Interestingly, the authors observed that TonB elicited a notably higher seroresponse in individuals with current *H. pylori* infection compared to those with a past infection history. This suggests that TonB could potentially serve as a sensitive marker for active *H. pylori* infection and could potentially play a role in disease progression, warranting further investigation into its diagnostic and prognostic utility. In our study, no correlation was found between the presence of *H. pylori* by 16S rRNA gene sequencing data in non-tumor and tumor tissues and seropositive IgG and IgA antibody response.

*H. pylori* has coevolved with humans, evolving sophisticated strategies to evade detection by the host immune system. Additionally, despite widespread exposure to *H. pylori*, there is no consensus on which bacterial proteins are universally recognized across all infected individuals.
(108). This highlights the variability in immune responses to H. pylori infection among individuals, resulting in differences in the recognition of bacterial proteins. This variability underscores the complex interplay between the bacterium and the host immune system, contributing to the challenges in understanding and effectively combating H. pylori-related diseases.

Furthermore, we observed that over 84.4% of participants exhibited seropositivity for IgG antibodies against S. gallolyticus. This aligns with the findings by Abdulamir et al. (109) where CRC patients exhibited elevated serological levels of S. gallolyticus IgG antibodies compared to control subjects and healthy volunteers as determined by ELISA testing (68% of seropositivity). Upon applying the cutoff value to healthy volunteers, CRC, and adenoma groups, the authors suggested that it could effectively screen high-risk population groups with sensitivities ranging from 68 to 78% for CRC and adenoma, with a specificity of 83.33%. Thus, this assay holds promise for developing an optimized kit for the screening and detection of CRC or colorectal adenoma in the population using a simple, inexpensive, and noninvasive procedure.

In a previous study also by Abdulamir et al. (110) reported a correlation between S. gallolyticus bacteremia and colorectal tumors in a significant percentage of cases, ranging from 25% to 80% (110). The hypothesis suggests that intestinal lesions serve as the point of entry for the commensal S. gallolyticus into the bloodstream, allowing the bacterium to transition from a harmless state to a pathogenic one, potentially leading to bacteremia or endocarditis. (111).

Furthermore, in a prospective study, it was noted that 45% of individuals diagnosed with S. gallolyticus endocarditis subsequently developed colonic neoplastic lesions within five years, providing compelling evidence for the association between this bacterium and CRC (112). Additionally, antibody responses to S. gallolyticus, in particular to two or more proteins seropositive among a 6-marker panel (Gallo0272, Gallo0748, Gallo1675, Gallo2018, Gallo2178 and Gallo2179, proteins not included in our study) were significantly associated with risk of
developing CRC, in a CRC case–control study nested within the prospective multinational EPIC cohort (European ancestry individuals) (113).

Interestingly, these findings prompt further investigation into two key areas. Firstly, exploring whether \textit{S. gallolyticus} may indirectly contribute to the development of colonic neoplastic lesions and tumors originating from distant sites, such as in cases of bacteremia and endocarditis. Secondly, determining the potential of antibody assessment against \textit{S. gallolyticus} as a preventive measure in controlling the occurrence of colon cancer. These avenues of research hold promise for advancing our understanding of the role of \textit{S. gallolyticus} in CRC development and potential preventive strategies.

Combined bacterial proteins of \textit{P. gingivalis} and \textit{F. nucleatum} are elevated in early right-sided colon cancer

Among the 42 bacterial proteins showing the highest seroprevalence in our study cohort, four were from \textit{F. nucleatum}. These included two proteins from the OmpA family, with seroprevalences of 14% and 11%, one protein from the cell wall endopeptidase M23/M37 family (11%), and another from the ABC transporter substrate-binding protein (18%). Additionally, only two proteins from \textit{P. gingivalis} triggered IgG antibody responses: the FimA protein (18%) and the Omp28 protein (11%). Other studies (114,115) has investigated IgA and IgG antibodies against \textit{F. nucleatum} using ELISA assays to assess their potential diagnostic value in CRC. While these studies reported higher optical density (OD) values for \textit{F. nucleatum} IgA and IgG in CRC patients compared to healthy controls, standalone ELISA tests lacked the discriminatory power necessary for accurate diagnosis. However, Wang et al. (115) demonstrated that integrating anti-Fn-IgA with carcino-embryonic antigen (CEA), a protein considered a tumor marker in serum, significantly improved CRC diagnostic accuracy. This highlights the potential of combining multiple targets to enhance their efficacy as biomarkers for colon cancer detection.
In our investigation, we evaluated the sensitivity of individual *F. nucleatum* proteins, observing a range of 8.9% to 17.8% when assessed individually. However, when these proteins were combined, the sensitivity notably increased, spanning a broader range of 37.8%. These results highlight the effectiveness of analyzing multiple bacterial proteins simultaneously, demonstrating enhanced sensitivity compared to assessing them individually. This approach shows potential for advancing diagnostic accuracy and precision in identifying biomarkers associated with CRC.

Interestingly, seropositivity for both *P. gingivalis* and *F. nucleatum* was significantly increased in right-sided colon location and early cancer stage for combined targets FimA + PgCD00783414 + PgCD00783523, FNCD00783099 + FNCD00783088 + FNCD00783084, and FNCD00783868 + FNCD00783100. Consistent with our findings, Wang et al. (115) demonstrated that combining anti-Fn-IgA with CEA and carbohydrate antigen 19-9 (CA19-9) improved the classification of CRC patients with early-stage cancer (stages I-II). These results suggest that Fn infection induces a robust serum anti-Fn antibody response in CRC patients, with serum anti-Fn-IgA levels potentially serving as a diagnostic biomarker for CRC. Further investigations into humoral immunoproteomic profiles and longitudinal studies are essential to validate the identified bacterial targets for potential application in diagnosing colon cancer and determining its location and stage.

Overall, African-descent Caribbean populations show lower rates in recommended screenings for CRC (116). Moreover, the prognosis disparity associated with right-sided colon cancer emphasizes the urgent need for a non-invasive diagnostic test tailored for this population. A targeted effort in developing and validating a non-invasive screening tool that specifically identifies right-sided colon cancer in its early stages could immensely benefit prognosis and treatment outcomes within this demographic.
In our study, we identified weak correlations between the relative abundance of *F. nucleatum* and *P. gingivalis* in colon tissues and IgG/IgA antibody responses in serum, suggesting potential interactions between tissue-based bacterial abundance and systemic immune response. Similarly, Kurt et al. (114) not only assessed IgA and IgG antibodies against *F. nucleatum* via ELISA but also detected *F. nucleatum* DNA in tumors using qPCR. They found *F. nucleatum* DNA present in 86.7% of tumors in CRC patients. Interestingly, they observed a significantly higher positivity rate of antibodies in patients with a high abundance of *F. nucleatum* in tissue compared to the negative group. These findings underscore the importance of considering both local tissue abundance and systemic antibody responses when investigating the role of *Fusobacterium nucleatum* in CRC.

Furthermore, we employed Fisher's Exact Test to examine the association between antibody response via NAPPA and the presence or absence of individual bacteria in tumors and non-tumors, utilizing 16S rRNA data. However, none of the obtained results reached statistical significance. One potential explanation is that the antibody response may have been elicited because of past bacterial infections, leading to the persistence of immune memory. Alternatively, it is conceivable that the bacteria detected by IgA or IgG antibodies in the serum may be localized in other anatomical sites, such as the oral cavity or respiratory tract, where they could have triggered an immune response. It is also possible that the 16S rRNA sequencing method used to detect the bacteria in the gut may not have been sensitive enough to detect low levels of the bacteria, or that the bacteria were present in a different part of the gut that was not sampled.

Our study is the first effort in characterizing the gut microbiome within tumor tissue compared to matched non-tumor tissue in a high-risk population, while also examining the antibody response against CRC-associated bacteria. Utilizing matched pairs of individual tumor and non-tumor tissues offers several advantages. Firstly, it enables a direct comparison between samples from the same individual, thereby mitigating individual-specific and anatomical site-
specific effects. This approach enhances the accuracy and reliability of our findings by minimizing confounding variables that may arise from inter-individual variability or distinct anatomical locations within the colon (57).

Our investigation encountered several limitations that warrant consideration. Firstly, the limited sample size, while providing valuable insights, underscores the need for further investigations with larger cohorts, the inclusion of control groups, and external validations to substantiate and generalize our findings. Secondly, the availability of frozen fresh colon tissues for analysis was limited to 33 patients, imposing constraints on the number of samples that could be effectively matched between different analyses. Thirdly, a more comprehensive assessment of oral health, coupled with the collection of oral samples for microbial analysis, would contribute to a deeper understanding of the intricate interplay within the oral-gut axis in CRC.

Future research efforts will focus on augmenting the number of oral samples and colon tissues collected from the same patients within our study cohort. This expansion aims to provide a more comprehensive investigation into the relationship between the oral-gut axis and their potential association with colon cancer in AA patients. Specifically, we aim to compare the microbial abundance and identify differential bacteria present in both niches. This will involve examining common elements between the oral cavity and this specific location in the colon to elucidate potential mechanisms underlying these observed similarities.

Crucial questions remain regarding the origin and behavior of bacteria found in colon tumors. Understanding whether these bacteria originate from the oral cavity or are unique to the intestines is essential. If they do originate from the mouth, it is important to determine whether they migrate through the GI tract as a biofilm structure, surviving the harsh conditions of the stomach and maintaining the structural elements of oral biofilms. Future research should focus on retrieving oral and colonic mucosal biofilms for comprehensive analyses, including bacterial
community composition and functional profiling, to identify commonalities that may shed light on their role in disease development. Meta-transcriptomic analysis of bacteria derived from human CRC samples could provide insights into the activity of oral microbes in the colon, examining factors such as biofilm matrix proteins, adhesins, quorum-sensing molecules, and virulence proteins. Additionally, utilizing a mouse model of infection may offer further insights once fundamental microbial groups or functions are identified.

Furthermore, experimental interventions such as testing treatments or probiotics in animal models, particularly mouse models of CRC, could offer invaluable insights into potential therapeutic strategies. By administering targeted treatments or probiotics and monitoring their effects on the microbiome and disease progression, researchers can elucidate the efficacy and mechanisms of action of these interventions by potentially competing against opportunistic bacteria as a preventive measure for CRC. Ultimately, these endeavors hold the promise of advancing our understanding of the complex interplay between microbial communities and disease development, paving the way for novel therapeutic approaches in the management of CRC and other related conditions.

Moreover, additional studies with larger sample sizes are warranted to validate the identified associations and to assess the potential utility of microbiota signatures as biomarkers for early detection or prevention of CRC. Additionally, longitudinal studies and thorough analyses of humoral immunoproteomic profiles are essential to validate bacterial targets for potential use in diagnosing, locating, and staging colon cancer. The underutilization of recommended screenings for CRC in AA populations, coupled with prognosis disparities in right-sided colon cancer, underscores the pressing need for a non-invasive diagnostic tool tailored to this at-risk population group. A targeted approach to developing and validating such a screening tool could significantly enhance prognosis and treatment outcomes for right-sided colon cancer in the early stages of the diseases.
**Conclusions - Aim 1**

In summary, our study shows differences in microbial composition between tumor and non-tumor tissues in AA colon cancer patients, with unique bacterial profiles. Tumor tissues displayed lower diversity, especially in right-sided tumors. *F. nucleatum* was significantly more abundant in tumors, especially in right-sided and late-stage cancers, impacting the microbial community by elevating oral pathogens and reducing beneficial bacteria. Antibody responses varied across cancer stages and colon tumor locations, most responsive proteins belonged to *H. pylori*, suggesting a significant immune recognition of these antigens. Weak correlations suggested potential links between tissue-based bacterial abundance and systemic immune response. The increased seropositivity for *P. gingivalis* and *F. nucleatum* in right-sided colon cancer and early cancer stages requires validation in larger diverse cohorts. Prospective studies employing specific antibody response targets could be pivotal for early diagnosis.
Aim 2: To evaluate systemic levels of inflammatory mediators, colon tumor gene expression and T cell subsets densities within tumors of African American colon cancer patients.

Subaims:

2a) To assess the gene expression of immune-oncology markers in colon cancer tumors

2b) To analyze serological levels of inflammatory mediators in colon cancer patients

2c) To evaluate T-cell subsets density in colon cancer tumors

Research design and methods - Aim 2

The authors’ contributions to this study are as follows:

- Sofia Tortora (ST) at SUNY Downstate, Division of Gastroenterology:
  • Project design, hypothesis, and research strategy.
  • IRB oral sample-related protocols and revisions
  • Processing of samples and storage
  • Recording of clinical data
  • RNA and DNA isolation from tissue samples
  • Serum, plasma, and buffy coat isolation from blood
  • Quantibody array analysis
  • Immunohistochemical staining analysis with the assistance of expert pathologist Dr. Jianying Zeng (JZ), pathology resident Dr. John Agboola (JA) (SUNY Downstate).

- Santhilal Subhash, PhD and Semir Beyaz, PhD, CSHL.
  • RNA isolation from tumor/non-tumor colon tissues for RNA sequencing
  • Data analysis of bulk RNA sequencing
**Cytokine Glass Slide-Based Array:** Serum samples (100 μL) from colon cancer patients was quantitively analyzed for the concentration of 20 proteins associated with human periodontal disease such as pro- and anti-inflammatory cytokines (IL-1α, IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, IL-12, IL-17, IFN-γ, TNF-α, TGF-β1); matrix metalloproteinases (MMP-9, and MMP-13) and their inhibitors (MIP-1α); bone metabolism related cytokines (osteofractoprotein (OPG), osteopontin (OPN), osteoactivin (OA), and receptor activator of nuclear factor kappa-B (RANK)); and acute phase protein (C-reactive protein) using a Quantibody® Human Periodontal Disease Array (QAH-PDD-1; RayBiotech, Inc., Norcross, GA, USA). The microarray used a pair of cytokine-specific antibodies for detection in which the capture antibody was bound to the glass surface of slides (117). Each microwell in the glass slide has four spots with antibodies for each specific cytokine. After incubation with samples, the target cytokine was trapped on the glass surface. Then biotin-labeled detection antibody was added, which recognized a different epitope of the target cytokine than the capture antibody. For protein quantification, the assay kit used recombinant protein standards at concentrations determined by the manufacturer. These reagents were used to generate a six-point standard curve for each protein using the microarray kit. The cytokine-antibody-biotin complex was detected with a streptavidin-labeled fluorescent dye by a microarray laser scanner. Data extraction was done with microarray analysis software, Q-Analyzer Tool, Excel-based program provided by RayBiotech for the quantitative analysis (118). The software can automatically identify the outlier spots for more accurate data analysis, utilizes the two positive controls in each array for normalization, and determines the values below or above the detection range.

