THE MICROBIAL ETIOLOGY OF COLORECTAL CANCER

Melissa C. Kordahi

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
University of Washington
2020

Reading Committee:
R. William DePaolo, Chair
Rosana Risques
Jeffrey McLean

Program Authorized to Offer Degree:
Pathology

University of Washington

ABSTRACT

Microbial Etiology of Colorectal Cancer
Melissa C. Kordahi

Chair of the Supervisory Committee:
R. William DePaolo
Department of Medicine

The microbiome of the Gastro-Intestinal tract is estimated at 100 trillion organisms which act in a symbiotic relationship with the surrounding tissue cells to maintain homeostasis. However, alterations in the gut microbiota caused by genetics or environmental factors can disturb this relationship and promote diseases such as colorectal cancer (CRC). CRC is the third most common form of cancer in both men and women and the second leading cause of cancer-related death worldwide, presenting a considerable disease burden. The intimate association between the microbiota and the cells of the colon sets the stage for a number of interactions that may contribute to carcinogenesis. While only a few specific commensal species may play a direct causal role in CRC, more general shifts in the composition may promote local inflammation through engagement of innate immune receptors encoded within the colonic tissue. Changes in gene expression within the microbiota may also be important by altering virulence factors and producing metabolites that may have detrimental effects on the tissue. In the first chapter of this Thesis, we explore the conceptual frameworks through which certain members of the microbiota are believed to cause CRC, the sensing of microbiota associated molecular patterns by innate immune receptors known as Toll-Like-receptors (TLRs) and the various strategies aimed at manipulating the microbiota and targeting the TLRs, in the hopes of developing new treatment approaches. In the second chapter of this body of Research, we focus on *Bacteroides fragilis*, a particular bacterial commensal species that has

been correlated to CRC development in human studies. Mouse models have also shown that *B. fragilis* is capable of remodeling the mucosal immune response and colonic bacterial community to promote oncogenic changes in the epithelium. In our study, we analyzed the mucosal microbiome of patients undergoing CRC screening and noted a high prevalence of *B. fragilis* in patients with early CRC lesions. We isolated *B. fragilis* from mucosal biopsy samples for deeper characterization and showed that they phenotypically differed according to their geographical location in the colon and the presence of pre-cancerous lesions in the microenvironment they were isolated from. The results we gathered explore the relationship between *B. fragilis* and early-stage CRC and provide biological framework for microbiota-based biomarkers and therapeutic targets. This study shows that the pre-cancerous mucosal environment alters the immunogenicity of a gut commensal. Environmental factors that influence the gut microbiota are then further explored in the context of individual species in Chapter 3 using Matrix-Assisted Laser Desorption Ionization Time of Flight Mass Spectrometry (MALDI TOF MS) technology. Another factor influencing the gut microbiota and, therefore the host immune response, is genetics. Development of the human immune system depends on various receptors capable of recognizing and responding to pathogens and commensals. As mentioned in Chapter 1, these receptors include toll-like receptors (TLRs) found on macrophages, dendritic cells, and intestinal epithelial cells. Using TLR6 knockout mice in chapter 4, our study aimed to understand how disruption in host recognition of the microbiota can exacerbate disease. Defective TLR6 signaling was shown to worsen the host susceptibility to inflammation associated colorectal cancer. Within the same study, analysis of the microbiota revealed a therapeutic potential by restoring microbial ecology. By better understanding how the gut microbiome influences the development of colorectal cancer, we can begin to think in terms of innovative therapies to approach the treatment of this disease that continues to be challenge for healthcare professionals and patients in the clinic.

“Hold fast to dreams
For if dreams die
Life is a broken-winged bird
That cannot fly”

Langston Hughes

“All disease begins in the gut”

Hippocrates, 2500 B.C

To my Family and Friends.

To Bertrand.

ACKNOWLEDGEMENTS

One of the most impactful childhood experiences I've had was reading my father's collection of the comic series "The Adventures of Tintin" by Hergé. I was mostly fascinated and amused by Professor Calculus, the half-deaf physicist who was behind the invention of many sophisticated devices used throughout the series, such as the Moon rocket or the invention of what turned out to be a much-coveted Ultra-sound weapon. Professor Calculus was the archetypical mad scientist and absent-minded professor, but also a genius who was an expert in many fields of science such as biology and chemistry on top of physics. I was also very intrigued by the tight relationship he had with his peers in various scientific fields and by the true dedication they all had for their respective fields, as well as by the mutual respect and admiration they held for each other's work. This paired with a natural curiosity, a passion for science, philosophy and logic, and taste for problem solving, sparked in me a desire to belong to the sphere of scientific academic research. A sphere that I believe essentially exists to serve a higher purpose. The purpose of enlightening and advancing humanity through visionary and innovative thinking and discoveries. As we are all in search for 'Belonging' in this world, I have personally struggled to define my identity in the early years of my adulthood due to a myriad of paradoxical influences stemming from my multicultural background. In that regard, the feeling of belonging to the world of scientific research is very warm and comforting for the individual that I am. That being said, my journey to graduate school was not a straightforward one. Growing up in Lebanon, many women in my immediate environment were not exactly expected to embrace the long and tedious path of an academic career. The pressure of finding a suitable husband was also present and unfortunately often became the priority. Not that an academic career and being married were incompatible, but it does take a lot of strength and will-power for any woman to build a successful scientific career in academia while still thriving to have a healthy work-life balance and a family. I also learned during my journey in graduate school, the paramount importance of having a flexible and supportive life partner and my experience would certainly not have been the same without the unconditional and perpetual support of my beloved Bertrand. So, after having completed a professional doctorate in Pharmacy at the Lebanese American University in Lebanon, I applied to a PhD program in the School of Pharmacy at the University of Southern California, in Los Angeles, CA, thinking that it was a logical extension of my professional path thus far. My experience at USC was an extremely enriching one on so many levels. The friendships I have acquired there amongst my peers in the PhD program are friendships that I will keep and cherish for life. Adriana Blachowicz, Xianhui Chen, Hsuan

Yao, Michelle Grau, are not only truly inspiring and talented Scientists but also amazing people in so many ways and I look forward to hearing about their many life achievements in the years to come. However, it turned out that USC was not the ultimate destination of my Journey in Graduate School as Dr. DePaolo's laboratory was offered a relocation at the University of Washington in Seattle, WA. The decision to leave LA where I was finally starting to feel "at home" was not an easy one. However, it was also a great decision because it allowed me to experience the amazing Research environment that is offered at the University of Washington and the truly supportive and inspiring mentors and friends I have acquired here over the last four years and for which I am forever grateful. On that note, I would like to thank the directors of the Molecular Medicine and Mechanisms of Diseases PhD program including Dr. William Mahoney, Dr. Nancy Maizels, Dr. Daniel Promislow, and Dr. Conrad Liles for allowing my transfer from USC to UW and helping me keep the momentum throughout this transition. A Huge Thanks to the members of my Thesis Committee including Dr. William Grady, Dr. Rosana Risques, Dr. Jeff Mc Lean, and Dr. Lynn Hajjar for their time and their mentorship on my committee, as well as their support throughout the PhD process. A special thanks to Dr. William Grady and Dr. Cynthia Ko for allowing me to collaborate with GiCaRes at UW Medicine, which made my main PhD project possible. Another special thanks to the amazing Research coordinators at GiCaRes: Jennie Huang and Wynn Burke whose help in recruiting patients for my study was absolutely essential. Thanks to the amazing people I have met at UW and in Seattle over the last few years for also making this place feel a bit more like home, especially my theatre crew, "the Seagulls".