**mRNA expression analysis:** RNA from tumor tissues from AA colon cancer patients were used for the unbiased detection of 770 immune-related genes that cover immune-oncology genetic pathways such as NF-kB, cytokines, chemokines, T and B cell activation and differentiation, myeloid cell recruitment and activation, ROS, natural killer cell activation, innate immunity, among
others. Gene expression was previously analyzed in tumor and non-tumor tissue samples by Dr. Jenny Paredes using NanoString Technologies; specifically, the "nCounter Analysis System" which is an analysis based on the binding of fluorescent-color probes to mRNA molecules and the further correlation of the probes to gene markers of specific cell populations. We selected specific genes for analysis by combining that list from nCounter® PanCancer Immune Profiling Panel (https://nanostring.com/products/ncounter-assays-panels/oncology/pancancer-immune-profiling/) and from the paper by Galeano Niño et al (119). The RNA-seq samples were aligned using STAR (v2.5.2a)(120) with parameters ‘-alignIntronMax 200000 --alignMatesGapMax 200000 --chimSegmentMin 15 --chimJunctionOverhangMin 15’ mode using GRCh38 human reference genome. The reference genome was indexed using STAR with overhang length of 124 before samples were subjected to alignment. Quantification of reads to genic features was done using feature ‘Counts’ from subread (v1.4.3-p1)(121). Batch effects were corrected using Limma Bioconductor package. The quantified and batch-corrected reads per genes were subjected to normalization and differential expression analysis using DESeq2 (122) from Bioconductor package. To obtain Gene Description from Gene ID we used GTEx Portal (https://gtexportal.org/) and for the Summary of the gene functionality we used Gene Tool from the NIH website (https://www.ncbi.nlm.nih.gov/gene/) (Supplementary Tables 1-6).

**Immunohistochemistry and quantification of the density of T-cells in tumor tissue:**
Five-micron whole tissue sections from 10% formalin-fixed paraffin-embedded blocks of primary colon cancer tissues were made. Sections were deparaffinized and stained with H&E to identify tumor foci and peritumor areas of tissue sections. Immunohistochemical (IHC) staining of primary colon cancer tissue for CD3, CD4, and CD8 were performed on automated immunostainer (Ventana BenchMark ULTRA, IN) according to the manufacturer’s protocol at the Department of Pathology of the University Hospital of Brooklyn. The antibodies used for IHC staining are prediluted and are as follows: CD3 (Rabbit monoclonal [2GV6], Ventana, IN); CD4 (Rabbit
monoclonal [SP35], Ventana, IN); CD8 (Rabbit monoclonal [SP57], Ventana, IN). The antigen retrieval was performed with Ventana ULTRA Cell Conditioning Solution (Ultra CC1) at 100°C. Appropriate positive controls were included using tonsil tissue sections.(123) On each scanned slide, the tumor intraepithelial and stroma areas were manually annotated by ST with the assistance of a trained surgical pathologist (JZ) and pathology resident (JA) using open source digital pathology software (QuPath v0.4.3). The number of lymphocytes per square millimeter (density) was determined using fixed parameters with the software’s automated positive cell detection function and posterior manual corroboration (9). We included between 5-7 fields of counting for each tissue section.

**Statistical analysis:** The analysis of the statistical significance of differences in serum levels was performed with the Mann-Whitney U test for non-normally distributed data. The correlation of serological levels of cytokines and the amount of *F. nucleatum* DNA in colon tumors was analyzed by Spearman correlation, α level of 0.05. For the RNA-seq analysis, the quantified gene reads, corrected for batch effects, underwent normalization and differential expression analysis using DESeq2. The genes with an absolute log2 fold change of at least 1 and a p-value below 0.01 were selected as differentially expressed. This set of genes, consisting of both significantly upregulated and downregulated ones, underwent functional enrichment using Gene Ontology and KEGG databases. Pathways were considered significant if they exhibited a p-value below 0.05. For the IHC staining analysis, the correlation between *F. nucleatum*/*P. intermedia* DNA amount in tumor tissue and density of T-cell subsets in tumor intraepithelial and stromal areas was analyzed by Spearman correlation test, α level of 0.05. Further examination of the T cell subsets density between *F. nucleatum*-negative and -positive groups was done by Mann-Whitney U test, α level of 0.05.
Results – Aim 2

2a) To study the gene expression of immune-oncology markers in colon tumors

We conducted bulk RNA sequencing on fresh adjacent colon tumor tissues obtained from 40 AA patients diagnosed with colon cancer (23 right-sided and 17 left-sided location; 25 early stage I-II and 15 late cancer stage III). The gene expression of 738 immune-oncology markers was analyzed in these tumor tissues to elucidate their molecular landscape and potential implications for the disease. The list of genes was selected based on the combination of genes analyzed in the paper published by Galeano Niño et al. (119) and the system nCounter® PanCancer Immune Profiling Panel (Nanostring, https://nanostring.com/products/ncounter-assays-panels). Genes with an absolute log2 fold change > 1 and a p-value > 0.01 were selected as differentially expressed. We summarized for the selected genes their description and functionality in Supplementary Tables 1-6.

First, we compared the gene expression between right- and left-sided tumor locations (23 and 17 patients, respectively) (Figure 25). The top upregulated genes in right-sided tumors included MUC6, a member of the mucin protein family, TFF2, which plays a role in mucus layer stabilization, MSLN, and CDH7, both functioning as cell adhesion proteins. Additionally, genes from the Homebox family (HOXC5, HOXC6, HOXC9, HOXC10, HOXC11), as well as genes associated with development such as EN1 and ARMC3, were upregulated. Moreover, numerous genes involved in the chemotaxis and activation of immune cells, including CXCL8, CXCL11, CXCL17, IL1A, LCN15, HCAR2, FPR2, and PPBP, were also upregulated (Supplementary Table 1).

The downregulated genes included those associated with anatomical functions specific to the left (distal) colon, such as CLDN8, which regulates sodium back-leakage, and PRAC1 reported to be specifically expressed in prostate, rectum and distal colon. Additionally, genes
related to neuro-endocrine functions, such as CHGA, PYY, GCG, and INSL5, were downregulated. Furthermore, several genes involved in controlling channels and transport mechanisms, including BEST2, SLC30A8, SCNN1G, and SCNN1B, were also downregulated. (Supplementary Table 2).

**Figure 25.** Volcano plots for right-sided versus left-sided tumor location of colon cancer for RNA sequencing data. Statistical comparison of gene expression levels by DESeq2, selection of differentially expressed genes based on an absolute log2 fold change >1 and a p-value >0.01. (Right-sided location n=23 and left-sided location n=17)

The top gene upregulated pathways in right-sided tumor location were TNFA signaling via NFKB, inflammatory response, cell adhesion, neutrophil chemotaxis, L-glutamate import across plasma membrane, chemokine-mediated signaling pathway, and others. The top gene downregulated pathways were ascorbate and aldarate metabolism, insulin secretion, bile secretion, pentose and glucuronate interconversions, vascular smooth muscle contraction, calcium-ion regulated exocytosis, among others (Table 5).
Table 5. The pathways associated with right- and left sided colon cancer from the RNA sequencing data.

<table>
<thead>
<tr>
<th>Right-sided colon cancer</th>
<th>p-value</th>
<th>Left-sided colon cancer</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNFA signaling via NFKB</td>
<td>0.000004</td>
<td>Ascorbate and aldarate metabolism</td>
<td>0.000001</td>
</tr>
<tr>
<td>Inflammatory response</td>
<td>0.00003</td>
<td>Insulin secretion</td>
<td>0.000002</td>
</tr>
<tr>
<td>Cell adhesion</td>
<td>0.0008</td>
<td>Bile secretion</td>
<td>0.000002</td>
</tr>
<tr>
<td>Neutrophil chemotaxis</td>
<td>0.001</td>
<td>Pentose and glucuronate interconversions</td>
<td>0.000002</td>
</tr>
<tr>
<td>L-glutamate import across plasma membrane</td>
<td>0.004</td>
<td>Vascular smooth muscle contraction</td>
<td>0.00003</td>
</tr>
<tr>
<td>Chemokine-mediated signaling pathway</td>
<td>0.004</td>
<td>Retinoic acid binding</td>
<td>0.0001</td>
</tr>
<tr>
<td>Regulation of keratinocyte differentiation</td>
<td>0.007</td>
<td>Taste transduction</td>
<td>0.0001</td>
</tr>
<tr>
<td>Bone mineralization</td>
<td>0.007</td>
<td>Calcium-ion regulated exocytosis</td>
<td>0.0002</td>
</tr>
<tr>
<td>Cytokine-cytokine receptor interaction</td>
<td>0.008</td>
<td>Salivary secretion</td>
<td>0.0002</td>
</tr>
<tr>
<td>Thiamine metabolism</td>
<td>0.011</td>
<td>Chloride channel activity</td>
<td>0.0003</td>
</tr>
<tr>
<td>Biosynthesis of cofactors</td>
<td>0.017</td>
<td>Regulation of lipolysis in adipocytes</td>
<td>0.002</td>
</tr>
<tr>
<td>cAMP signaling pathway</td>
<td>0.021</td>
<td>Steroid hormone biosynthesis</td>
<td>0.002</td>
</tr>
<tr>
<td>Folate biosynthesis</td>
<td>0.029</td>
<td>Glucagon signaling pathway</td>
<td>0.003</td>
</tr>
<tr>
<td>Antifolate resistance</td>
<td>0.037</td>
<td>Serotonergic synapse</td>
<td>0.004</td>
</tr>
<tr>
<td>beta-Alanine metabolism</td>
<td>0.039</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Most significant expressed genes were subjected to functional enrichment analysis using GeneSCF by incorporating the Gene Ontology and KEGG databases, p-value <0.05. (Right-sided location n=23 and left-sided location n=17)*

We also analyzed the differences in gene expression between late (n=15) versus early cancer stages (n=25) (Figure 26). The top genes upregulated in the late cancer stage included members of the keratin gene family (KRT5, KRT13, KRT17, and KRT81). Additionally, MUC1, a marker for metastatic breast cancer diagnosis, and MUC15, which inhibits EGFR and PI3K-AKT signaling, were upregulated. Furthermore, PTGS2 antisense NFKB1 complex-mediated
expression regulator (PACERR), which interacts with NF-kB transcriptional regulators, was also upregulated. (Supplementary Table 3).

The downregulated genes included several associated with immune response, such as C-X-C motif chemokine receptor 1 and 2 (CXCR1 and CXCR2), colony stimulating factor 3 receptor (CSF3R), interleukin 1 alpha (IL1A), CD300e molecule (CD300E), triggering receptor expressed on myeloid cells 1 (TREM1), and serum amyloid A2 (SAA2), a major acute phase protein. Additionally, genes related to antimicrobial and cytotoxic peptides, including defensin alpha 6 (DEFA6), Fc fragment of IgG receptor (FCGR3B), and APOBEC3A, which restricts viral transmission, were also downregulated (Supplementary Table 4).

Figure 26. Volcano plots for colon cancer late versus early cancer stages for RNA sequencing data. Statistical comparison of gene expression levels by DESeq2, selection of differentially
expressed genes based on an absolute log2 fold change >1 and a p-value >0.01, late cancer stage n=15 and early cancer stages n=25.

Top gene upregulated pathways in early cancer stage were IL8R activity, TNFA signaling via NF-kB, and Phospholipase D signaling pathway. Top gene upregulated pathways in late cancer stage were tyrosine metabolism, KRAS signaling, Wnt/B-Catenin signaling, and fatty acid degradation (Table 6).

**Table 6.** The pathways associated with early and late colon cancer stages by RNA sequencing.

<table>
<thead>
<tr>
<th>Early cancer stages</th>
<th>p-value</th>
<th>Late cancer stage</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL8R activity</td>
<td>0.0001</td>
<td>Tyrosine metabolism</td>
<td>0.0002</td>
</tr>
<tr>
<td>TNFA signaling via NF-kB</td>
<td>0.0006</td>
<td>KRAS signaling</td>
<td>0.003</td>
</tr>
<tr>
<td>Phospholipase D signaling pathway</td>
<td>0.001</td>
<td>Wnt/B-Catenin signaling</td>
<td>0.008</td>
</tr>
<tr>
<td>Transmembrane signaling receptor activity</td>
<td>0.001</td>
<td>Fatty acid degradation</td>
<td>0.008</td>
</tr>
<tr>
<td>GPCR downstream signaling</td>
<td>0.005</td>
<td>Pyruvate metabolism</td>
<td>0.010</td>
</tr>
<tr>
<td>C-C chemokine binding</td>
<td>0.007</td>
<td>Basal cell carcinoma</td>
<td>0.016</td>
</tr>
<tr>
<td>Cell recruitment (pro-inflammatory response)</td>
<td>0.008</td>
<td>Glycolysis/Gluconeogenesis</td>
<td>0.018</td>
</tr>
<tr>
<td>Indoleamine 2,3-dioxygenase activity</td>
<td>0.015</td>
<td>Retinol metabolism</td>
<td>0.019</td>
</tr>
<tr>
<td>Myogenesis</td>
<td>0.020</td>
<td>Salivary secretion</td>
<td>0.033</td>
</tr>
<tr>
<td>IFN-gamma response</td>
<td>0.020</td>
<td>Inflammatory mediator regulation of TRP channels</td>
<td>0.037</td>
</tr>
<tr>
<td>Osteoclast differentiation</td>
<td>0.029</td>
<td>Protein digestion and absorption</td>
<td>0.040</td>
</tr>
<tr>
<td>Natural killer cell mediated cytotoxicity</td>
<td>0.031</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Most significant expressed genes were subjected to functional enrichment analysis using GeneSCF by incorporating the Gene Ontology and KEGG databases, p-value <0.05, late cancer stage n=15 and early cancer stages n=25.

Lastly, we analyzed the differences in gene expression between the presence or absence of *F. nucleatum* DNA in tumors previously determined by qPCR (*F.nucleatum*-positive n=20 and *F.nucleatum*-negative n=20) (Figure 27). The top upregulated genes in *F. nucleatum*-positive tumors were MUC6 a member of mucin protein, ONERCUT1 a member of the Cut homeobox family and keratin 16 (KRT16). Also, Wnt family member 7B (WNT7B), RasGEF domain family
member 1A (RGS7) involved in positive regulation of Ras protein and in cell migration, BCL2 family apoptosis regulator (BOK), and G0/G1 switch 2 (G0S2) a gene involved in extrinsic apoptosis signalling pathway. Additionally, genes associated with immune response such pro-platelet basic protein (PPBP) a potent chemoattractant and activator of neutrophils, mesothelin (MSLN), cytokine that can stimulate colony formation of bone marrow megakaryocytes, colony stimulating factor 3 receptor (CSF3R) a cytokine that controls the production, differentiation, and function of granulocytes, and pyrin innate immunity regulator (MEFV) were upregulated. And other genes associated to metabolism such as PFKFB3 and APOE, to insulin such as RGS7 and ECHDC3, and to members of superfamily of channels such as SYT12, CHRFAM7A, and KCNJ5 (Supplementary Table 5). The top downregulated genes were bestrophin 4 (BEST4), calneuron 1 (CALN1), X-[rolyl aminopeptidase 2 (XPNPEP2), phytanoyl-CoA 2-hydroxylase interacting protein like (PHYHIPL), flavin containing monooxygenase (5FMO5), XK related 7 (XRR7), and others (Supplementary Table 6).
Figure 27. Volcano plots for *Fusobacterium nucleatum* positive and negative presence in tumors for RNA sequencing data. Statistical comparison of gene expression levels by DESeq2, selection of differentially expressed genes based on an absolute log2 fold change >1 and a p-value >0.01, *F.nucleatum*-positive n=20 and *F.nucleatum*-negative n=20.

The top upregulated pathways in *F. nucleatum* positive tumors were neutrophil chemotaxis, viral protein interaction with cytokine-cytokine receptor, chemokine signaling pathway, cytokine-cytokine receptor interaction, TGF-β signaling, T cell differentiation, and osteoclast differentiation, NF-κB signaling, Wnt signaling pathways, and Toll-like receptor signaling pathway. The top upregulated pathways in *F. nucleatum* negative tumors were bile secretion, phototransduction, ascorbate and aldarate metabolism, pentose and glucuronate interconversions, and protein digestion and absorption (Table 7).
Table 7. The pathways associated with *F. nucleatum* positive and negative by RNA sequencing.

<table>
<thead>
<tr>
<th>Pathway</th>
<th><em>F. nucleatum</em> positive</th>
<th>p-value</th>
<th><em>F. nucleatum</em> negative</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutrophil chemotaxis</td>
<td>0.0002</td>
<td>Bile secretion</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Viral protein interaction with cytokine-cytokine receptor</td>
<td>0.0007</td>
<td>Phototransduction</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>Chemokine signaling pathway</td>
<td>0.0009</td>
<td>Ascorbate and aldarate metabolism</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>Cytokine-cytokine receptor interaction</td>
<td>0.001</td>
<td>Pentose and glucuronate interconversions</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>TGF-β signaling</td>
<td>0.002</td>
<td>Protein digestion and absorption</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>T cell differentiation</td>
<td>0.002</td>
<td>Porphyrin metabolism</td>
<td>0.018</td>
<td></td>
</tr>
<tr>
<td>Osteoclast differentiation</td>
<td>0.002</td>
<td>Regulation of lipolysis in adipocytes</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td>Epithelial Mesenchymal transition</td>
<td>0.003</td>
<td>Phenylalanine, tyrosine and tryptophan biosynthesis</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>Aldosterone synthesis and secretion</td>
<td>0.004</td>
<td>D-Amino acid metabolism</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>Inflammatory response</td>
<td>0.004</td>
<td>Mineral absorption</td>
<td>0.032</td>
<td></td>
</tr>
<tr>
<td>Wnt signaling pathways</td>
<td>0.015</td>
<td>Steroid hormone biosynthesis</td>
<td>0.033</td>
<td></td>
</tr>
<tr>
<td>MAPK family signaling cascades</td>
<td>0.019</td>
<td>Retinol metabolism</td>
<td>0.040</td>
<td></td>
</tr>
<tr>
<td>NF-κB</td>
<td>0.023</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toll-like receptor signaling pathway</td>
<td>0.023</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Most significant expressed genes were subjected to functional enrichment analysis using GeneSCF by incorporating the Gene Ontology and KEGG databases, p-value <0.05. Statistical comparison of gene expression levels by DESeq2, selection of differentially expressed genes based on an absolute log2 fold change >1 and a p-value >0.01, *F. nucleatum*-positive n=20 and *F. nucleatum*-negative n=20.

2b) To analyze serological levels of inflammatory mediators in colon cancer patients

20 proteins associated with human periodontal disease such as pro- and anti-inflammatory cytokines (IL-1α, IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, IL-12, IL-17, IFN-γ, TNF-α, TGF-β1); matrix metalloproteinases (MMP-9, and MMP-13) and their inhibitors (MIP-1α); bone metabolism related cytokines (osteoprotegerin (OPG), osteopontin (OPN), osteoactivin (OA), and receptor activator of nuclear factor kappa-B (RANK)); and acute phase protein (C-reactive protein) were
analyzed in serum of 47 colon cancer patients (23 right-sided and 24 left-sided tumor location; 28 early I-II and 19 late cancer stage III) using a quantitative slide-based array. This array has been categorized as a slide-based system that enables accurate quantification of multiple cytokines in biological samples simultaneously, using ELISA immunochemistry and fluorescence detection. This platform is more sensitive and has a greater dynamic range as compared to a conventional ELISA saving time and cost of assay. This platform provides a versatile system to rapidly quantify a wide variety of proteins in a multiplex format (Figure 28) (117,124).

**Figure 28.** Slide based multiplex arrays. **A.** Robotic printing of specific antibodies into the activated glass slide surface. Each antibody plus positive controls are arrayed in quadruplicate. **B.** The array functions like ELISA wells. After appropriate blocking, sample exposure, and washing, the detection of each target analyte is done by adding a combination of specific second antibodies to form a sandwich with a reporter system. **C.** High-resolution scanning allows for measuring signal intensity for each feature, indicating individual detection analyses. (Modified from Tighe et al. (124))

The differences in the serum levels of periodontal-related proteins between tumor location and cancer stage in patients were evaluated by Mann-Whitney test. The analysis was only performed on groups with a sufficient level of detection and enough samples number per group,
presented in Supplementary Table 7 and 8 in the Appendix section. Notably, no significant differences were detected in the mediators between tumor location and cancer stage. The clinicopathological characteristics of the patients with colon cancer are summarized in Figure 29. The analysis revealed elevated levels of MMP-9, OPG, OPN, osteoactivin, and CPR across the entire cohort of colon cancer patients. Figure 30 summarizes the functionality and the role of these mediators in CRC.

Figure 29. Heatmap of serological levels of inflammatory mediators of each colon cancer patient including clinicopathological variables and absence/presence of oral pathogens (qPCR assay). High/low levels were defined by cutoff (median values for each marker). Within the highlighted red box are the mediators exhibiting levels surpassing the limit of detection for the entire cohort.
of patients. (23 right-sided and 24 left-sided tumor location; 28 early I-II and 19 late cancer stage III).

**Figure 30.** Summary of functionality and role in CRC of CRP, MMP-9, MMP-13, Osteopontin, and Osteoactivin.

The correlation of serological levels of cytokines and the amount of *F. nucleatum* DNA in colon tumors was analyzed by Spearman correlation. IL-12, TNF-α, OPN and *F. nucleatum* DNA amount was observed to have a significant positive correlation. A significant positive correlation was found between abundance of *F. nucleatum* and TNF-α (*r* = 0.355, *p* = 0.014). The correlation for IL-12 (*r* = 0.300, *p* = 0.042) and OPN (*r* = 0.302, *p* = 0.039) was borderline significant.
Sub-aim 2c) To evaluate T-cell subset densities in colon cancer tumors

The abundance of T-cells infiltrating tumor tissues has been associated with improved clinical outcomes of CRC patients (30). However, the precise mechanism is unknown, though the adaptive immune system may be playing an important role in suppressing tumor progression. Tumor-infiltrating T-cells may be an indicator of the host immune response to the tumor and a target for immunotherapy. CD8+ T cells and CD4+ T cells are key players in antitumor immunity, and their functions are influenced by gut microbes. CD4+ T cells activate CD8+ T cells, enhancing their effector functions, proliferation, and survival (9).

As demonstrated in Aim 1, AA colon cancer patients exhibit elevated levels of F. nucleatum in their colonic tumor tissues. Our study represented one of the first human studies investigating whether the amount of F. nucleatum DNA in tumor tissue correlates with specific T-cell subset densities in AA colon cancer patients. Understanding the interaction between oral pathogens and different immune cell types could offer valuable insights into designing therapeutic strategies aimed at improving T-cell infiltration and enhancing the efficacy of immunotherapy.

In our investigation, we explored the association between F. nucleatum and the density of CD3+, CD4+, and CD8+ T cell subsets within colon tumor regions. Using IHC staining, we quantified T cell densities separately in intraepithelial and stromal regions of colon tumor tissues. The analysis revealed an intriguing relationship between the densities of these T cell subsets and the presence of F. nucleatum DNA, previously determined by qPCR. We utilized QuPath software to precisely count positive nuclei within neoplastic areas, excluding non-neoplastic mucosa and necrotic regions (Figure 31). Among the samples analyzed (n=13), F. nucleatum DNA was detected in approximately 6 tumors (46.1% of the total of cases). Our correlation analysis showed an inverse relationship between the amount of F. nucleatum and certain T cell densities, specifically CD3+ cells in both stromal and intraepithelial areas (r=-0.585, p=0.019, and r=-0.489, p= 0.046, respectively), as well as CD4+ cells in stromal regions (r=-0.770, p= 0.001) (Table 8).
This indicates a potential influence of *F. nucleatum* on T cell infiltration within these tumor microenvironments.

**Figure 31.** T cells densities analysis in digitalized whole colon tissue sections using QuPath Software. A. Digital images were acquired for all slides. B. Selection of stromal areas (dark green) and intraepithelial areas (light green) in tumor sites. C. Example of CD8+ positive cells detection is represented as red circles in the respective areas of the tumor.

**Table 8.** Correlation between *F. nucleatum* DNA amount in colon tumor tissue and density of T-cell subsets in tumor intraepithelial and stromal areas.

<table>
<thead>
<tr>
<th>T cell subsets</th>
<th>Region</th>
<th>r</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD3+</td>
<td>Stroma</td>
<td>-0.585</td>
<td>-0.863 to -0.032</td>
<td>0.019,*</td>
</tr>
<tr>
<td></td>
<td>Intraepithelial</td>
<td>-0.489</td>
<td>-0.825 to 0.102</td>
<td>0.046,*</td>
</tr>
<tr>
<td>CD4+</td>
<td>Stroma</td>
<td>-0.770</td>
<td>-0.930 to -0.365</td>
<td>0.001, **</td>
</tr>
<tr>
<td></td>
<td>Intraepithelial</td>
<td>-0.358</td>
<td>-0.767 to 0.257</td>
<td>0.114, ns</td>
</tr>
<tr>
<td>CD8+</td>
<td>Stroma</td>
<td>0.191</td>
<td>-0.417 to 0.681</td>
<td>0.265, ns</td>
</tr>
<tr>
<td></td>
<td>Intraepithelial</td>
<td>0.465</td>
<td>-0.132 to 0.815</td>
<td>0.056, ns</td>
</tr>
</tbody>
</table>

Spearman correlation test, *p > 0.05, **p > 0.001, ns= not significant.

Further examination of the T cell density between *F. nucleatum*-negative (n=7) and positive (n=6) colon tumor tissues showed a significant increase (about 2.5-fold) in the mean value of CD8+ T cells in the intraepithelial regions of *F. nucleatum*-positive cases compared to *F. nucleatum*-negative cases (217.4 ± 135.7 and 85.2 ± 46.3, respectively, p=0.036) (Table 9). This finding, supported by images in Figure 32 demonstrating increased CD8+ positive T cells in *F. nucleatum*-positive cases.
**nucleatum**-positive intraepithelial regions, suggests a potential association between the bacterium’s presence and CD8+ T cell recruitment within the tumor.

**Table 9.** Comparison of mean values of T-cell subset densities (cell count/mm²) between *F. nucleatum*-negative and -positive colon tumor tissues.

<table>
<thead>
<tr>
<th>T cell subsets</th>
<th>Region</th>
<th>F. nucleatum</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>CD3+</td>
<td>Stroma</td>
<td>127.4 ± 95.3</td>
<td>372.4 ± 261.5</td>
</tr>
<tr>
<td></td>
<td>Intraepithelial</td>
<td>38.1 ± 26.3</td>
<td>135.3 ± 218.4</td>
</tr>
<tr>
<td>CD4+</td>
<td>Stroma</td>
<td>37.8 ± 23.2</td>
<td>173.9 ± 135.9</td>
</tr>
<tr>
<td></td>
<td>Intraepithelial</td>
<td>18.2 ± 12.9</td>
<td>47.7 ± 44.3</td>
</tr>
<tr>
<td>CD8+</td>
<td>Stroma</td>
<td>222.4 ± 162.9</td>
<td>146.4 ± 62.4</td>
</tr>
<tr>
<td></td>
<td>Intraepithelial</td>
<td>217.4 ± 135.7</td>
<td>85.2 ± 46.3</td>
</tr>
</tbody>
</table>

Mann-Whitney test, *p > 0.05, ns= not significant, *F. nucleatum*-negative n=7 and -positive n=6

As we previously mentioned in Aim 1, *F. nucleatum* infection might be associated with co-abundance of other oral pathogens. Moreover, we expanded our investigation to explore the correlation between another oral pathogen, *P. intermedia* and T cell subset densities within colon tumor regions. Of the analyzed cases, 8 cases (61%) were positive for *P. intermedia*. Our analysis revealed an inverse correlation between the density of CD3+ cells in stromal and intraepithelial areas and the amount of *P. intermedia* (r= -0.599, p=0.017, and r= -0.5482, p=0.028, respectively), similar to correlation found for *F. nucleatum*. Additionally, a directly proportionate correlation was observed between CD8+ densities in intraepithelial areas and the amount of *P. intermedia* (r= 0.6217, p= 0.014), suggesting a potential effect on increased CD8+ T cell recruitment (**Table 10**). These findings underscore potential interactions between these oral bacteria and the tumor microenvironment, influencing T cell subset densities within specific
regions of colon tumors. Further research is necessary to fully elucidate these associations and their implications in colon cancer progression and response to therapy.

Figure 32. Immunohistochemistry staining for CD3+, CD4+, and CD8+ T cells for representative *F. nucleatum*-positive and -negative tumors. Red arrows indicate specific T cell subsets present in each tissue.
Table 10. Correlation between *P. intermedia* DNA amount in tumor tissue and density of T-cell subsets in tumor intraepithelial and stromal areas.

<table>
<thead>
<tr>
<th>T cell subsets</th>
<th>Region</th>
<th>r</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD3+</td>
<td>Stroma</td>
<td>-0.5991</td>
<td>-0.8692 to -0.05359</td>
<td>0.017, *</td>
</tr>
<tr>
<td></td>
<td>Intraepithelia</td>
<td>-0.5482</td>
<td>-0.8494 to 0.02225</td>
<td>0.028, *</td>
</tr>
<tr>
<td>CD4+</td>
<td>Stroma</td>
<td>-0.2996</td>
<td>-0.7385 to 0.3177</td>
<td>0.159, ns</td>
</tr>
<tr>
<td></td>
<td>Intraepithelia</td>
<td>-0.2148</td>
<td>-0.6943 to 0.3969</td>
<td>0.239, ns</td>
</tr>
<tr>
<td>CD8+</td>
<td>Stroma</td>
<td>0.1696</td>
<td>-0.4357 to 0.6692</td>
<td>0.290, ns</td>
</tr>
<tr>
<td></td>
<td>Intraepithelia</td>
<td>0.6217</td>
<td>0.08944 to 0.8778</td>
<td>0.014, *</td>
</tr>
</tbody>
</table>

Spearman correlation test, *p > 0.05, ns= not significant.