Thanks to all the members of the DePaolo lab that I have had the pleasure of working with and learning from. Thanks to Andrew Johnson for the guidance and support. Thanks to volunteers Anika Parker, Vince Friedman and Paul McCleary for the assistance and enthusiasm for science. Thanks to Leandra Brettner and Arushi Verma for the great support in the DePaolo Lab. Special thanks to Denise Chac and Marion Avril for the continuous support and guidance through the PhD process. Finally, huge thanks to William DePaolo for accepting me as a graduate student and providing me the opportunity to become a researcher.

Thanks to my family and friends for their unconditional love and support. Thanks to Fouad for being the best friend one can wish for. Special thanks to my parents for being supportive of my choices and standing by me, always. Last, but not least, thanks to Bertrand for being my backbone and support through it all.

Table of Contents

LIST OF FIGURES AND TABLES

1. Introduction

The Influence of the Microbiota on the Etiology of Colorectal Cancer

By Kordahi MC, DePaolo RW

Adapted from Chapter 8 of the Book "Mechanisms Underlying Host-Microbiome Interactions in Pathophysiology of Human Diseases" 2018, pages167-193, Boston, MA, Springer US
https://doi.org/10.1007/978-1-4939-7534-1_8

Keywords: Colon cancer, microbes, microbiota, innate immunity, toll-like receptors

Colorectal cancer (CRC) is the third most common form of cancer in both men and women and the second leading cause of cancer-related death in the USA[1]. CRC can be divided into three subtypes: heritable, sporadic, and inflammation-associated[2]. Family studies have identified mutations in dominant genes such as the adenomatous polyposis coli gene (APC) and mismatch repair genes in only 10–15% of CRC[2]. These tumors typically develop at earlier ages than the sporadic and inflammation-associated CRC[2]. Epidemiological studies have shown that the most CRC cases are sporadic, arising from non-shared environmental factors, and rapidly increase in incidence beyond age 50[1]. A third subset of CRC, comprising less than 1% of all CRCs, is associated with chronic inflammatory diseases such as Crohn's disease and ulcerative colitis, the two main forms of inflammatory bowel disease[1,3].

The intestine is home to a large microbial ecosystem that provides protective, structural, and metabolic functions. Due to the proximity of the microbiota to the intestinal epithelium and underlying immune cells, tightly regulated communication must occur to prevent abnormal tissue responses that could lead to chronic inflammation and malignancy. Coordination of intestinal responses are initiated through the recognition of both microbial-associated and host cellular-associated ligands by innate immune receptors, such as the toll-like receptor (TLR) and NOD-like receptor (NLR) families. Thus, mucosal homeostasis in the healthy intestine depends largely on the interplay between commensal microbiota, host genetics, and the immune status of the mucosal tissue. Disruption or imbalances of these signals can lead to uncontrolled inflammation and changes within the microbiota, which play a significant role in tumorigenesis and tumor progression[4].

The etiological importance of the intestinal microbiome in CRC has been most clearly illustrated in models using germ-free (GF) mice, which develop less inflammation and fewer tumors than conventionally housed mice[5]. Moreover, studies using GF mice colonized with the microbiota from tumor-bearing mice showed a significant increase in tumorigenesis in the colon compared with GF animals colonized with a healthy gut microbiome, further suggesting that the gut microbiome contributes directly to tumorigenesis[6]. Alterations in the composition of the microbiota have also been identified in both mouse models of cancer and patients with CRC[7,8,9]. However, it is important to consider that the initiation and progression of colon cancer is likely not due to one unique bacterial species, as many members of the microbiota have been identified as contributors to colon cancer pathogenesis[10]. In that regard, each bacterial species may contribute to carcinogenesis by a distinct microbial signature that could include the production of metabolites and other by-products, stimulation of innate immunity, changes in location, and/or changes in bacterial gene expression.

1.1. MICROBIOTA ASSOCIATED WITH COLORECTAL CANCER:

The number of commensal bacteria in a normal healthy gut is equal to the number of our own eukaryotic cells, but there is an even more astonishing amount of genetic diversity that these bacteria contribute to our physiology. It is estimated that for every one of our genes, there are approximately 145 microbial genes. This roughly equals 3.3 million bacterial genes in the gut to the 23,000 in the human genome[11]. Therefore, although it is important to consider how commensal composition changes in a diseased state, we must also pay close attention to changes in the gene expression of the microbiota. Changes in microbial gene expression may be influenced by intrinsic factors such as polymorphisms in the host genome or the immune status of the mucosal tissue, and they may be influenced by extrinsic factors such as diet, infection, and exposure to xenobiotics. These factors may induce genetic programs in the commensal microbiota, that modulate virulence protein expression, metabolites, genotoxins, and/or carcinogenic molecules, leading to direct neoplastic effects[10]. They may also intensify neoplasia through the induction of local inflammation[4]. For example, bacterially produced toxins can lead to DNA damage, inhibition of apoptosis or induction of cellular proliferation. In addition to toxin production, bacteria also produce reactive oxygen species (ROS). ROS can directly promote the initiation and progression

of carcinogenesis by causing DNA damage or alter cellular signaling and activation pathways, leading to cell survival and proliferation signals[12].

1.2. MICROBIOTA SPECIES MOST COMMONLY CORRELATED WITH CRC PATHOGENESIS

Streptococcus gallolyticus: *Streptococcus gallolyticus* is a low-grade pathogen known for its involvement in bacteremia and endocarditis. It belongs to the Firmicutes family and is a frequent colonizer of the intestinal tract. Interestingly, 25–80% of patients with *S. gallolyticus* in the bloodstream were shown to have concomitant colon adenomas[13]. Further studies have shown that *S. gallolyticus* bacteremia is specifically associated with an aggressive form of polyp, the tubular villous adenoma, necessitating careful clinical screening in a certain subset of patients[14]. The etiological role of *S. gallolyticus* in CRC is thought to be mediated through specific virulence mechanisms involving adherence and induction of inflammatory factors. Structurally, *Streptococcus gallolyticus* possesses a pilus protein, encoded by the *pil1* locus, with a collagen-binding domain allowing it to attach to mucosal surfaces and translocate into systemic circulation. The *pil1* locus is further induced through metabolic processes or promoted by the dysbiosis of the microbiota. This microorganism has been shown to translocate efficiently through a para-cellular epithelial route and promote vasodilatation and capillary permeability, thereby promoting vascularization of neoplasms. *S. gallolyticus* also induces strong inflammatory signals such as cyclooxygenase-2 (COX2), interferon-γ (IFN-γ), interleukin-1β (IL-1β), which may lead to alterations in apoptosis and proliferation, formation of nitric oxide and free radicals that directly cause DNA damage, or production of angiogenic factors such as interleukin-8[15,16].

Enterococcus faecalis: Another microorganism belonging to the Firmicutes phylum is *Enterococcus faecalis*. *E. faecalis* has been linked to CRC pathogenesis because certain strains have the capacity to produce ROS. The high levels of ROS can damage DNA and create genomic instabilities; two events that can lead to transformation in the colon's epithelium. The involvement of certain *E. faecalis* strains in CRC pathogenesis has further been assessed in studies showing that the ROS produced by the microorganism were involved in distal colitis, DNA damage, and cancer in Germ-Free mice that were deficient in mounting an anti-inflammatory Interleukin-10 (IL-10) mediated response. On the other hand, *E. faecalis* strains that did not produce ROS induced

inflammation, but not tumorigenesis. In addition to ROS, an *E. faecalis* symbiont found in the oral cavity could induce mucosal macrophages to produce another chromosomal-breaking factor called 4-hydroxy-2-nonenal, a breakdown product of omega-6 polyunsaturated fatty acids. Despite the abundant experimental literature, human studies linking superoxide-producing *E. faecalis* strains to tumorigenesis are lacking[17].