**Discussion – Aim 2**

Right-sided tumors are associated with immunological and metabolic pathways

We performed bulk RNAseq to analyze the gene expression of 738 immune-oncology markers in the tumors of 40 AA colon cancer patients. In our study, the pathways that are significantly enriched in right-sided colon cancer tumors highlight aspects primarily associated with inflammation, immune responses, and metabolic processes. These include pathways related to TNFA signaling, inflammatory responses, cell adhesion, chemotaxis of neutrophils, and various biosynthetic pathways involving thiamine, cofactors, and folate. Interestingly, right-sided colon cancer also exhibits upregulated pathways associated with bone mineralization, glutamate import, and resistance to certain compounds like antifolates.

First, several genes from the Homeobox family (HOXC5, HOXC6, HOXC9, HOXC10, HOXC11) were upregulated in right-sided tumors, as well as genes associated with development such as engrailed homeobox 1 (EN1) and armadillo repeat containing 3 (ARMC3), which makes sense since the larger portion (the proximal two-thirds) of the transverse colon develops from the midgut during embryonic stages, while only the remaining one-third originates from the hindgut (125). Notably, there are differential gradients of various homeobox (HOX) and other genes
establishing craniocaudal polarization (125). A previous study found that the upregulated HOXC10 expression was found significantly higher in tumor tissues relative to normal tissues, and may serve as a biomarker for the diagnosis of CRC and as a potential therapeutic target (126). Additionally, HOXC6 was identified to be overexpressed in various cancer types and has been linked to poorer overall survival in right-sided colon cancer (127). Further investigation should be conducted to the HOXC cluster as biomarkers and its association with proximal colon cancer.

Additionally, we observed numerous upregulated genes in right-sided tumors involved in chemotaxis and immune cell activation, including those associated with macrophages, neutrophils, T cells, and dendritic cells (CXCL8, CXCL11, CXCL17, IL1A, LCN15, HCAR2, FPR2, PPBP). The observed increase in IL1A gene expression, interleukin-1-alpha gene, in right-sided colon cancer among AA patients might have implications. Previously in our laboratory, Dr. Marzia Spagnardi (128) analyzed gene expression differences by RNAseq between AA cell lines (CHTN-06 and SB-521) (129), in comparison to CA cell lines (HT-29 and HCT-116). IL1A gene expression was increased for both CHTN-06 and SB-521 (25.7- and 7.67-fold change, respectively). In colon cancer patients, after assessing the gene expression of cytokines belonging to this immunogenic pathway, a 3.3-fold increase in IL1A in AA patients when compared to CA patients was observed (63). Its increased expression in AA colon cancer cell lines may suggest a potential contribution to the aggressive behavior of these tumors. Investigating IL1A levels provides insights into the factors driving the malignant properties of colon cancer in AA individuals.

The chemokine CXCL17, secreted by gastrointestinal epithelial cells, plays a crucial role in recruiting immune cells like neutrophils, macrophages, dendritic cells (DCs), and myeloid-derived suppressor cells (MDSCs) to mucosal tissues. In colon cancer, elevated expression of CXCL17 is linked to unfavorable survival outcomes. We found increased expression of CXCL17
in right-sided tumors. Recent research led by Dr. Jenny Paredes in our laboratory (63) revealed distinct IL-17A expression patterns in AA colon cancer patients compared to CA patients. Together, these results suggest that AA patients may fail to promote the anti-tumor activity of the Th1 and Th2 subsets of T cells and secrete cytokines associated with the presence of Th17 and inflammation. Further investigations are warranted to elucidate the precise mechanisms underlying these observations and their implications for therapeutic interventions.

Additionally, CXCL17 contributes to maintaining the integrity of the mucosal epithelium, aiding in the clearance of pathogenic infections like *Staphylococcus aureus*, *Escherichia coli*, *Pseudomonas aeruginosa*, *Salmonella*, and *Candida albicans* (130). Moreover, CXCL17 is a novel factor regulating T cell homeostasis and function. In studies involving Cxcl17−/− mice (131), researchers observed elevated levels of CD4+ and CD8+ T cells in the lymph nodes and spleen compared to wild-type mice. Their reduced presence within the TME, however, can delay effective immune surveillance and tumor eradication, allowing the tumor to evade immune attack and progress. CXCL17 expression was also shown to be tightly co-regulated with VEGF expression contributing to angiogenesis (132). This dual role of CXCL17 highlights its significance in immune cell recruitment and its impact on tumor-related vascular formation, underscoring its potential implications in the complex landscape of tumorigenesis. Further investigations are necessary to fully comprehend the relationship between CXCL17 and the influence of oral pathogens on colon cancer, especially in the right-sided location of the colon.

In our study, mucin 6 (MUC6) was upregulated by 9-fold (p<0.0001) in right-sided tumors when compared to the left location. MUC6 mucin plays a crucial protective role in the stomach, pancreas, and duodenum under normal conditions, yet it shows irregular expression patterns during the advancement of certain GI cancers. *In vitro* studies conducted on pancreatic, CRC, and breast cancer cell lines suggest a potential role for MUC6 in restraining the invasion of tumor cells through the basement membrane (133). This mechanism could potentially elucidate why
patients with tumors that demonstrate high expression of MUC6 tend to have a more favorable prognosis (127). Further investigation should be carried to correlate MUC6 with survival and colon location.

At a biological level, patients with right-sided colon cancer often present at more advanced tumor stages, with larger tumor sizes, and frequently poorly differentiated tumors, along with increased instances of lymphatic vascular invasion compared to left-sided colon cancer patients (11). Evidence from systematic reviews and meta-analyses consistently links right-sided colon cancer with poorer prognosis and overall survival outcomes. Molecular and multi-omics analyses have highlighted distinctions between both locations, encompassing variances in miRNA regulation, transcriptional control, and the immune landscape (14,134). In-depth analysis indicates that the reduced ability of right-sided colon cancer to neutralize carcinogens potentially fosters a more genotoxic tumor environment, leading to a more aggressive phenotype compared to left-sided colon cancer. Moreover, emerging evidence points to altered metabolic processes, a key hallmark in colon cancer, being more pronounced in right location, further strengthening the association between metabolism and cancer aggressiveness (135).

Recent studies (43,44) have highlighted an intriguing correlation between periodontitis and an increased risk of CRC, particularly in the proximal or right side of the colon. This observation underscores the significance of oral health factors, including a history of periodontal disease and tooth loss, in potentially increasing the risk of developing precancerous lesions in the colon. Understanding the shared characteristics between right-sided colon cancers and periodontitis in affected populations, especially among AA patients since there is a predominance of right-sided tumors, holds promise for elucidating the underlying mechanisms and improving clinical outcomes. Molecular profiling and innovative diagnostic approaches offer avenues for personalized treatment and early detection strategies tailored to this demographic.
Our investigation also revealed that pathways associated with TNFA signaling via NF-kB, inflammatory responses, regulation of keratinocyte differentiation, bone mineralization, cell adhesion, and neutrophil chemotaxis are upregulated in right-sided colon tumor locations, mirroring key components implicated in the pathogenesis of periodontitis (64,136). Moreover, we observed an increased presence of periodontal pathogens, including *F. nucleatum* and *P. gingivalis*, in right-sided colon cancer specimens. This raises intriguing questions regarding the potential role of periodontal pathogens in triggering a pro-inflammatory response akin to that seen in periodontitis within the gut environment. Furthermore, studies have indicated that immune cells residing in the oral draining lymph nodes possess the capacity to migrate to various lymphoid organs, including the gut (137). Consequently, oral inflammation may foster the generation of T cells reactive to oral pathobionts. These activated T cells could potentially migrate from the oral mucosa to the intestine, where they might encounter specific microbes, thus triggering intestinal inflammation. Further exploration of the immune mechanisms underlying this oral-gut axis is essential to unravel the full extent of their interconnectedness.

**Left-sided tumors present a metabolic and signaling features compared to right-sided tumors immune profile**

The right colon presents distinct biological characteristics and molecular features compared to the left colon. Multiple studies have categorized CRC subtypes using gene expression clustering. An international CRC Subtyping Consortium applied various independent classification methods to 3,962 samples, revealing an association among these different schemes and leading to the identification of 4 consensus molecular subtypes (CMSs). CMS1, mostly found in right-sided colon cancers, exhibits enhanced immune infiltration signatures. Heterogeneity exists within right-sided colorectal cancers. While CMS1 is more prevalent among them, all 4 CMS categories are present within this subgroup of CRCs. Conversely, CMS2 tumors were mainly left-sided. CMS2 tumors displayed epithelial differentiation and strong upregulation of
WNT and MYC downstream targets, both of which have classically been implicated in CRC carcinogenesis. They also observed a strong negative correlation between CMS1 and CMS2 indicating that these pairs are more easily distinguishable (10). The study did not provide information regarding the race of the included patients so this is a limitation.

Our results showed that left-sided colon cancer appears to have upregulated pathways connected to metabolic and signaling functions more like CMS3 'metabolic' profile: such as ascorbate, aldarate, pentose, and glucuronate metabolism, bile secretion, and steroid hormone biosynthesis. Additionally, signaling pathways such as calcium-ion regulated exocytosis and serotonergic synapse were upregulated. Additionally, a notable association with calcium homeostasis, sensing, mobilization, and absorption, was identified. This interrelation potentially contributes to the distinct immune landscape observed in left-sided tumors, influencing their decreased potential for proliferation and metastasis when compared to the right location (135).

These differences could potentially necessitate alternative or supplementary screening methods that specifically target specific CRC locations. Identifying markers or innovative screening techniques specific to right-sided colon cancer could overcome the limitations of colonoscopy in this context. These markers might aid in identifying individuals at elevated risk for right-sided CRC, allowing tailored and more frequent screenings for this subset of patients. Additionally, a thorough analysis could unveil potential connections between genetic, metabolic, and immune differences in both locations and their possible association with specific bacteria associated with colon cancer locations.

**Bone metabolism and colon cancer**

Recent metabolic research has underscored the intricate relationship between immune cell function and bone metabolism. In a comprehensive review (132), authors investigated the
interplay between bone metabolism markers and their implications for CRC development. Notably, osteoprotegerin (OPG), a member of the tumor necrosis factor receptor family (TNFR), emerged as a pivotal player with a dual role in CRC. While OPG is recognized for safeguarding bone tissue from osteoclast-induced damage, it also exhibits paradoxical behavior in CRC. Acting as both an oncogene and an onco-suppressor, OPG can either promote cancer development or impede tumor growth. Additionally, osteopontin (OPN) has garnered attention for its association with tumor progression, angiogenesis, and metastasis in colon cancer. Elevated levels of OPN may indicate its involvement in fostering aggressive cancer phenotypes, potentially influencing the invasive and migratory capacities of cancer cells.

In our study, we found elevated levels of MMP-9, OPG, OPN, osteoactivin, and CPR across the entire cohort of colon cancer patients suggesting potential implications for the pathobiology of the disease. A study described the association of OPG with an aggressive phenotype and the formation of metastases in CRC patients through mRNA expression analysis of immunohistochemical samples. Moreover, overexpression of the OPG protein was an independent risk factor for CRC recurrence (139). Additionally, another study revealed a correlation between OPG levels and advanced stages of CRC. Serum OPG was identified as an independent predictor of survival for metastatic CRC patients (140).

OPN serves as a potential CRC tumor marker, with elevated gene and protein expression in colon cancer cell lines and CRC tissue samples. In tumor tissues, OPN mRNA and protein expression were markedly elevated compared to non-tumor tissues and levels correlated significantly with lymph node metastasis, lymphatic or venous invasion, and TNM stage. Patients with high OPN mRNA expression had notably shorter disease-free and overall survival rates than those with low expression. Furthermore, OPN mRNA expression status independently predicted CRC patient prognosis (141). Deciphering their diverse functionalities unveils promising prospects for tailored therapies targeting these precise molecular pathways, potentially transforming CRC treatment approaches (138).
Bone metabolism mediators play crucial roles in initiating inflammation, promoting tumoral cell transformation, inducing epithelial–mesenchymal transition, facilitating intravasation and extravasation, promoting mesenchymal–epithelial transition, and ultimately contributing to the formation of metastases (138). It is noteworthy that our investigation revealed the upregulation of pathways linked to bone metabolism, particularly in early-stage cancers, right-sided tumors, and those positive for *F. nucleatum*. This finding underscores the multifaceted involvement of bone-related pathways in distinct subsets of CRC, offering valuable insights into the molecular nuances of these subtypes. Thus, the link between periodontitis-associated bacteria in the colon, heightened bone mineralization pathways, and potential systemic effects via inflammatory responses warrants comprehensive exploration for a deeper understanding and potential clinical implications. Further investigations into these connections could elucidate the mechanisms underlying this association.

The top upregulated pathway in early cancer stage was IL8R

Examining gene expression variations among racial groups, notably in AA patients, holds vital importance in comprehending distinctive tumor biology and formulating precise treatments. We analyzed the gene expression in late versus early cancer stages. The early cancer stage exhibited a notable upregulation of genes closely associated with immune response modulation. Among these, C-X-C motif chemokine receptor 1 and 2 (CXCR1 and CXCR2), essential for the recruitment and activation of neutrophils and other immune cells. Interleukin 1 alpha (IL1A), known for its involvement in inflammatory responses Additionally, colony stimulating factor 3 receptor (CSF3R), responsible for regulating granulocyte production and function, was prominently upregulated and CD300e molecule (CD300E), which plays a role in immune regulation. Moreover, the upregulation of serum amyloid A2 (SAA2) and triggering receptor expressed on myeloid cells 1 (TREM1) further underscored the activation of immune pathways in the early stages of cancer development.
We also found that the top upregulated pathway in early cancer stage was IL8R. Increased levels of IL-8 and CXCR2 in the TME has been associated with CRC growth, progression, and recurrence in patients. IL-8 mainly induced by pro-inflammatory cytokines such as IL-1β and TNFα, recruit and activate neutrophils and granulocytes to the tumor site. IL-8 has an affinity for two types of receptors: CXCR1 and CXCR2. Ligation of IL-8 to CXCR1 leads to CRC cell proliferation and neutrophils chemotaxis (142). However, when it binds to CXCR2, it affects angiogenic activity, impacting the formation of new blood vessels. Similarly, CXCL1 plays a crucial role in creating a premetastatic niche by recruiting myeloid-derived suppressor cells that express CXCR2 (143).

Jin et al. (144) showed that by incorporating modifications to the IL-8 receptor, specifically CXCR1 or CXCR2, into chimeric antigen receptors (CARs), there is a notable enhancement observed in the migration and persistence of T cells within the tumor microenvironment. This enhancement leads to remarkable outcomes, including complete tumor regression and the establishment of enduring immunologic memory. These effects have been particularly prominent in pre-clinical models representing aggressive tumors such as glioblastoma, ovarian cancer, and pancreatic cancer (144). This therapeutic approach, specifically tailored to interact with IL-8 receptors, showcases promising potential for advancing the efficacy of immunotherapeutic strategies in the treatment of challenging malignancies such as colon cancer.

Jovov et al. (51) conducted a thorough examination of gene expression patterns in tumor and non-tumor tissues obtained from both AA and European American EA CRC patients. Their analysis revealed significant differences in the expression of 95 genes between these two racial groups, with three out of six affected pathways linked to inflammatory and immune responses. Notably, one of the genes upregulated in AA patients was SAA2, same as in our study. The SAA
Gene family has been implicated in inflammatory bowel disease, suggesting its potential as a therapeutic target for mitigating chronic inflammation and reducing the associated risk of developing colitis-associated cancer. This finding underscores the importance of racial/ethnic-specific molecular profiling in identifying novel targets for therapeutic intervention in CRC. Further research into the mechanistic role of SAA2 in CRC pathogenesis and its potential as a therapeutic target is warranted.

*F. nucleatum* presence in tumors is associated with proinflammatory tumor-promoting microenvironment

We observed several genes that were significantly upregulated in *F. nucleatum*-positive tumors associated to immune response such as pro-platelet basic protein (PPBP) a potent chemoattractant and activator of neutrophils, mesothelin (MSLN) that stimulate megakaryocytes, colony stimulating factor 3 receptor (CSF3R), and pyrin innate immunity regulator (MEFV).

Also, upregulated pathways in *F. nucleatum* positive tumors were NF-κB, Toll-like receptor signaling, and Wnt signaling pathways. We know that the mechanisms by which *F. nucleatum* promotes tumor progression include generating a proinflammatory tumor-promoting microenvironment (145) and accelerating proliferation of colon cancer cells through activation of Wnt/beta-catenin signaling (28,146) and through TLR4-activated signaling to NF-κB (26). The stimulation of TLR signaling triggers the NF-κB pathway, resulting in the production of numerous cytokines. IL-1, a pivotal cytokine, is a central product of this pathway. Signaling via surface IL-1 receptors and various TLRs shares a common route, converging through MyD88 to activate NF-κB. This activation initiates an inflammatory response, prompting the expression of IL-1 (147).

Lee et al. (148) investigated the correlation between *F. nucleatum* presence, pathway mutations, and patient prognosis in 246 CRC patients. Their methodology closely resembled ours,
utilizing quantitative qPCR to detect *F. nucleatum* in both tumor and adjacent normal tissue, employing the same primer sequences (nusG gene for *F. nucleatum* and SLCO2A1 gene as a reference). Additionally, they conducted an analysis of targeted sequencing data for 40 genes linked to 5 key CRC pathways. Consistent with our findings, they observed higher levels of *F. nucleatum* in tumor tissue compared to adjacent normal tissue (p<0.001). Furthermore, they identified an association between *F. nucleatum* presence in tumors and the TGF-β pathway (p = 0.020), same than our study (148).

To fully comprehend the relationship between oral pathogens and inflammation in the TME of AA patients, further investigation into locally secreted inflammatory mediators within the tumor is crucial. Additionally, exploring whether increased circulatory cytokine levels are linked to active oral diseases like periodontitis caused by oral pathogens in these patients is necessary. This deeper analysis can shed light on the interplay between oral health and tumor-related inflammation.

*F. nucleatum* presence in tumors is associated with T cells recruitment and differentiation

Our RNA-seq analysis revealed a significant upregulation in the T differentiation pathway within *F. nucleatum*-positive tumors, implying its potential involvement in the oncogenesis of colon cancer. Tumors often attract immune cells due to their ability to display antigens. This presentation makes them vulnerable to recognition and subsequent targeting by the adaptive immune responses orchestrated by T cells (8) CD8+ T cells, termed cytotoxic T cells, possess the capability to eliminate virus-infected cells and combat cancerous cells directly. On the other hand, CD4+ T cells, often referred to as helper T cells, perform a supportive function by stimulating memory B cells and cytotoxic T cells. This activation leads to an amplified immune reaction. The adaptive immune response guided by the T helper cell is determined by its subtype (e.g., Th1, Th2, Th17, regulatory T-cell), each characterized by the specific cytokines they
produce. The differentiation of naïve CD4+ T cells into T helper type 1 cells producing IFN-γ promotes CD8+ T cell-mediated adaptive immunity (8).

Given these connections, we pursued to explore and spatially identify T cells subsets (CD3, CD4, and CD8) within tumor stromal and intraepithelial regions. This is the first study to evaluate the potential link between the quantity of F. nucleatum DNA in tumor tissue and the densities of specific T-cell subsets in AA colon cancer patients. Our findings showed an inverse relationship between F. nucleatum and specific T cell densities, such as CD3+ (Pan T cells) and CD4+ (T helper cells) T cells in tumor stromal and intraepithelial areas. Interestingly, F. nucleatum-positive cases exhibited a significant increase in CD8+ T cell density within the tumor, hinting at a potential connection between F. nucleatum presence and CD8+ T cell recruitment.

Borowsky et al. (149) showed that in tumor stromal areas, the densities of CD3+ and CD4+ T cells exhibited an inverse correlation with F. nucleatum levels (p < 0.005), same than the findings in our study. Same results found by Mima et al. (97) where F. nucleatum was inversely associated with the density of CD3+ T cells. However, in both studies by Borowsky et al. (149) and Mima et al. (97) no significant association was found between F. nucleatum with the density of CD8+ T cells. The race of the patients included in these studies was not disclosed and rectal cancers were included in the analysis.

Golby et al. (150) analyzed the proportions of tumor-infiltrating T-cell subsets by IHC staining in the intraepithelial and stromal compartments, similar to the methodology of our study minus the association with F. nucleatum infection. CD4+ cells were found to be significantly concentrated in the stroma and CD8+ cells were significantly concentrated in the intraepithelial region, aligned to our findings. Furthermore, they assessed T-cell subset proliferation, identified through Ki67 co-expression, showing a significantly higher proliferation frequencies in the intraepithelial compartment compared to the stromal compartment across CD3+, CD4+, and
CD8+ subsets. This contrasted with the typically non-proliferative state observed in T cells within the normal gut's intraepithelial and lamina propria compartments. This could mean that tumor-infiltrating T cells in direct contact with the tumor (intraepithelial region) exhibited substantially higher proliferation rates than stromal T cells, particularly evident in the CD8+ subset. The proliferation frequency of CD8+ intraepithelial cells surpassed that of stromal CD8+ T cells and even exceeded the proliferation rates of CD8+ T cells in any area of normal mucosal lymphoid tissue. While the precise factors driving CD8+ T cell proliferation in CRC remain elusive, it is posited that the tumor cells themselves might play a role.

Since we found in our study that *F. nucleatum*-positive cases exhibited a significant increase in CD8+ T cell density within the tumor, we speculate that the bacterium presence could be one of the factors driving the proliferation of CD8+ T cells. Interestingly, Roelands et al. (91) found that the abundance of *Fusobacterium* or *F. nucleatum* in the tumor environment is associated with the presence of cytotoxic T cells and natural killer (NK) cells, along with an increase in myeloid markers and signaling molecules (such as CD68, TREM1, and IL8 signature). But despite these immune responses, *F. nucleatum* may not be associated with a favorable outcome due to their ability to inhibit the antitumor activity by T and NK cells through activation of inhibitory receptors like TIGIT and CEACAM1.

Comparably, Gur et al. (151) isolated TILs from tumor samples and observed that approximately half of the CD3+ cells were CD8+ (cytotoxic T cells), half were CD4+ (helper T cells), and most TILs expressed TIGIT. Intriguingly, the TILs were found dispersed throughout the tumor tissue rather than being localized to specific areas with the bacterial presence of *F. nucleatum*. The study demonstrated that *F. nucleatum* adhered to tumor cells and, within tumors containing *F. nucleatum*, hindered the cytotoxic activity of natural killer (NK) cells and the functional abilities of TILs. These inhibitory effects were linked to the interaction between a protein produced by *F. nucleatum* called Fap2 and the human TIGIT receptor.
In a previous study by Dr. Jenny Paredes, gene expression was analyzed in tumor and non-tumor colon tissue samples from AA and EA colon cancer patients (63). The data showed that the scores of most of immune cell populations varied between tumors from EA and AA colon cancer patients, and AA tumors had significantly higher scores of CD8+ T cells and NK cells (p = 0.04). AA tumors also had a significant increase in scores of exhausted CD8+ cells. This observation suggests that while we found that AA colon cancer tumors do have significantly higher numbers of CD8+ T cells, these cells might not be functioning as effectively against the cancer cells due to possible exhaustion.

Galon et al. (152) characterized the TILs in a large cohort of human CRCs by gene expression profiling and IHC staining. They found that tumors from patients without recurrence had higher immune cell densities of CD3, CD8, GZMB and CD4RO. Moreover, patients with low density of CD3+ cells had a very poor prognosis, similar to patients with concomitant distant metastasis (152). Moreover, AA patients showed increased CD4+ and B-cell responses along with a greater proportion of exhausted CD8+ T cells relative to the overall CD8+ T-cell population (63). T cell exhaustion is a state of dysfunction characterized by the progressive loss of effector functions and sustained expression of inhibitory receptors, such as PD-1, being less capable of recognizing and eliminating cancer cells effectively. AA patients have fewer clinically actionable alterations than patients from other ancestries and tend to have tumors less responsive to immunotherapy and targeted therapy. Unfortunately, a smaller portion AA patients displayed molecular profiles suitable for targeting with these therapeutic approaches (153). This emphasizes the potential for immune checkpoint inhibitors targeting this exhaustion pattern in CD8+ T cells as a treatment approach (154).

Our study a pilot effort to comprehensively explore the multifaceted landscape of colon cancer in AA patients, integrating factors such as the microbiome, inflammatory mediators, genetic expression, and immune response. While our findings offer valuable insights, several
limitations warrant consideration. Firstly, the sample size in our study, though informative, is relatively small. To strengthen the robustness of our conclusions, future investigations should involve larger cohorts, incorporate control groups, and undergo external validations to ensure the reproducibility and generalizability of our results. Secondly, the scarcity of blank slides for IHC analysis restricted the number of samples that could undergo comprehensive staining, imposing constraints on the number of specimens that could be effectively matched between different analyses. Additionally, we were constrained to analyzing only one slide per T cell subset staining, potentially increasing both intra- and inter-individual variability. Lastly, in our serological analysis of inflammatory mediators, several samples fell either below the level of detection or outside the parameters of the assay. Despite following the manufacturer's instructions for sample volume and concentration, determining the correct values for these targets remains challenging. Addressing this issue may require methodological refinements or alternative assay approaches.

For future directions, it is imperative to address the issue of incomplete follow-up data regarding cancer outcomes and therapy responses among a subset of patients. Efforts should be intensified to ensure comprehensive data collection, allowing for more robust survival and prognosis studies. Additionally, it is noteworthy that tumors in AA patients exhibit lower responsiveness to immunotherapy and targeted therapy compared to other demographics. Unfortunately, only a minority of AA patients display molecular profiles amenable to these therapeutic modalities. This underscores the urgent need to explore alternative treatment approaches, such as immune checkpoint inhibitors targeting the exhaustion pattern observed in CD8+ T cells. By focusing on novel therapeutic strategies tailored to the unique characteristics of AA patients' tumors, we can potentially improve treatment outcomes and prognosis in this population.

Fostering multidisciplinary collaboration among experts from diverse fields will be instrumental in advancing our understanding of colon cancer and the development of
immunotherapeutic interventions that target, for example, specific oral pathogens. In summary, the oral-gut-circulatory axis represents a promising frontier in colon cancer research, potentially reshaping our conception of this disease and opening innovative opportunities for prevention and treatment.

**Conclusions- Aim 2**

In the context of colon cancer, our study delves into the distinct molecular landscapes and pathways associated with different tumor locations and stages, shedding light on potential microbial influences and specific immune cell responses. Analyzing right-sided colon cancer tumors, we identified significantly upregulated pathways emphasizing inflammation, immune responses, and metabolic processes. Furthermore, the presence of *F. nucleatum* DNA in tumors correlated with pathways involving inflammation, chemotaxis, Wnt, NF-κB, and the TLR signaling pathway. *Fusobacterium nucleatum*-positive tumors displayed increased gene expression levels related to T cell differentiation and a higher density of CD8+ T cells. These findings underscore the complexity of colon cancer, urging further research to unravel the underlying mechanisms and implications for disease progression and therapeutic interventions.
Summary

Our study explored the complex roles of inflammation, immune responses, and the microbiome in the context of colon cancer disparities. In this comprehensive investigation, we utilized DNA and RNA sequencing of tumor and non-tumor tissues, combined with the analysis of inflammatory mediators and antibody responses to bacterial proteins in serum. Furthermore, immunohistochemistry staining was employed to assess the association of T cells subset within the tumor in relation to the presence of *F. nucleatum*. DNA Notably, our research marks the pioneering effort to explore the interplay between periodontal bacteria, inflammation, and colon cancer within the AA demographic.

Microbiome in AA colon cancer patients (see Figure 33 and 34)

- There are significant differences in microbial composition between tumor and non-tumor tissues in AA colon cancer patients.
- The microbial communities within tumor tissues were predominantly comprised of *Proteobacteria, Firmicutes, Actinobacteria,* and *Bacteroidetes*. Additionally, *Fusobacteriota* was exclusively present in the tumor group and notably absent in the non-tumor group.
- Right-sided tumors had reduced diversity and about half the observed species compared to left-sided tumors.
- Novel oral-origin bacteria identified with significant differential abundance in these population; *Treponema socranskii* and *Prevotella spp.* in left-sided and early cancer stage, *Leptotrichia spp.* in early stage, *Fusobacterium necrophorum* spp. in right-sided and late stage, and *Veillonella spp.* in non-tumors tissues.
Inflammatory mediators and antibody response in serum (see Figure 33 and 34)

- IgG responses against specific bacterial proteins varied among cancer stages and colon tumor locations.
- The top five proteins with the highest seroprevalence were TonB/HP1341 (68% for IgG and 20% for IgA response), HP-1564 (61% for IgG), GroEL (50% for IgG) from *H. pylori*, a transglycosylase SLT domain-containing protein (64% for IgG and 11% IgA response), and an amino acid ABC transporter substrate-binding protein (57% for IgG and 11% for IgA response) from *S. gallolyticus*.
- Antibody responses varied across cancer stages and colon locations, notably showing increased seropositivity for *P. gingivalis* and *F. nucleatum* in right-sided colon and early cancer stages. Weak correlations suggested potential links between tissue-based bacterial abundance and systemic immune response.
- Prospective studies employing specific antibody response targets could be pivotal in early diagnosis.
- Levels of MMP-9, osteoprotegerin, osteopontin, osteoactivin, and CPR were within the detectable range across the entire cohort of colon cancer patients.