Enterotoxigenic *Bacteroides fragilis:* Enterotoxigenic *Bacteroides fragilis* (ETBF) belongs to the Bacteroidetes phylum, and may be considered oncogenic under certain circumstances, because of its virulence factor *B. fragilis* toxin (BFT), or fragilysin. BFT, like ROS, can induce DNA damage *in vivo*[10]. Further, BFT has been shown to rapidly alter the structure and function of colonic epithelial cells, including the cleavage of the tumor suppressor protein, E-cadherin. E-cadherin is a transmembrane protein confined to epithelial cells and responsible for maintaining the tight junctions between neighboring cells. The extracellular domain of one E-cadherin molecule interacts with E-cadherin molecules on neighboring cells. A pre-requisite for intercellular adhesion is the cytoplasmic linkage of E-cadherin to β-catenin. The association of β-catenin at the cell membrane prevents its nuclear translocation and activation of oncogenic signals. The cleavage of E-cadherin by BFT increases cytosolic levels of β-catenin, allowing it to translocate to the nucleus and increase epithelial cell proliferation and expression of proto-oncogenes such as myelocytomatosis viral oncogene (MYC)[10]. In Apc(Min/+) mice, BFT has been shown to induce colonic hyperplasia and tumor initiation via induction of signal transducer and activator of transcription 3 (STAT3) and a TH17 inflammatory response[18]. Apc(Min/+) mice develop tumors in the small bowel with limited formation in the colon. However, colonization with ETBF increases tumorigenesis in the distal colon, but not in the small intestine and histological findings showed that colonic adenomas are detectable much faster in mice colonized with ETBF than in Apc(Min/+) mice that were ETBF-free. In humans, one study detected ETBF at a significantly higher frequency in the stools of consecutive cases of CRC compared with concurrent hospital-based, age- and gender-matched patients without CRC[19]. However, the development of an IL-17 immune response has been linked to a worse prognosis in human CRC, indicating that long-term ETBF colonization may promote colon carcinogenesis in certain predisposed individuals.

Escherichia coli: Unlike ETBF, and *E. faecalis*, whose importance in CRC has been identified using mouse models or in pre-clinical studies, *E. coli* has been isolated from human CRC patients

and the importance of this microorganism in the immunopathophysiology of CRC has been verified experimentally[20]. *E. coli* is a member of the Enterobacteriaceae family and constitutes less than 1% of a healthy individual's fecal microbiota, when performing 16S rDNA sequence analysis. Despite it being found in relatively low numbers compared with other commensal bacteria, *E. coli* is a very common cause of intestinal disease[21]. During inflammation, *E. coli* often becomes a dominant member in the gut microbiota after 16S fecal rDNA sequence analysis, a phenotype particularly associated with clinical irritable bowel disease (IBD) and, in animal models, chronic inflammation[22]. Although the molecular mechanism that *E. coli* uses to expand during tumorigenesis is not known, Enterobacteriaceae and other Proteobacteria have evolved several strategies to utilize products or by-products formed during an inflammatory response. *Escherichia coli* is particularly interesting because, in addition to the changes observed in its abundance, it seems that it can also alter its gene expression in an inflamed gut. Evidence for changes in gene expression come from the analysis of clinical isolates from patients with a chronic disease, such as IBD and CRC. These studies demonstrate that *E. coli* alters its functional characteristics by inducing a more pathogenic phenotype, including an increase in its adherence and invasive abilities[23]. Analysis of *E. coli* strains isolated from IBD and CRC patients has identified several genes that encode factors influencing tumorigenesis. The first is cytolethal distending toxin or CDT-V, which can directly cause DNA damage[24]. However, only a small number of *E. coli* strains carry this gene. More recently, a natural peptide–polyketide genotoxin called colibactin was identified in *E. coli* isolated from IBD and CRC patients as well[25]. This genotoxin is encoded by the 54-kb polyketide synthase (PKS) genotoxicity island. The importance of this genotoxicity island was first demonstrated by Jobin and colleagues when a mutant *E. coli* strain harboring a deletion in PKS was still able to induce inflammation, but had less DNA damage, tumor numbers, and bacterial invasion in mice lacking the gene encoding the anti-inflammatory cytokine interleukin-10 (IL-10)[26]. Thus, in the context of IL-10 deficiency, carcinogenesis requires expression of a bacterial genotoxin in addition to the genotype-dependent inflammation. PKS may promote oncogenesis via its ability to directly bind DNA and causes double-stranded breaks[27].

E. coli has been associated with gender-specific differences in CRC development as well. It was demonstrated that hemolytic type I *E. Coli* is significantly associated with adenoma and CRC in female patients only, upon the analysis of many clinical *E. coli* isolates. This was linked to the activation of the expression of the tumor suppressor *bim* by acting in part on hypoxia-induced α-subunit[28].

Fusobacterium nucleatum: *Fusobacterium Nucleatum* is a fastidious anaerobe, belonging to the Fusobacteria family, that has also been isolated from patients with CRC and its pro-tumorigenic effects have been verified in experimental models. Despite its presence in the colon, *F. nucleatum* is most abundant in the oral cavity where it is associated with dental plaques and gum disease. Although considered an oral bacterium, *F. nucleatum* has been intimately linked to gut inflammation and carcinogenesis, because of its isolation from IBD and CRC patients[29]. Like *E. coli*, there seem to be pro-tumorigenic effects due to inflammation caused by the expression of the microorganism's own genes. There is a strong correlation between the abundance of *F. nucleatum* and the magnitude of the inflammatory response, especially in terms of tumor necrotic factor-α (TNF-α) and IL-10 expression[30]. However, the molecular signals that induce these cytokines are not well understood. In fact, the findings to date demonstrate that the induction of pro-inflammatory cytokine responses are likely TLR- and NLR-independent yet rely on the sensing of *F. nucleatum* by viral-associated innate receptors such as cytoplasmic retinoic acid-inducible gene I[31]. The genetics of *F. nucleatum* are still widely undefined, as this microorganism is difficult to isolate, culture, and manipulate experimentally and clinically. Studies using a periodontal disease-derived *F. nucleatum* strain suggested that the invasive and carcinogenic properties of *F. nucleatum* might be mediated by the activated complex of the FadA adhesin (FadAc), a well-characterized virulence protein[32]. *In vitro* colon carcinoma cell-line studies and *in vivo* tumor xenograft models revealed that FadAc binds to a select extracellular domain of E-cadherin. These binding triggers invasion of the organism and activation of *β-catenin*/*Wnt* signaling with stimulation of cell proliferation or tumor growth. Evaluation of tumor tissue from adenoma and adenocarcinoma patients compared with normal colon tissue from non-tumorous individuals revealed elevated gene copy numbers of *fadA*, but more interesting was the fact that the highest *fadA* gene copies were detected in cancer tissues and associated with increases in expression of representative *Wnt* and nuclear factor kappa B (*NF-κB*) genes[32]. In another study, *F. nucleatum* was shown to be associated with the promotion of colonic tumor formation in mice through the identification of a host polysaccharide Gal-GalNAc and fusobacterial lectin (Fap2) that explained the abundance of the microorganism in CRC. Indeed, Fap2 was shown to mediate *F. nucleatum* binding to overexpressed Gal-GalNAc in CRC. Targeting host Gal-GalNAc or Fap2 may provide a way to reduce the *F. nucleatum* drive of CRC[33]. Anecdotally, *F. nucleatum* was also shown to be more abundant on colon tumors from Spanish individuals compared with tumors from

individuals in the USA or Vietnam. This suggests that *F. nucleatum* colonization might vary regionally, indicating that environmental causes may affect the virulence of this microorganism and its involvement in CRC[34].

These five CRC-associated members of the microbiota, *Streptococcus gallolyticus*, *Enterococcus faecalis*, Enterotoxigenic *Bacteroides fragilis*, *Escherichia coli*, and *Fusobacterium nucleatum* have also been suggested to breach the colonic mucus layer and persistently adhere to the mucosa. This mucosal adherence is likely necessary for their oncogenic potential because it allows a more intimate contact with the epithelium. This intimate association could affect the rate of initiation and progression of CRC by promoting inflammation via the stimulation of innate receptors. It could also allow direct targeting of factors that may cause DNA damage, or in the context of an individual's genetic makeup and environment, have an impact on cellular turnover[35,36]. Furthermore, alterations in the composition of the microbiome can favor changes in bacterial virulence and metabolism genes, supporting the hypothesis that specific commensal microbes might act sequentially, and in synergy, with certain microbial communities involved in colon carcinogenesis. The findings concerning these microorganisms demonstrate that there are three major mechanisms through which bacteria may contribute to human CRC pathogenesis. The first is that certain members of the colonic microbial community are capable of triggering signaling pathways classically activated during carcinogenesis, such as Wnt and E-cadherin. The second mechanism is through the induction of oxidative stress pathways leading to DNA damage. The last mechanism is through the activation of signaling pathways leading to an inflammatory cytokine response and the production of interleukin-17 (IL-17), IFN-γ, and TNF-α[10]. However, simplification of these bacterial processes in CRC does not rule out the possibility that additional signaling pathways may be involved, nor does it diminish the possibility that, yet undefined regulatory elements could be targeted by the microbiota. These microorganisms and their genomic potential to induce CRC are summarized in Fig. 1.1.

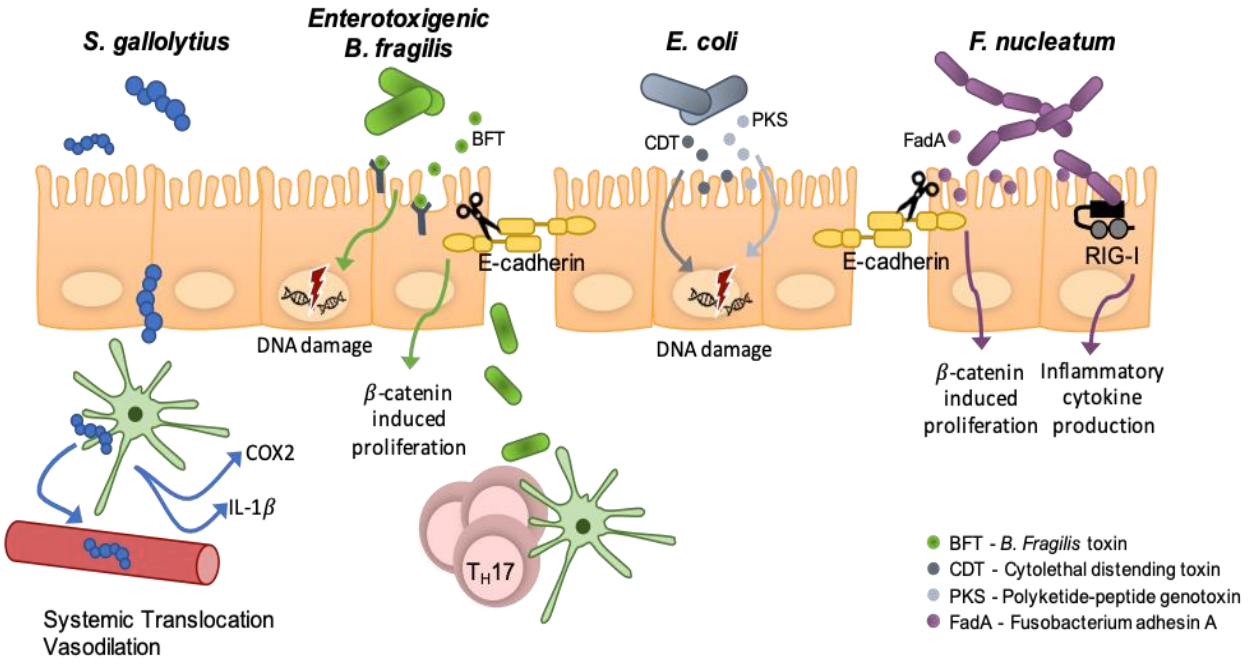


Figure 1.1 Common commensal bacteria associated with CRC

The commensal species most commonly associated with colorectal cancer (CRC) are *Streptococcus gallolyticus* (blue) enterotoxigenic *Bacteroides fragilis* (green), *Escherichia coli* (gray), *Fusobacterium nucleatum* (purple) and *Enterococcus faecalis* (not shown). *S. gallolyticus* promotes inflammation via its para-cellular translocation and subsequent stimulation of mucosal dendritic cells causing elevated Cox2 and inflammatory cytokines. *S. gallolyticus* also promotes vascularization and vasodilation. Enterotoxigenic *B. fragilis* (ETBF) produces fragilysin or BFT, which can directly cause DNA damage or increase proliferation by freeing intracellular β-catenin from E-cadherin. Chloride ions produced by ETBF and the dissociation of E-cadherin allows bacterial translocation and stimulation of TH17-mediated immunity. *E. coli* species are predominant during inflammation and in CRC alter gene expression to become more mucosally associated. *E. coli* expresses CDT and PKS, which both have DNA-damaging effects. *F. nucleatum* expresses a virulence factor, FadA, which binds to the extracellular portion of E-cadherin and dissociates β-catenin, allowing activation of proto-oncogenic and proliferative pathways. *F. nucleatum* also contributes to CRC via the induction inflammation after recognition by intracellular retinoic acid inducible gene I.

1.3.METABOLITES THAT CONTRIBUTE TO CRC

The Influence of Microbiota-Derived Metabolites on CRC

As discussed earlier, the microbiota contributes to the immunopathophysiology of CRC via inflammation and modulation of pathways, leading to carcinogenesis. In addition to these pro-tumorigenic effects of the microbiome, accumulating evidence suggests that there might be an influence of the wider microbial community on CRC through its secreted metabolites. Some of the metabolites secreted from the microbiota exert an important beneficial influence on human health, whereas others have been linked to the pathogenesis of cancer by influencing inflammation, DNA damage, and apoptosis, as summarized in Fig. 1.2.

Nitrogenous Compounds: A subset of both Bacteroidetes and some Firmicutes ferment aromatic amino acids from proteins and produce potentially bioactive ammonia and nitrogenous products, particularly N-nitroso compounds (NOCs). Both ammonia and NOCs are carcinogenic agents. In the case of NOCs, their carcinogenic nature is due to their ability to alkylate DNA, resulting in genetic mutations. NOCs are positively associated with CRC in Europeans and although some pre-formed NOCs are taken in as part of the diet, they can also be formed via endogenous microbial fermentation that occurs in the colon through the expression of nitro- and nitrate-reductases encoded by Proteobacteria[37,38].