Gene expression analysis in tumors (see Figure 33 and 34)

- Upregulated pathways in right-sided colon tumors significantly emphasize inflammation and immune responses, in contrast to metabolic processes and signaling pathways in left-sided location, showcasing distinct biological mechanisms unique to each location.
- Increased IL1A expression in right-sided colon cancer among AA patients may suggest a potential contribution to the aggressive behavior of these tumors.
- These differences could potentially serve as screening methods that specifically target the unique nature of right-sided colon cancer.
• The top upregulated pathway in early cancer stage was IL8R in addition to genes CXCR1 and CXCR2, which could hold promising potential for advancing efficacy testing of immunotherapeutic strategies.

**F. nucleatum and colon cancer in AA patients (see Figure 35 and 36)**

• *F. nucleatum* was significantly more abundant in tumors, especially in right-sided and late-stage cancers, impacting the microbial community by elevating oral pathogens and reducing beneficial bacteria.

• There were significant increases in the relative abundance of other oral bacteria in Fn-positive tumors including: *Leptotrichia* (168-fold), *Campylobacter* (78-fold), *Pseudomonas* (51.5-fold), *Treponema* (50-fold), *Porphyromonas* (10-fold), *Fusobacterium* (10-fold), *Diallister* (8.5-fold), and *Prevotella* (4-fold).

• *F. nucleatum* aggregation phenomenon mirrors its behavior within the oral cavity and may play a pivotal role in the context of the gut, influencing microbial interactions and potentially contributing to the pathogenesis of colon cancer.

• This intriguing association underscores the need for further exploration into the specific mechanisms through which *F. nucleatum* influences bacterial dynamics and its impact on colon cancer development.

• Increase in seropositivity for both *P. gingivalis* and *F. nucleatum* in the right-sided colon and early cancer stages, particularly for combined targets such as FimA + PgCD00783414 + PgCD00783523, FNCD00783099 + FNCD00783088 + FNCD00783084, and FNCD00783868 + FNCD00783100. This finding warrants further investigation and validation of these targets as potential serological biomarkers for colon cancer.

• There was a positive weak correlation found for *F. nucleatum* relative amount in colon tissue and IgA antibody response in serum for antigens FnCD00783087 and FnCD00783099, and IgG against FnCD00783081.
• The presence of *F. nucleatum* DNA in tumors is associated with inflammation, neutrophils chemotaxis, chemokines, Wnt, NF-κB, and TLR signaling pathways.

• *F. nucleatum*-positive tumors showed increased gene expression levels related to T cell differentiation and higher density of CD8+ T cells.

**Figure 33.** Summary of the microbial and gene expression profiles in non-tumor, early stage and late cancer tumors in AA colon cancer patients.
**Figure 34.** Summary of the microbial and gene expression and antibody response against CRC-associated bacteria in right- and left-sided colon tumors in AA colon cancer patients.

**Figure 35.** Summary of the interplay between *F. nucleatum* and cell mediators in AA tumors.
Figure 36. Schematic overview of the key findings for *F. nucleatum* and colon cancer in African American patients.
References


34. Mima K, Nishihara R, Rong Qian Z, Cao Y, Sukawa Y, Nowak JA, et al. Fusobacterium nucleatum in colorectal carcinoma tissue and patient prognosis were responsible for collection of tumour tissue, and acquisition of epidemiologic, clinical and tumour tissue


52. Royston KJ, Adedokun B, Olopade OI. Race, the microbiome and colorectal cancer.


81. Barderas R. Protein Microarrays for Disease Analysis.


166
Appendix

Supplementary Figure 1. Efficiency of the amplification reaction using the slope of the standard curve of qPCR for *F. nucleatum*, *P. gingivalis*, and *P. intermedia*. 
### Supplementary Table 1. Top upregulated genes for right- versus left-sided location by RNAseq.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene description</th>
<th>Summary</th>
<th>log2 Fold Change</th>
<th>pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>MUC6</td>
<td>mucin 6, oligomeric mucus/gel-forming member of the mucin protein family</td>
<td>8.93</td>
<td>9.3E-13</td>
<td></td>
</tr>
<tr>
<td>HOXC11</td>
<td>homeobox C11</td>
<td></td>
<td>7.11</td>
<td>4.3E-09</td>
</tr>
<tr>
<td>CALB1</td>
<td>calbindin 1</td>
<td></td>
<td>6.12</td>
<td>1.6E-07</td>
</tr>
<tr>
<td>HOXC10</td>
<td>homeobox C10</td>
<td>homeobox family of genes that play an important role in morphogenesis</td>
<td>5.54</td>
<td>3.7E-08</td>
</tr>
<tr>
<td>HABP2</td>
<td>hyaluronan binding protein 2 role in the coagulation and fibrinolysis systems</td>
<td>5.28</td>
<td>2.9E-05</td>
<td></td>
</tr>
<tr>
<td>EN1</td>
<td>engrailed homeobox 1 role during development in segmentation</td>
<td>5.26</td>
<td>3.3E-03</td>
<td></td>
</tr>
<tr>
<td>RP11-550H2.1</td>
<td>novel transcript</td>
<td></td>
<td>5.22</td>
<td>3.5E-05</td>
</tr>
<tr>
<td>SPRR1B</td>
<td>small proline rich protein 1B envelope protein of keratinocytes</td>
<td>4.91</td>
<td>3.2E-04</td>
<td></td>
</tr>
<tr>
<td>LCN15</td>
<td>lipocalin 15 anti-microbial peptide originally identified in neutrophil granules</td>
<td>4.78</td>
<td>1.9E-06</td>
<td></td>
</tr>
<tr>
<td>DMRTA2</td>
<td>DMRT like family A2 regulation of transcription by RNA polymerase II</td>
<td>4.44</td>
<td>8.6E-06</td>
<td></td>
</tr>
<tr>
<td>CTD-2527J2.15</td>
<td>novel transcript</td>
<td></td>
<td>4.31</td>
<td>9.5E-06</td>
</tr>
<tr>
<td>HOXC6</td>
<td>homeobox C6 homeobox family of genes that play an important role in morphogenesis</td>
<td>4.29</td>
<td>9.3E-11</td>
<td></td>
</tr>
<tr>
<td>ARMC3</td>
<td>armadillo repeat containing 3 development, cell adhesion and mobility, and tumor initiation and metastasis</td>
<td>4.03</td>
<td>4.3E-03</td>
<td></td>
</tr>
<tr>
<td>HOXC5</td>
<td>homeobox C5 homeobox family of genes that play an important role in morphogenesis</td>
<td>3.98</td>
<td>1.3E-04</td>
<td></td>
</tr>
<tr>
<td>AFAP1-AS1</td>
<td>AFAP1 antisense RNA 1 overexpressed in tumor cells and may promote cancer cell metastasis</td>
<td>3.97</td>
<td>4.6E-04</td>
<td></td>
</tr>
<tr>
<td>RP11-678G14.4</td>
<td>novel transcript</td>
<td></td>
<td>3.95</td>
<td>7.6E-03</td>
</tr>
<tr>
<td>HOXC8</td>
<td>homeobox C8 homeobox family of genes that play an important role in morphogenesis</td>
<td>3.77</td>
<td>1.8E-05</td>
<td></td>
</tr>
<tr>
<td>HOXC9</td>
<td>homeobox C9 homeobox family of genes that play an important role in morphogenesis</td>
<td>3.68</td>
<td>1.2E-05</td>
<td></td>
</tr>
<tr>
<td>ALPP</td>
<td>alkaline phosphatase, placental hydrolysis of phosphoric acid monoesters</td>
<td>3.60</td>
<td>3.5E-05</td>
<td></td>
</tr>
<tr>
<td>NPSR1</td>
<td>neuropeptide S receptor 1 receptor for neuropeptide S</td>
<td>3.43</td>
<td>3.4E-03</td>
<td></td>
</tr>
<tr>
<td>CITED1</td>
<td>Asp/Glu-rich C-terminal domain (CITED) family play a role in pigmentation of melanocytes</td>
<td>3.18</td>
<td>8.4E-07</td>
<td></td>
</tr>
<tr>
<td>MSLN</td>
<td>mesothelin cell adhesion protein</td>
<td>3.10</td>
<td>5.4E-06</td>
<td></td>
</tr>
<tr>
<td>HCAR2</td>
<td>hydroxycarboxylic acid receptor 2 positive regulation of neutrophil apoptotic process</td>
<td>2.88</td>
<td>7.8E-04</td>
<td></td>
</tr>
<tr>
<td>CXCL17</td>
<td>C-X-C motif chemokine ligand 17 mucosal chemokine that attracts immature dendritic cells</td>
<td>2.81</td>
<td>2.4E-03</td>
<td></td>
</tr>
<tr>
<td>TFAP2A</td>
<td>transcription factor AP-2 alpha</td>
<td>2.74</td>
<td>4.9E-07</td>
<td></td>
</tr>
<tr>
<td>CDH7</td>
<td>cadherin 7 type II classical cadherin</td>
<td>2.67</td>
<td>1.1E-03</td>
<td></td>
</tr>
<tr>
<td>IL1A</td>
<td>interleukin 1 alpha This cytokine is produced by monocytes and macrophages as a proprotein,</td>
<td>2.51</td>
<td>6.4E-04</td>
<td></td>
</tr>
<tr>
<td>TFF2</td>
<td>trefoil factor 2 protect the mucosa from insults, stabilize the mucus layer and healing of the epithelium</td>
<td>2.48</td>
<td>1.3E-03</td>
<td></td>
</tr>
<tr>
<td>CXCL11</td>
<td>C-X-C motif chemokine ligand 11 chemotactic for activated T cells</td>
<td>2.29</td>
<td>4.8E-04</td>
<td></td>
</tr>
<tr>
<td>CXCL8</td>
<td>C-X-C motif chemokine ligand 8 chemotactic factor by guiding the neutrophils to the site of infection</td>
<td>2.26</td>
<td>4.3E-03</td>
<td></td>
</tr>
<tr>
<td>FPR2</td>
<td>formyl peptide receptor 2 positive regulation of monocyte chemotaxis</td>
<td>2.25</td>
<td>2.9E-03</td>
<td></td>
</tr>
<tr>
<td>PPBP</td>
<td>pro-platelet basic protein potent chemoattractant and activator of neutrophils</td>
<td>2.24</td>
<td>6.5E-04</td>
<td></td>
</tr>
</tbody>
</table>