Sulfides: Hydrogen sulfide is a major product of the gut and occurs through the reduction of diet-derived sulfate. Sulfides cause both a breakdown in the epithelial barrier and DNA damage via the activation of ROS. An increase in sulfate-reducing bacteria such as Desulfovibrio spp. are not likely driving these pathways in CRC, as they have not been increased in the stool of patients. Therefore, the increase in hydrogen sulfide may be due to changes in bacterial gene expression and activity rather than composition[12].

Bile Acids: Gut bacteria are also important contributors to bile acid metabolism and thus may play a role in the biology linking bile acids to colon cancer. For instance, prolonged consumption of red meat and saturated fatty acids has been associated with increased risk of CRC development. This was observed among descendants of low-risk individuals who moved to developed countries

and converted to a western-type diet[39]. At the gut microbiome level, a high-fat diet may alter the composition of the microbiome by resulting in a bloom of the sulfur-reducing bacterium *Bilophila wadsworthia*, which in turn was shown to exacerbate colitis and inflammation in IL-10 deficient mice as well as modify the composition of the microbiota owing to their strong antimicrobial activities [40]. Bile acids have also been directly implicated in carcinogenesis and increases in bile acid concentrations have been observed in the stool of patients with CRC[41]. The mechanism by which bile salts contribute to CRC is likely due to the generation of DNA-damaging ROS and reactive nitrogen species, which can lead to increased DNA damage and an increase in the mutation rates[42].

It is important to note that although certain metabolites such as lithocholic and deoxycholic acid have been shown to be pro-inflammatory and linked to the development of colon cancer, others, such as ursodeoxycholic acid have been shown to have certain health benefits in pre-clinical and clinical studies. Clostridium, Ruminococcus, and Eubacterium strains, from the commensal microflora, have the capacity to convert chenodeoxycholic acid to ursodeoxycholic acid. The latter have been shown to be beneficial for CRC prevention in patients with a history of adenomas and IBD in small retrospective studies[43].

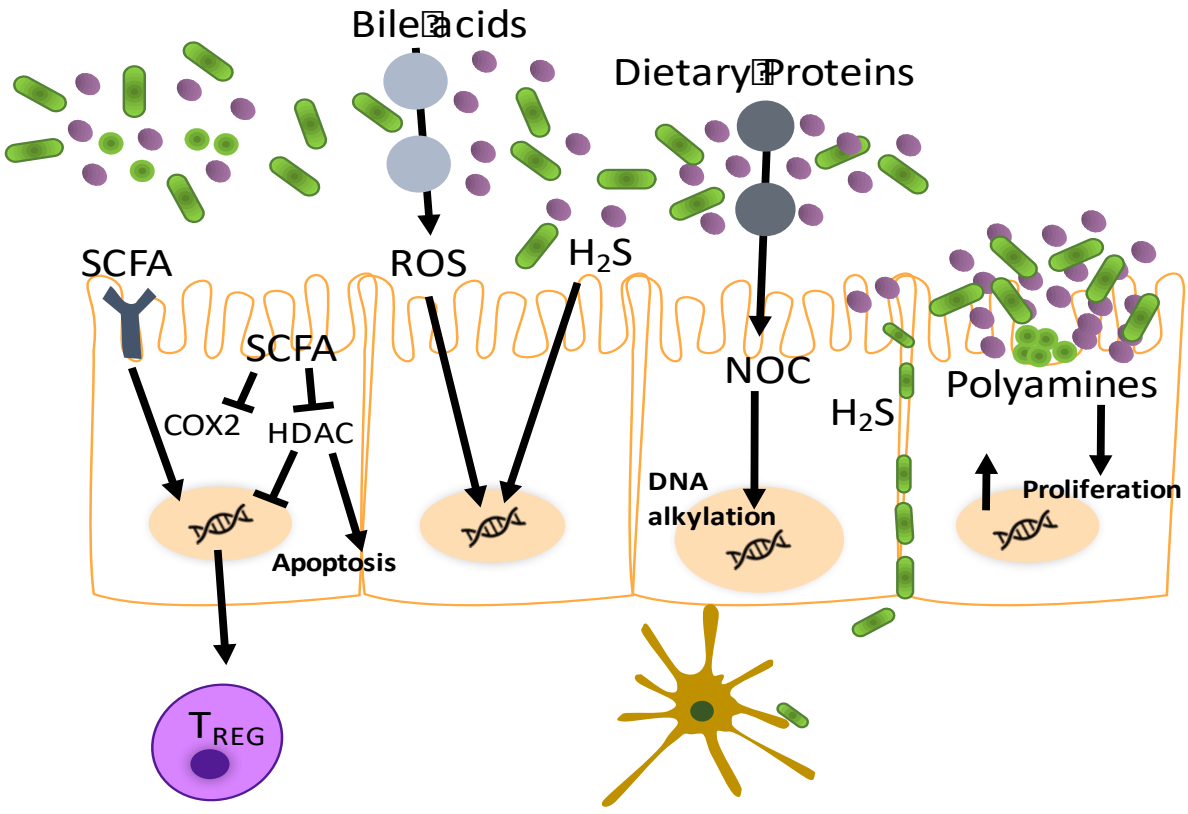


Figure 1.2. Pro- and anti-tumor effects of bacterial metabolites.

Short chain fatty acids (SCFAs) include butyrate, propionate, and acetate. Extracellular or secreted SCFAs can bind to a Growth Protein and provide signals to upregulate important genes needed to modulate T regulatory cell (TREG) differentiation. Intracellular SCFAs block COX2. SCFAs also promote TREGs and inhibit apoptosis through modulation of histone deacetylases. Bile acids promote carcinogenesis via the breakdown by certain commensals that generate DNA-damaging ROS. Hydrogen sulfide (H2S) is produced by many bacteria and can damage DNA or cause a breach in the epithelial barrier, allowing translocation of commensals and generating inflammation. Dietary proteins can be metabolized to N-nitroso compounds that alkylate DNA and cause cancer. Biofilms also can generate polyamines, which can increase cellular proliferation, as discussed in the next section of this chapter.

1.4.MICROBIOTA, BIOFILM FORMATION AND CRC

The commensal bacteria found within our intestines are not all free-floating entities in a planktonic state. Some of these microbes live in complex communities called biofilms. Biofilms are populations of microbes held to each other, to surfaces or at an interface by microbial-produced polymeric matrices. To form and maintain biofilms, bacteria must induce a different set of genes that aids in increasing the concentration of bacteria, a regulation referred to as quorum-sensing. Bacteria in a biofilm may also differ from their planktonic counterparts in antimicrobial resistance and expression of different virulence genes[44]. The five steps of biofilm formation are illustrated and described in Fig. 1.3. It is believed that the intestinal microbiota can form biofilms along the mucosal surface in healthy individuals. However, these data come from human sudden death studies and may be an artifact resulting from the embedding time, as healthy subjects lacked these biofilm-like structures in previous reports[45]. Despite their controversial presence in healthy individuals, biofilms have been associated with both non-malignant pathological conditions such as Inflammatory Bowel Disease and with colon-associated malignancy[45,46,47]. Interestingly, in human CRC, the presence of bacterial biofilms was linked to right-sided colon adenomas and cancers, whereas left-sided cancers were biofilm-negative[47]. Siuzdak and colleagues also went on to show a direct correlation between biofilms and the upregulation of a specific polyamine metabolite, N1, N12-diacetylspermine[47]. However, polyamine production was connected to

biofilm formation and was not specific to the cancer cells associated with the biofilm, as measurement of up-regulated polyamines was identified in paired normal tissue and in the small number of biofilm-positive left-sided cancers[47]. The expression of polyamine is known to enhance eukaryotic cellular proliferation, microbial growth, and cell wall formation[48]. Thus, it could be that the presence of biofilms potentially increases polyamines production creating an environment more amenable to epithelial hyperproliferation and oncogenic transformation.