Statistical comparison of gene expression levels by DESeq2, selection of differentially expressed genes based on an absolute log2 fold change >1 and a p-value >0.01. (Right-sided location n=23 and left-sided location n=17)
**Supplementary Table 2.** Top downregulated genes for right- versus left-sided location by RNAseq.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene description</th>
<th>Summary</th>
<th>log2 Fold Change</th>
<th>pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLDN8</td>
<td>claudin 8</td>
<td>prevent sodium back-leakage in distal colon</td>
<td>-6.35</td>
<td>7.5E-06</td>
</tr>
<tr>
<td>PRAC1</td>
<td>PRAC1 small nuclear protein</td>
<td>expressed in prostate, rectum and distal colon</td>
<td>-6.06</td>
<td>0.00004</td>
</tr>
<tr>
<td>SCNN1G</td>
<td>sodium channel epithelial 1 gamma subunit</td>
<td>transport across epithelia in many organs</td>
<td>-4.63</td>
<td>0.0000002</td>
</tr>
<tr>
<td>KRTAP13-2</td>
<td>keratin associated protein 13-2</td>
<td>located in cytosol</td>
<td>-4.50</td>
<td>0.0004</td>
</tr>
<tr>
<td>RP11-384J4.2</td>
<td>novel transcript</td>
<td></td>
<td>-4.44</td>
<td>0.0002</td>
</tr>
<tr>
<td>GCG</td>
<td>glucagon</td>
<td>secreted from gut endocrine cells</td>
<td>-4.32</td>
<td>0.0008</td>
</tr>
<tr>
<td>PYY</td>
<td>peptide YY</td>
<td>member of the neuropeptide secreted by endocrine cells in the gut</td>
<td>-4.00</td>
<td>0.00004</td>
</tr>
<tr>
<td>ATP12A</td>
<td>ATPase H+/K+ transporting non-gastric alpha2 subunit</td>
<td>family of P-type cation transport ATPases</td>
<td>-3.99</td>
<td>0.0000001</td>
</tr>
<tr>
<td>RP11-386E5.1</td>
<td>novel transcript</td>
<td></td>
<td>-3.97</td>
<td>0.0024</td>
</tr>
<tr>
<td>RIPPLY2</td>
<td>rippy transcriptional repressor 2</td>
<td>novel family of proteins required for vertebrate somitogenesis</td>
<td>-3.94</td>
<td>0.0098</td>
</tr>
<tr>
<td>INSL5</td>
<td>insulin like 5</td>
<td>classical signature of the insulin superfamily</td>
<td>-3.94</td>
<td>0.0017</td>
</tr>
<tr>
<td>ZG16</td>
<td>zymogen granule protein 16</td>
<td>enable carbohydrate binding activity and peptidoglycan binding activity</td>
<td>-3.75</td>
<td>0.0005</td>
</tr>
<tr>
<td>RP11-13K12.1</td>
<td>uncharacterized LOC100507351</td>
<td></td>
<td>-3.59</td>
<td>0.0000002</td>
</tr>
<tr>
<td>ANKRD20A4</td>
<td>ankyrin repeat domain 20 family member A4</td>
<td></td>
<td>-3.39</td>
<td>0.0000</td>
</tr>
<tr>
<td>CA1</td>
<td>carbonic anhydrase 1</td>
<td>zinc metalloenzymes for hydration of carbon dioxide</td>
<td>-3.39</td>
<td>0.0026</td>
</tr>
<tr>
<td>RP11-348F1.3</td>
<td>novel transcript</td>
<td></td>
<td>-3.38</td>
<td>0.0020</td>
</tr>
<tr>
<td>LHFPL4</td>
<td>LHFPL tetraspan subfamily member 4</td>
<td>member of the lipoma HMGIC fusion partner</td>
<td>-3.35</td>
<td>0.0004</td>
</tr>
<tr>
<td>SLC30A8</td>
<td>solute carrier family 30 member 8</td>
<td>zinc efflux transporter</td>
<td>-3.27</td>
<td>0.0004</td>
</tr>
<tr>
<td>FEV</td>
<td>FEV, ETS transcription factor</td>
<td>ETS transcription factor family</td>
<td>-3.23</td>
<td>0.0032</td>
</tr>
<tr>
<td>CHGA</td>
<td>chromogranin A</td>
<td>chromogranin/secretogranin family of neuroendocrine secretory proteins</td>
<td>-3.19</td>
<td>0.0052</td>
</tr>
<tr>
<td>SCNN1B</td>
<td>sodium channel epithelial 1 beta subunit</td>
<td>sodium channels control fluid and electrolyte transport</td>
<td>-3.16</td>
<td>0.0001</td>
</tr>
<tr>
<td>BEST2</td>
<td>bestrophin 2</td>
<td>family of anion channels</td>
<td>-3.08</td>
<td>0.0029</td>
</tr>
<tr>
<td>RP11-60A8.1</td>
<td>novel transcript</td>
<td></td>
<td>-3.05</td>
<td>0.0002</td>
</tr>
<tr>
<td>CD177</td>
<td>CD177 molecule</td>
<td>cell surface glycoprotein that plays a role in neutrophil activation</td>
<td>-3.00</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Statistical comparison of gene expression levels by DESeq2, selection of differentially expressed genes based on an absolute log2 fold change >1 and a p-value >0.01. (Right-sided location n=23 and left-sided location n=17)
### Supplementary Table 3. Top upregulated genes for late versus early cancer stage by RNAseq.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene description</th>
<th>Summary</th>
<th>log2 Fold Change</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTD-2384A14.1</td>
<td>long intergenic non-protein coding RNA 2327</td>
<td></td>
<td>6.02</td>
<td>0.001248</td>
</tr>
<tr>
<td>C17orf78</td>
<td>chromosome 17 open reading frame 78</td>
<td></td>
<td>5.28</td>
<td>0.002322</td>
</tr>
<tr>
<td>DBH</td>
<td>dopamine beta-hydroxylase</td>
<td>catalyzes the conversion of dopamine to norepinephrine</td>
<td>5.22</td>
<td>0.000050</td>
</tr>
<tr>
<td>PRR9</td>
<td>proline rich 9</td>
<td></td>
<td>5.12</td>
<td>0.001268</td>
</tr>
<tr>
<td>MUC15</td>
<td>mucin 15, cell surface associated</td>
<td>inhibits dimerization of EGFR and PI3K-AKT signaling</td>
<td>5.01</td>
<td>0.007978</td>
</tr>
<tr>
<td>RP11-148E17.1</td>
<td>novel transcript</td>
<td></td>
<td>4.82</td>
<td>0.002043</td>
</tr>
<tr>
<td>FADS6</td>
<td>fatty acid desaturase 6</td>
<td>lipid metabolic process</td>
<td>3.79</td>
<td>0.000049</td>
</tr>
<tr>
<td>VWA5B1</td>
<td>von Willebrand factor A domain containing 5B1</td>
<td></td>
<td>3.74</td>
<td>0.000117</td>
</tr>
<tr>
<td>KRT13</td>
<td>keratin 13</td>
<td>member of the keratin gene family</td>
<td>3.27</td>
<td>0.000334</td>
</tr>
<tr>
<td>NXF3</td>
<td>nuclear RNA export factor 3</td>
<td>member of a family of nuclear RNA export factor genes</td>
<td>3.23</td>
<td>0.000341</td>
</tr>
<tr>
<td>PADI1</td>
<td>peptidyl arginine deiminase 1</td>
<td>member of the peptidyl arginine deiminase family of enzymes</td>
<td>3.06</td>
<td>0.000001</td>
</tr>
<tr>
<td>KRT17</td>
<td>keratin 17</td>
<td>member of the keratin gene family</td>
<td>3.01</td>
<td>0.001805</td>
</tr>
<tr>
<td>KRT5</td>
<td>keratin 5</td>
<td>member of the keratin gene family</td>
<td>2.94</td>
<td>0.000979</td>
</tr>
<tr>
<td>MUC1L1</td>
<td>mucin like 1</td>
<td>marker for the diagnosis of metastatic breast cancer</td>
<td>2.82</td>
<td>0.002954</td>
</tr>
<tr>
<td>L1CAM</td>
<td>immunoglubulin supergene family</td>
<td>important role in nervous system development, including neuronal migration and differentiation</td>
<td>2.58</td>
<td>0.003253</td>
</tr>
<tr>
<td>MYH7B</td>
<td>myosin heavy chain 7B</td>
<td>member of the motor-domain superfamily</td>
<td>2.53</td>
<td>0.000016</td>
</tr>
<tr>
<td>PACERR</td>
<td>PTGS2 antisense NFKB1 complex-mediated expression regulator</td>
<td>interacts with NF-kB transcriptional regulators to promote expression of PTGS2</td>
<td>2.42</td>
<td>0.002556</td>
</tr>
<tr>
<td>PSCA</td>
<td>prostate stem cell antigen</td>
<td>glycosylphosphatidylinositol-anchored cell membrane glycoprotein</td>
<td>2.40</td>
<td>0.000490</td>
</tr>
<tr>
<td>MMP13</td>
<td>matrix metallopeptidase 13</td>
<td>member of the peptidase M10 family of matrix metalloproteinases (MMPs)</td>
<td>2.29</td>
<td>0.001541</td>
</tr>
<tr>
<td>KRT81</td>
<td>keratin 81</td>
<td>member of the keratin gene family</td>
<td>2.29</td>
<td>0.003566</td>
</tr>
</tbody>
</table>

Statistical comparison of gene expression levels by DESeq2, selection of differentially expressed genes based on an absolute log2 fold change >1 and a p-value >0.01, late cancer stage n=15 and early cancer stages n=25.
### Supplementary Table 4. Top downregulated genes for late versus early cancer stage by RNAseq.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene description</th>
<th>Summary</th>
<th>log2 Fold Change</th>
<th>pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>EN1</td>
<td>engrayed homebox 1</td>
<td>role in controlling development</td>
<td>-6.38</td>
<td>0.00087</td>
</tr>
<tr>
<td>C11orf87</td>
<td>chromosome 11 open reading frame 87</td>
<td>integral component of membrane</td>
<td>-5.20</td>
<td>0.00137</td>
</tr>
<tr>
<td>EMX2</td>
<td>empty spiracles homeobox 2</td>
<td>homeobox-containing transcription factor</td>
<td>-4.87</td>
<td>0.00007</td>
</tr>
<tr>
<td>DGKK</td>
<td>diacylglycerol kinase kappa</td>
<td>enzyme that phosphorlates diacylglycerol</td>
<td>-3.81</td>
<td>0.00047</td>
</tr>
<tr>
<td>CTB-61M7.2</td>
<td>novel transcript</td>
<td></td>
<td>-3.35</td>
<td>0.00014</td>
</tr>
<tr>
<td>DEFA6</td>
<td>defensin alpha 6</td>
<td>family of antimicrobial and cytotoxic peptides</td>
<td>-3.19</td>
<td>0.00570</td>
</tr>
<tr>
<td>CXCR1</td>
<td>C-X-C motif chemokine receptor 1</td>
<td>receptor for interleukin 8</td>
<td>-3.03</td>
<td>0.00013</td>
</tr>
<tr>
<td>FCGR3B</td>
<td>Fc fragment of IgG receptor IIIb</td>
<td>low affinity receptor for the Fc region of gamma immunoglobulins (IgG)</td>
<td>-2.86</td>
<td>0.00000</td>
</tr>
<tr>
<td>FPR2</td>
<td>formyl peptide receptor 2</td>
<td>enables amyloid-beta binding activity</td>
<td>-2.83</td>
<td>0.00007</td>
</tr>
<tr>
<td>APOBEC3A</td>
<td>apolipoprotein B mRNA editing enzyme catalytic subunit 3A</td>
<td>role in immunity, by restricting transmission of foreign DNA such as viruses</td>
<td>-2.75</td>
<td>0.00002</td>
</tr>
<tr>
<td>KB-1507C5.4</td>
<td></td>
<td></td>
<td>-2.58</td>
<td>0.00002</td>
</tr>
<tr>
<td>G0S2</td>
<td>G0/G1 switch 2</td>
<td>extrinsic apoptotic signaling pathway</td>
<td>-2.55</td>
<td>0.00024</td>
</tr>
<tr>
<td>CSF3R</td>
<td>colony stimulating factor 3 receptor</td>
<td>member of the family of cytokine receptors</td>
<td>-2.53</td>
<td>0.00003</td>
</tr>
<tr>
<td>IL1A</td>
<td>interleukin 1 alpha</td>
<td>pleiotropic cytokine involved in various immune responses, inflammatory processes, and hematopoiesis</td>
<td>-2.50</td>
<td>0.00023</td>
</tr>
<tr>
<td>CD300E</td>
<td>CD300e molecule</td>
<td>glycoprotein family of cell surface proteins expressed on myeloid cells</td>
<td>-2.29</td>
<td>0.00099</td>
</tr>
<tr>
<td>SAA2</td>
<td>serum amyloid A2</td>
<td>major acute phase protein</td>
<td>-2.25</td>
<td>0.00500</td>
</tr>
<tr>
<td>CXCR2</td>
<td>C-X-C motif chemokine receptor 2</td>
<td>receptor for interleukin 8</td>
<td>-2.23</td>
<td>0.00006</td>
</tr>
<tr>
<td>GRM1</td>
<td>glutamate metabotropic receptor 1</td>
<td>activates both ionotrophic and metabotropic glutamate receptors</td>
<td>-2.10</td>
<td>0.00534</td>
</tr>
<tr>
<td>TREM1</td>
<td>triggering receptor expressed on myeloid cells 1</td>
<td>Ig superfamily that is expressed on myeloid cells</td>
<td>-2.09</td>
<td>0.00332</td>
</tr>
</tbody>
</table>

Statistical comparison of gene expression levels by DESeq2, selection of differentially expressed genes based on an absolute log2 fold change >1 and a p-value >0.01, late cancer stage n=15 and early cancer stages n=25
Supplementary Table 5. Top upregulated genes for *F. nucleatum*-positive versus -negative by RNAseq.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene description</th>
<th>Summary</th>
<th>log2 Fold Change</th>
<th>pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>MUC6</td>
<td>mucin 6, oligomeric mucus/gel-forming</td>
<td>member of the mucin protein family</td>
<td>8.08</td>
<td>1.38E-09</td>
</tr>
<tr>
<td>ONECUT1</td>
<td>one cut homeobox 1</td>
<td>member of the Cut homeobox family of transcription factors</td>
<td>5.14</td>
<td>2.79E-05</td>
</tr>
<tr>
<td>KRT16</td>
<td>keratin 16</td>
<td>member of the keratin gene family</td>
<td>4.40</td>
<td>1.80E-06</td>
</tr>
<tr>
<td>LA16c-380H5.3</td>
<td>novel protein</td>
<td></td>
<td>4.15</td>
<td>1.13E-07</td>
</tr>
<tr>
<td>ARHGAP40</td>
<td>Rho GTPase activating protein 40</td>
<td>enable GTPase activator activity</td>
<td>4.07</td>
<td>5.49E-05</td>
</tr>
<tr>
<td>STAR</td>
<td>steroidogenic acute regulatory protein</td>
<td>key role in the acute regulation of steroid hormone synthesis</td>
<td>3.91</td>
<td>9.13E-05</td>
</tr>
<tr>
<td>WNT7B</td>
<td>Wnt family member 7B</td>
<td>member of the WNT gene family</td>
<td>3.38</td>
<td>7.71E-06</td>
</tr>
<tr>
<td>FRMPD2</td>
<td>FERM and PDZ domain containing 2</td>
<td>establishment and maintenance of cell polarization</td>
<td>3.31</td>
<td>0.00027</td>
</tr>
<tr>
<td>RGS7</td>
<td>regulator of G protein signaling 7</td>
<td>Enables G-protein beta-subunit binding activity</td>
<td>3.17</td>
<td>0.00012</td>
</tr>
<tr>
<td>IGF1</td>
<td>IGF like family member 2</td>
<td>IGF1 belongs to the insulin-like growth factor</td>
<td>3.09</td>
<td>1.72E-06</td>
</tr>
<tr>
<td>KL7</td>
<td>kallikrein related peptidase 7</td>
<td>many kallikrein genes are biomarkers for cancer.</td>
<td>2.98</td>
<td>0.00042</td>
</tr>
<tr>
<td>G0S2</td>
<td>G0/G1 switch 2</td>
<td>Involved in extrinsic apoptotic signaling pathway</td>
<td>2.83</td>
<td>1.77E-05</td>
</tr>
<tr>
<td>MSLN</td>
<td>mesothelin</td>
<td>cytokine that can stimulate colony formation of bone marrow megakaryocytes</td>
<td>2.59</td>
<td>0.00037</td>
</tr>
<tr>
<td>CREG2</td>
<td>cellular repressor of E1A stimulated genes 2</td>
<td>active in extracellular space</td>
<td>2.58</td>
<td>2.03E-05</td>
</tr>
<tr>
<td>LINCO0941</td>
<td>long intergenic non-protein coding RNA 941</td>
<td></td>
<td>2.49</td>
<td>5.72E-05</td>
</tr>
<tr>
<td>PPBP</td>
<td>pro-platelet basic protein</td>
<td>potent chemoattractant and activator of neutrophils</td>
<td>2.47</td>
<td>0.00041</td>
</tr>
<tr>
<td>SYT12</td>
<td>synaptotagmin 12</td>
<td>mediate calcium-dependent regulation of membrane trafficking in synaptic transmission</td>
<td>2.44</td>
<td>1.06E-05</td>
</tr>
<tr>
<td>TNNT1</td>
<td>tropinin T1, slow skeletal type</td>
<td>regulatory complex located on the thin filament of the sarcomere</td>
<td>2.35</td>
<td>0.00036</td>
</tr>
<tr>
<td>CARN51</td>
<td>carnosine synthase 1</td>
<td>member of the ATP-grasp family of ATPases</td>
<td>2.33</td>
<td>0.00013</td>
</tr>
<tr>
<td>C2CD4B</td>
<td>C2 calcium dependent domain containing 4B</td>
<td>Involved in positive regulation of acute inflammatory response</td>
<td>2.32</td>
<td>2.21E-05</td>
</tr>
<tr>
<td>CHRFAM7A</td>
<td>CHRNA7 (exons 5-10) and FAM7A (exons A-E) fusion</td>
<td>members of a superfamily of ligand-gated ion channels</td>
<td>2.31</td>
<td>0.00036</td>
</tr>
<tr>
<td>CCL4L2</td>
<td>C-C motif chemokine ligand 4 like 2</td>
<td>one of several cytokine genes that are clustered on the q-arm of chromosome 17</td>
<td>2.25</td>
<td>0.00024</td>
</tr>
<tr>
<td>RASGEF1A</td>
<td>RasGEF domain family member 1A</td>
<td>Involved in cell migration and positive regulation of Ras protein signal transduction</td>
<td>2.25</td>
<td>0.00043</td>
</tr>
<tr>
<td>CSF3R</td>
<td>colony stimulating factor 3 receptor</td>
<td>cytokine that controls the production, differentiation, and function of granulocytes</td>
<td>2.15</td>
<td>0.00025</td>
</tr>
<tr>
<td>ECHDC3</td>
<td>enoyl-CoA hydratase domain containing 3</td>
<td>Involved in positive regulation of cellular response to insulin stimulus</td>
<td>2.13</td>
<td>0.00046</td>
</tr>
<tr>
<td>MGP</td>
<td>matrix Gla protein</td>
<td>vitamin K-dependent protein is secreted by chondrocytes and vascular smooth muscle cells</td>
<td>2.07</td>
<td>0.00027</td>
</tr>
<tr>
<td>ARHGEF4</td>
<td>Rho guanine nucleotide exchange factor 4</td>
<td>orm complex with G proteins and stimulate Rho-dependent signals</td>
<td>2.07</td>
<td>0.00028</td>
</tr>
<tr>
<td>KCNJ5</td>
<td>potassium voltage-gated channel subfamily J member 5</td>
<td>subunit of the potassium channel which is homotetrameric</td>
<td>2.04</td>
<td>0.00010</td>
</tr>
<tr>
<td>MEFV</td>
<td>pyrin innate immunity regulator</td>
<td>pyrin or marenostrin, that is an important modulator of innate immunity</td>
<td>2.02</td>
<td>0.00022</td>
</tr>
<tr>
<td>APOE</td>
<td>apolipoprotein E</td>
<td>essential for the normal catabolism of triglyceride-rich lipoprotein constituents</td>
<td>1.54</td>
<td>0.00044</td>
</tr>
<tr>
<td>PFKFB3</td>
<td>6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3</td>
<td>synthesis and degradation of fructose-2,6-biphosphate,</td>
<td>1.23</td>
<td>0.00013</td>
</tr>
<tr>
<td>SF2L1</td>
<td>stromal cell derived factor 2 like 1</td>
<td>Enables misfolded protein binding activity</td>
<td>1.04</td>
<td>0.00016</td>
</tr>
<tr>
<td>BOK</td>
<td>BCL2 family apoptosis regulator</td>
<td>anti- or proapoptotic regulators that are involved in a wide variety of cellular processes</td>
<td>1.04</td>
<td>2.83E-06</td>
</tr>
</tbody>
</table>

Statistical comparison of gene expression levels by DESeq2, selection of differentially expressed genes based on an absolute log2 fold change >1 and a p-value >0.01, *F.nucleatum*-positive n=20 and *F.nucleatum*-negative n=20.
### Supplementary Table 6. Top downregulated genes for *F. nucleatum*-positive versus -negative by RNAseq.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene description</th>
<th>Summary</th>
<th>log2 Fold Change</th>
<th>pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT83</td>
<td>cancer/testis antigen 83</td>
<td></td>
<td>-12.25</td>
<td>5.87E-06</td>
</tr>
<tr>
<td>XKR7</td>
<td>XK related 7</td>
<td>involved in apoptotic process involved in development</td>
<td>-3.69</td>
<td>0.00022</td>
</tr>
<tr>
<td>PHYHIPL</td>
<td>phytanoyl-CoA 2-hydroxylase interacting protein like</td>
<td></td>
<td>-3.46</td>
<td>0.00022</td>
</tr>
<tr>
<td>CALN1</td>
<td>calneuron 1</td>
<td>protein with high similarity to the calcium-binding proteins of the calmodulin family</td>
<td>-2.98</td>
<td>5.87E-05</td>
</tr>
<tr>
<td>BEST4</td>
<td>bestrophin 4</td>
<td>transmembrane proteins that contain a homologous region rich in aromatic residues</td>
<td>-2.86</td>
<td>9.43E-07</td>
</tr>
<tr>
<td>XPNPEP2</td>
<td>X-prolyl aminopeptidase 2</td>
<td>common to several collagen degradation products</td>
<td>-2.35</td>
<td>9.38E-05</td>
</tr>
<tr>
<td>RP11-286E11.1</td>
<td>uncharacterized LOC101929595</td>
<td></td>
<td>-1.93</td>
<td>0.00015</td>
</tr>
<tr>
<td>FMO5</td>
<td>flavin containing monoxygenase 5</td>
<td>member of the FMO family of proteins, best known for their roles in the detoxification of foreign chemicals</td>
<td>-1.53</td>
<td>0.00017</td>
</tr>
<tr>
<td>PLEKHG7</td>
<td>pleckstrin homology and RhoGEF domain containing G7</td>
<td></td>
<td>-1.24</td>
<td>0.00026</td>
</tr>
<tr>
<td>GNG12-AS1</td>
<td>GNG12 antisense RNA 1</td>
<td></td>
<td>-1.23</td>
<td>0.00014</td>
</tr>
<tr>
<td>RP11-44F21.5</td>
<td>long intergenic non-protein coding RNA 2562</td>
<td></td>
<td>-1.17</td>
<td>0.00046</td>
</tr>
<tr>
<td>RP6-201G10.2</td>
<td>novel transcript</td>
<td></td>
<td>-1.15</td>
<td>0.00011</td>
</tr>
<tr>
<td>RP1-193H18.2</td>
<td></td>
<td></td>
<td>-1.12</td>
<td>0.00026</td>
</tr>
<tr>
<td>PRR26</td>
<td>proline rich 26</td>
<td></td>
<td>-1.08</td>
<td>0.00026</td>
</tr>
<tr>
<td>RP11-10C24.2</td>
<td></td>
<td></td>
<td>-1.05</td>
<td>2.56E-05</td>
</tr>
</tbody>
</table>

Statistical comparison of gene expression levels by DESeq2, selection of differentially expressed genes based on an absolute log2 fold change >1 and a p-value >0.01, *F. nucleatum*-positive n=20 and *F. nucleatum*-negative n=20.
**Supplementary Table 7.** Serum levels of inflammatory mediators in the entire study cohort of colon cancer patients (n=47).

<table>
<thead>
<tr>
<th>Mediators</th>
<th>Median</th>
<th>CI 95%</th>
<th>LOD</th>
<th>Max</th>
<th>Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-2</td>
<td>0</td>
<td>0-0.3</td>
<td>3.9</td>
<td>8,000</td>
<td>1/47</td>
</tr>
<tr>
<td>IL-1α</td>
<td>0</td>
<td>0-4</td>
<td>5.4</td>
<td>4,000</td>
<td>4/47</td>
</tr>
<tr>
<td>IFNγ</td>
<td>0</td>
<td>0.7-7.5</td>
<td>17.1</td>
<td>20,000</td>
<td>5/47</td>
</tr>
<tr>
<td>IL-4</td>
<td>0</td>
<td>0.3-10.4</td>
<td>6.7</td>
<td>4,000</td>
<td>5/47</td>
</tr>
<tr>
<td>IL-17</td>
<td>0</td>
<td>0.6-116</td>
<td>29.1</td>
<td>20,000</td>
<td>5/47</td>
</tr>
<tr>
<td>TGF-β1</td>
<td>0</td>
<td>57.5-528.9</td>
<td>355.2</td>
<td>200,000</td>
<td>8/47</td>
</tr>
<tr>
<td>RANK (TNFRSF11A)</td>
<td>0</td>
<td>76.5-379.6</td>
<td>320.7</td>
<td>200,000</td>
<td>9/47</td>
</tr>
<tr>
<td>IL-12</td>
<td>0</td>
<td>0.3-3.3</td>
<td>1.0</td>
<td>4,000</td>
<td>14/47</td>
</tr>
<tr>
<td>MMP-13</td>
<td>0</td>
<td>43.7-415.4</td>
<td>21.8</td>
<td>40,000</td>
<td>14/47</td>
</tr>
<tr>
<td>IL-1β</td>
<td>0</td>
<td>5.1-40.4</td>
<td>5.6</td>
<td>4,000</td>
<td>15/47</td>
</tr>
<tr>
<td>IL-10</td>
<td>0</td>
<td>0.8-3.1</td>
<td>2.1</td>
<td>4,000</td>
<td>16/47</td>
</tr>
<tr>
<td>IL-8 (CXCL8)</td>
<td>0</td>
<td>0.4-2.6</td>
<td>0.6</td>
<td>800</td>
<td>17/47</td>
</tr>
<tr>
<td>TNF-α</td>
<td>96.9</td>
<td>83.3-174.9</td>
<td>80.1</td>
<td>4,000</td>
<td>24/47</td>
</tr>
<tr>
<td>MIP-1α (CCL3)</td>
<td>30.2</td>
<td>5-557.9</td>
<td>25.6</td>
<td>20,000</td>
<td>36/47</td>
</tr>
<tr>
<td>IL-6</td>
<td>16.4</td>
<td>17.2-28.6</td>
<td>6.9</td>
<td>4,000</td>
<td>39/47</td>
</tr>
<tr>
<td>CRP</td>
<td>20529.1</td>
<td>20138-22777</td>
<td>491.1</td>
<td>80,000</td>
<td>47/47</td>
</tr>
<tr>
<td>MMP-9</td>
<td>16014.6</td>
<td>15112-19808</td>
<td>15.6</td>
<td>20,000</td>
<td>47/47</td>
</tr>
<tr>
<td>Osteoprotegerin (TNFRSF11B)</td>
<td>920.3</td>
<td>988.3-1486</td>
<td>11.9</td>
<td>40,000</td>
<td>47/47</td>
</tr>
<tr>
<td>Osteopontin (SPP1)</td>
<td>60226.6</td>
<td>54392-82336</td>
<td>651.9</td>
<td>200,000</td>
<td>47/47</td>
</tr>
<tr>
<td>Osteoactivin (GPNMB)</td>
<td>4095.4</td>
<td>3547-4493</td>
<td>14.3</td>
<td>20,000</td>
<td>47/47</td>
</tr>
</tbody>
</table>
**Supplementary Table 8.** Serum levels of inflammatory mediators in colon cancer patients by right/left-sided location and early/late stage (above the limit of detection values).

<table>
<thead>
<tr>
<th>Mediators</th>
<th>Right n=23</th>
<th>Left n=24</th>
<th>p-value</th>
<th>Early n=28</th>
<th>Late n=19</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>CRP</td>
<td>20056.4</td>
<td>3504.8</td>
<td>22801.8</td>
<td>4979.0</td>
<td>0.051</td>
<td>20533</td>
</tr>
<tr>
<td>IFNγ</td>
<td>3.8</td>
<td>10.6</td>
<td>4.3</td>
<td>12.6</td>
<td>0.983</td>
<td>3.1</td>
</tr>
<tr>
<td>TNF-α</td>
<td>138.6</td>
<td>143.2</td>
<td>120.0</td>
<td>170.1</td>
<td>0.507</td>
<td>111.7</td>
</tr>
<tr>
<td>TGF-b1</td>
<td>283.6</td>
<td>708.6</td>
<td>302.4</td>
<td>899.3</td>
<td>0.863</td>
<td>323.7</td>
</tr>
<tr>
<td>IL-1α</td>
<td>1.9</td>
<td>6.2</td>
<td>2.1</td>
<td>7.2</td>
<td>&gt;0.999</td>
<td>1.8</td>
</tr>
<tr>
<td>IL-1β</td>
<td>23.3</td>
<td>57.0</td>
<td>22.2</td>
<td>64.2</td>
<td>0.621</td>
<td>18.2</td>
</tr>
<tr>
<td>IL-2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.9</td>
<td>&gt;0.999</td>
<td>0.2</td>
</tr>
<tr>
<td>IL-4</td>
<td>5.2</td>
<td>16.3</td>
<td>5.5</td>
<td>18.6</td>
<td>0.768</td>
<td>4.7</td>
</tr>
<tr>
<td>IL-6</td>
<td>25.7</td>
<td>18.3</td>
<td>20.2</td>
<td>20.2</td>
<td>0.150</td>
<td>24.3</td>
</tr>
<tr>
<td>IL-8</td>
<td>1.4</td>
<td>2.7</td>
<td>1.6</td>
<td>4.3</td>
<td>0.602</td>
<td>1.2</td>
</tr>
<tr>
<td>IL-10</td>
<td>2.2</td>
<td>3.7</td>
<td>1.8</td>
<td>4.2</td>
<td>0.275</td>
<td>1.5</td>
</tr>
<tr>
<td>IL-12</td>
<td>1.7</td>
<td>5.0</td>
<td>1.9</td>
<td>5.6</td>
<td>0.868</td>
<td>1.5</td>
</tr>
<tr>
<td>IL-17</td>
<td>6.3</td>
<td>17.3</td>
<td>5.9</td>
<td>20.1</td>
<td>0.768</td>
<td>6.2</td>
</tr>
<tr>
<td>MIP-1α</td>
<td>95.5</td>
<td>157.5</td>
<td>459.6</td>
<td>1296.7</td>
<td>0.755</td>
<td>80.8</td>
</tr>
<tr>
<td>MMP-9</td>
<td>18909.0</td>
<td>9294.9</td>
<td>16072.7</td>
<td>6413.5</td>
<td>0.347</td>
<td>16040</td>
</tr>
<tr>
<td>MMP-13</td>
<td>236.9</td>
<td>597.4</td>
<td>222.5</td>
<td>678.1</td>
<td>0.239</td>
<td>205.2</td>
</tr>
<tr>
<td>OPG</td>
<td>1307.3</td>
<td>863.7</td>
<td>1169.9</td>
<td>844.7</td>
<td>0.326</td>
<td>1244.6</td>
</tr>
<tr>
<td>OPN</td>
<td>76920.1</td>
<td>48674.9</td>
<td>60164.9</td>
<td>46035.9</td>
<td>0.326</td>
<td>69104</td>
</tr>
<tr>
<td>OA</td>
<td>3207.0</td>
<td>984.0</td>
<td>4225.2</td>
<td>2044.1</td>
<td>0.891</td>
<td>3826.2</td>
</tr>
<tr>
<td>RANK</td>
<td>267.8</td>
<td>501.7</td>
<td>190.0</td>
<td>537.5</td>
<td>0.370</td>
<td>140.1</td>
</tr>
</tbody>
</table>

Mann-Whitney test, *p > 0.05, ns= not significant. Matrix metalloproteinases (MMP-9, and MMP-13) and their inhibitors (MIP-1α); bone metabolism related cytokines osteoprotegerin (OPG), osteopontin (OPN), osteoactivin (OA), and receptor activator of nuclear factor kappa-B (RANK) and acute phase protein (C-reactive protein).