Regarding the microbial composition of biofilms, it seems that the organization of the biofilms rather than specific compositions may be more important in the pathogenesis of CRC. As mentioned earlier, Dejea et al. (2014) found that most right-sided tumors are associated with biofilms and that the normal colon tissue from these patients was also biofilm-positive[45]. In contrast, normal tissue from patients with biofilm-negative tumors was always biofilm-negative[45]. Furthermore, compositional analysis found that the communities from non-malignant biofilm-free tissues of CRC patients were significantly closer to those found in tissue of healthy volunteers than to those found in biofilm-positive tumor-associated tissues. Yet, non-malignant tissues from CRC patients containing biofilms showed a significantly closer structure to CRC-associated tissue than to the biopsies from healthy, cancer-free volunteers. Thus, the observations from examining these tumor-associated and non-tumors associated tissues suggest that the presence of a biofilm might correlate with changes in bacterial composition.

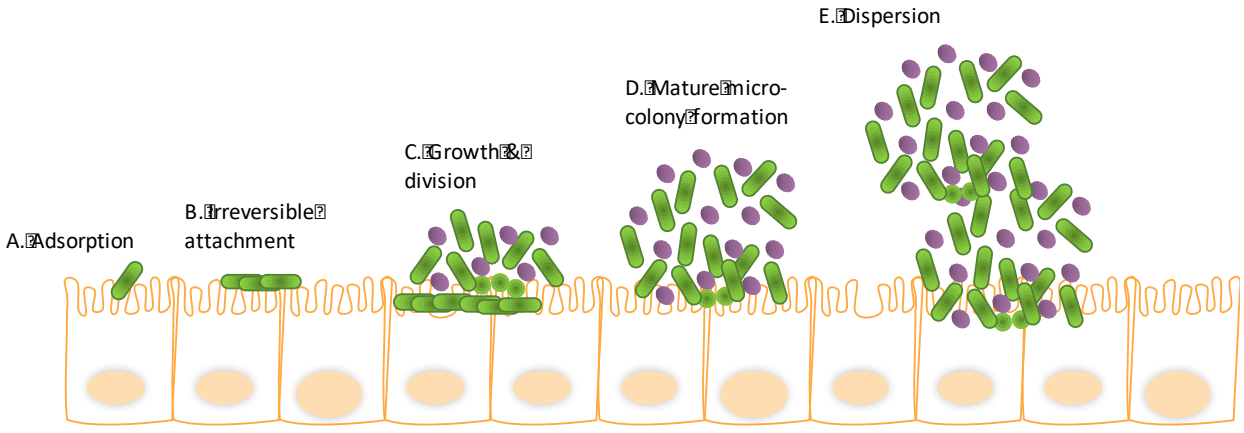


Figure 1.3. The five stages of bacterial biofilm formation.

(A) Bacteria reversibly attach to solid support. (B) Bacteria become irreversibly attached and aggregate to form matrix. (C) Maturation phase: cells become layered and effects of quorum

sensing begin. (D) Clusters reach maximum thickness. (E) Escape of planktonic bacteria from matrix dispersion

1.5. MICROBIOTA-DEPENDENT SENSING BY THE INNATE IMMUNE SYSTEM AND CRC

The function of the inflammatory response is to protect us against microbes and to help repair and regenerate tissue damage caused by both infectious and non-infectious agents. Thus, it is not surprising that there is an intrinsic link between inflammation and cancer. This is demonstrated by the fact that 15% of worldwide cancers are microbial-associated, and that several cancers are associated with chronic inflammatory diseases, such as gastritis, inflammatory bowel disease, and thyroiditis[49]. Our innate immune system imparts our bodies with sensors for tissue damage and infection, with major examples being TLRs and NOD-like receptors. Defective innate immune responses may lead to inadequate pathogen eradication, recurrent tissue injury or failure of anti-inflammatory responses, which can cause chronic inflammation and support tumorigenesis. In CRC, the severity of inflammation strongly correlates with the risk of CRC in patients with IBD and we believe that one of the major factors that drives chronic inflammation in genetically susceptible IBD patients is the commensal microbiota[50,51]. In that light, it makes sense that many of the genetic polymorphisms identified in large-scale, genome-wide association studies of IBD patients were genes directly or indirectly involved in microbial sensing (e.g., TLR, NLR), microbial handling (e.g., autophagy) or innate inflammation (cytokines, chemokines).

1.6. TOLL-LIKE RECEPTORS BIOLOGY

Toll-like receptors are evolutionarily conserved, type I transmembrane pattern recognition receptors (PRRs) that sense conserved microbial motifs also called pathogen associated molecular patterns[52]. Currently, 10 TLRs have been identified (TLR1–TLR10) in humans and 12 in the mouse (TLR1–9, TLR11–13)[53]. TLR1, -2, -4, -5, and -6 are localized at the cell surface, TLR3, -4, -7, -8, and -9 are present in the intracellular compartment[54], and TLR4 can be expressed both intracellularly and extracellularly[52]. TLRs can recognize a variety of ligands such as lipids and lipopeptides (TLR1, -2, -4, -6), bacterial Flagellin (TLR5), and fragments of nucleic acids (TLR3, -7, -8, -9). The expression of TLRs is not limited to immune cells as their expression has also been found on non-immune cells such as epithelial and endothelial cells. Most of the TLRs are expressed

throughout the small and large intestines; however, the localization and function of all the individual TLRs are still the object of active investigations.

Activation of TLRs can induce a few signaling pathways, which results in the up regulation of genes involved in co-stimulation, inflammation, cellular metabolism, survival, and death. Receptor dimerization of a TLR results in differential recruitment of specific adaptor proteins, including MyD88, MyD88 adaptor-like (Mal, also known as TIR domain-containing adaptor protein), TIR domain-containing adaptor-inducing interferon-β (TRIF), or TRIF-related adaptor molecule, which drive subsequent signaling. This results in the activation of a few downstream pathways, including NF-κB, mTOR, interferon regulatory factor, PI3K/Akt and mitogen-activated protein kinase (MAPK) pathways that lead to signals affecting inflammation, regeneration, cell survival, and proliferation[55,56]. To expand their PAMP-recognizing capacity, TLRs also have the capacity to form heterodimers and homodimers and bind different accessory proteins, leading to different signal transduction pathways, such as MD-2 and CD14, that form a complex with TLR4 in response to lipopolysaccharide (LPS). Another example is TLR2, which can switch its ability to produce pro- and anti-inflammatory responses by dimerization with several co-receptors such as TLR2 itself, TLR1, TLR6, and TLR10. Recent studies suggest that TLR2/TLR6 dimerization might activate a signal transduction pathway that induces the transcription of pro-inflammatory molecules, whereas TLR2/ TLR1 dimerization promotes the anti-inflammatory pathway that leads to the expression of IL-10 and the trans-differentiation of Th17 and Treg cells[57].

Toll-like Receptors Associated with CRC & Specific Contributions of Individual TLRs to CRC

The idea that TLRs might be involved in tumorigenesis came about at the end of the nineteenth century, when William Coley observed that repeated injections of a mixture of bacterial toxins from the Gram-positive Bacterium *Streptococcus pneumoniae* and the Gram-negative bacterium *Serratia marcescens* yielded efficient anti-tumor effects, providing proof that microbial products, rather than the infection itself, mediate an anti-tumor effect[58]. Shear and Turner later discovered that LPS, which is a component of the membrane of Gram-negative bacteria, was the fraction of Coley's toxin that accounted for its anti-tumor effect[59]. As LPS stimulates TLR4, these results suggest that the antitumor effect of Coley's toxin was a result of TLR activation.

Toll-like receptors play many reported and sometimes conflicting roles in tumorigenesis. The number of potential roles is likely due to the number of different TLRs, the different sets of genetic signatures induced by each TLR, and the differences in expression within the GI tract.

Toll-like receptor signaling may provide activation of anti-tumor immunity by inducing downstream mediators such as type I interferons. This activation has recently been used to generate TLR agonists as potential candidates for cancer immunotherapy. In contrast, there is much literature on the role of TLRs as tumor promoters through the induction of pathways linked to inflammation, wound healing, tissue regeneration, cell survival, and cell death. Owing to their ability to promote cell survival and proliferation and regulate apoptosis and cell death, the function of a given TLR signal may be highly influenced by many factors such as the tissue microenvironment, the host's genetic signature or the microbiota. TLR sensing of the microbiota is even more complex, as an individual TLR can influence either tumor promotion or anti-tumor immunity, depending on the context in which it is activated. Regardless of the effect, it is well accepted that TLR recognition of the microbiota plays an influential role in CRC, and this has been demonstrated in a number of studies. Indeed, the deletion of the TLR adapter molecule Myd88, the absence of the microbiota in GF mice on the Apc(Min/+) background, or treating mice with broad spectrum antibiotics decreases the incidence and severity of cancer in both sporadic and colitis-associated cancer models[55,56,60,61].

TLR1: Our group has previously shown that genetic deficiency in TLR1 promotes acute enteric infection by gut pathogen Yersinia enterocolitica[62]. Examining that model further, we uncovered an altered cellular immune response that promotes the recruitment of neutrophils, which in turn increases metabolism of the respiratory electron acceptor tetrathionate by Yersinia. These events drive permanent alterations in anti-commensal immunity, microbiota composition, and chronic inflammation, which persist long after Yersinia is cleared[62]. These data demonstrate that acute infection can drive long-term immune and microbiota alterations, leading to chronic inflammatory disease in genetically predisposed individuals and potentially predispose them to cancer. Interestingly, humans express a variant of TLR1 in which a hydrophobic isoleucine is replaced by a hydrophilic serine at the (I602S) transmembrane domain. When the variant allele is expressed, TLR1 does not localize to the surface membrane, but is rather trapped intracellularly in the cytosol of the cell[63]. This reduces the TLR1-mediated activation of NF-κB by extracellular ligands. Recently, we have shown that metastatic CRC patients expressing one or both variant alleles of TLR1 I602S have a better response to treatment in addition to improved progression-free survival, when treated with FOLFIRI (a combination therapy of irinotecan, 5-fluorouracil, and folic acid)

and bevacizumab, a humanized monoclonal antibody and angiogenesis inhibitor, suggesting that endogenous TLR2/1 ligands might play a pro-tumorigenic role[64].

TLR2: The role of TLR2 remains controversial in CRC. In one study, there were no differences in tumor burden between wild-type and TLR2-deficient mice, using a colitis-associated neoplasia model of azoxymethane (AOM) and dextran sulfate sodium (DSS). Increased tumor number and higher IL-6, IL-17A, and phosphor-STAT3 levels were identified in TLR2-deficient mice by a different group using a similar AOM-DSS model[65,66]. The different outcomes seen in TLR2 studies may be explained by the fact that TLR2 can form heterodimers with TLR1, TLR6, and, in humans, TLR10. The ability to bind different TLRs expands the number of ligands that TLR2 recognizes, and potentially alters the downstream signaling molecules associated with activation. Moreover, TLR1 and TLR6 allow TLR2 to sense tri-acylated and di-acylated lipoproteins respectively, and, depending upon it binding TLR2, has been shown to induce different MAPK and genetic signatures[67]. Therefore, if the microbiota of an animal colony differed greatly between two groups, with one containing more TLR2/6-activating bacteria and the other more TLR2/1, then even the baseline microenvironment could be different, and knocking out TLR2 may modulate a response in one instance and not in the other.

TLR4: Toll-like receptor 4 can activate both MyD88-dependent and TRIF-dependent signaling pathways, depending upon its cellular expression[68]. Using the same AOM/DSS model, TLR4 deletion was shown to strongly reduce inflammation and tumor burden, whereas transgenic mice overexpressing constitutively activated TLR4 in the intestine exhibit a higher sensitivity to colitis-associated neoplasia due to activation of β-catenin signaling pathways[69,70]. In contrast, one recent study showed that intestinal overexpression of constitutively activated TLR4 in the Apc(Min/+) model reduces tumor load by increasing tumor cell apoptosis[71].

An anti-tumor role for TLR4 has also been shown. The release of HMGB1, a TLR4 ligand, by damaged or necrotic tumor cells can trigger TLR4 activation in local immune cells, enhancing antigen presentation and promoting anti-tumor immunity. In humans, a TLR4 loss-of-function allele is associated with less cross presentation of antigens and results in relapse and an increase in metastasis in patients with breast cancer[72]. Another human variant, TLR4D299G was recently identified as an aberrant innate immune mediator that may create an auto-inflammatory environment, favoring excessive intestinal epithelial cell (IEC) remodeling and driving tumor

progression. This polymorphism seems to compromise the recruitment of the signaling adaptors MyD88 and TRIF, thereby impairing the downstream activation of NF-κB target genes. Instead, activation of STAT3 is likely the principal target of this SNP, thereby promoting malignant tumor progression in human IECs[73]. Moreover, primary human sporadic adenocarcinomas from patients carrying the TLR4-D299G are more frequently associated with advanced tumor stage.
Lastly, LPS, which can also be derived from Gram-negative commensal bacteria, has been used in phase II clinical trials for the treatment of CRC and was shown to lead to tumor regression when directly injected into adoptively transferred tumors[74].

TLR2/6: Lactic acid bacteria are a group of commensal bacterial strains that were shown to have anti-tumor potential in several probiotic studies (see Sect. "Probiotics"). A recent study also found that several strains such as *Lactobacillus plantarum* CCFM634, *L. plantarum* CCFM734, *L. fermentum* CCFM381, *Lactobacillus acidophilus* CCFM137, and *Streptococcus thermophilus* CCFM218 stimulated TLR2/TLR6, providing an insight into lactic acid bacteria-specific host–microbe interactions[75].

TLR10: Like TLR1 and TLR6, TLR10 can form heterodimers with TLR2. Although TLR6/2 dimerization allows recognition of di-acylated lipoproteins and lactic acid bacteria, no ligand has yet been found for TLR10 and little is known about TLR10 and CRC.

TLR9: Toll-like receptor 9 is activated by both bacterial and viral DNA, immunoglobulin– DNA complexes, and synthetic oligodeoxynucleotides (ODNs), which contain unmethylated CpG sequences[76]. Apical expression and activation of TLR9 on epithelial cells by bacterial DNA fragments have been reported to maintain colonic homeostasis. TLR9-induced type 1 interferons have also been shown to mediate the anti-inflammatory effects in experimental colitis[77]. In CRC, expression of TLR9 has been shown to be higher in adenomatous polyps but decreased in hyperplastic and villous polyps from patients who developed CRC, suggesting that TLR9 expression might play a protective role against malignant transformation[78]. Recently, much research has focused on antitumor immunity induced by TLR9 antagonists and inhibitory ODNs (inh-ODNs; discussed in next section).

TLR Adaptor Molecules: Similar to TLR2, conflicting results were also found in studies looking at the effect of Myd88 deletion on CRC in the Apc(Min/+) model. One study demonstrated that Apc(Min/+) /Myd88-/- mice develop fewer colonic tumors than Apc(Min/+) mice, indicating that bacterial signaling contributes to tumorigenesis in the context of APC mutations[79]. However, a subsequent study found that MyD88-dependent activation of ERK stabilizes β-catenin. In this way, the absence of MyD88 protected against APC-dependent tumors[80] (Lee et al. 2010). Thus, depending on the model used, MyD88-deficiency either protects or increases tumorigenesis.

Manipulation of the Microbiome as a Treatment for CRC

Specific microbiota species and their products constitute potential targets for modulating colon cancer because of their immunological potential and/or protective effects on colon carcinogenesis. A large body of evidence exists chronicling the influence of certain commensal micro-organisms on both the development of the mucosal immune system and the modulation of innate inflammatory responses to maintain homeostasis. The commensal microbiota has evolved several mechanisms that help to modulate the inflammatory response as it has been shown to induce anti-inflammatory cytokines such as IL-10 and TGF-β that directly suppress inflammatory signaling, reduce antigen presentation, and can induce regulatory T cells. Other commensal species may directly inhibit colon cancer through the generation of ROS that can, in combination with platinum compounds, greatly enhance the effect of chemotherapies.

In addition to their immune-modulatory effects, commensal microbiota has been shown to reduce cell cycle progression, induce apoptosis of tumor cells, and alter the host's metabolome, by producing specific bacterial enzymes that enhance the production of beneficial or protective metabolites. Moreover, the microbiota and its products may be exploited for diagnosis and detection purposes. For example, comparing the microbial characteristics of patients with CRC with those of IBD patients could begin to dissociate whether the changes observed in the microbiota are inflammation- or cancer-dependent.

1.7. SPECIFIC THERAPIES TARGETING THE MICROBIOTA TO TREAT CRC

Probiotics According to the US Food and Drug Administration (FDA), probiotics are live microorganisms that are intended to have health benefits. Products sold as probiotics include foods, dietary supplements, and products that can be used topically. Although some probiotics have shown promise in research studies, strong scientific evidence to support the specific uses of probiotics for most health conditions is lacking. The FDA has not yet approved any probiotics for preventing or treating any health problems. Some experts have also cautioned that the rapid growth in marketing and use of probiotics may have outpaced scientific research for many of their proposed uses and benefits, as we have yet to define the therapeutic windows for probiotic dosing, the potential side effects, and the accurate stratification of patients most likely to benefit from such therapies.

However, *in-vitro* and *in-vivo* studies, in addition to some clinical trials, do suggest that certain microbial species in colon cancer might have a promising beneficial effect. For instance, a particular study showed that the administration of *Lactobacillus rhamnosus* GG (LGG) induced a significant reduction in polyamine biosynthesis in HGC-27 and DLD-1 cancer lines, suggesting that LGG might have the potential to change the host's metabolome and consequently halt the proliferation of tumor cells[81] (Orlando et al. 2009). As mentioned earlier, polyamine is a metabolite that is necessary for cell proliferation. It was also found to be produced by microbial biofilms. Thus, targeting polyamine production and biofilm interactions could be another strategy for treating CRC, for example, by the administration of probiotic strains with adhesive properties, that can prevent the establishment of pathogenic biofilms by competitive exclusion. This ability of probiotics to adhere to GI mucus is of considerable importance in their potential to exert a modulatory effect *in situ* and the adhesion of probiotic bacteria to epithelial cells has been shown to prevent the establishment of pathogens[82]. Resta-Lenert and Barret showed that exposure of biofilms to live, but not heat-inactivated, probiotic *S. thermophilus* and *L. acidophilus* strains significantly limited adhesion, invasion, and physiological dysfunction induced by exposure to an entero-invasive strain of *E. coli*[83]. A similar effect has been demonstrated for a probiotic strain of *L. plantarum*, which had a protective effect against damage to the integrity of epithelial cancer cell line Caco-2 monolayers and the structure and distribution of TJ proteins by enteroinvasive *E.*

coli[84]. Another study showed that Caco-2 cells exposed to *L. plantarum* bacteria significantly induced human anti-microbial peptide beta-defensin 2 mRNA expression and secretion in a dose-dependent manner compared with controls. This was inhibited by anti-TLR2 neutralizing antibodies, suggesting that *L. plantarum* may signal through this microbial pattern recognition receptor and generate an anti-cancer response[85]. *In vivo* examples include a study that showed that daily oral administration of micro-encapsulated *L. acidophilus* in a yogurt formulation to Apc(Min/+) mice resulted in significant suppression of colon tumor incidence, tumor multiplicity, and reduced tumor size. Moreover, the treated animals exhibited fewer gastro-intestinal intra-epithelial neoplasia with a lower grade of dysplasia in tumors, that this probiotic might have a potential benefit[86]. In another study, Park et al. fed F344 male rats with *Bacillus polyfermenticus*, showing that these rats displayed significantly lower numbers of aberrant crypt foci than the control group. Supplementation with *B. polyfermenticus* induced less leukocytic DNA damage and plasma lipid peroxidation levels, in addition to a lower plasma total antioxidant potential, suggesting that *B. polyfermenticus* might exert a protective effect on the antioxidant system and the process of colon carcinogenesis[87].

Last, the potential benefit of certain microorganisms was also tested in some clinical trials involving human subjects. In one study, the administration of *Lactobacillus casei* was evaluated as a method of preventing the occurrence of colorectal tumors. The occurrence of tumors with a grade of moderate atypia or higher was significantly lower in the patient group after 2–4 years of treatment with *L. casei* compared with the control group[88]. In another 12-week clinical trial completed back in 2007, polypectomies patients were treated with LGG, *Bifidobacterium lactis* Bb12, and oligofructose-enriched inulin. The treatment resulted in significant changes in fecal microbiota of the patients, as the proportion of Bifidobacterium and Lactobacillus increased, and *Clostridium perfringens* decreased. The intervention was also associated with a significant reduction of colorectal epithelial cell proliferation in the patients[89]. Another study reported in a prospective trial that in 31 CRC patients, *Lactobacilli johnsonii* (La1), but not *Bifidobacterium longum* (BB536), affects intestinal microbiota by reducing the concentration of pathogens and modulating intestinal dendritic cell (DC) and T-cell specific responses exemplified by a reduction in the release of anti-inflammatory cytokines and a reduction of potentially pathogenic bacteria. Moreover, DCs isolated from patients colonized with La1 had a significantly blunted ability to proliferate, and it is possible that La1 might prevent excessive activation of DCs and the development of Th1-polarized immunity in the intestinal mucosa[90]. Recently, a cohort study with

12 years' follow-up on 45,241 volunteers determined that high yogurt intake was significantly associated with decreased CRC risk, suggesting that the long-term administration of probiotic formulations might reduce the incidence of CRC[91]. These studies show promising findings, but we have yet to elucidate the mechanisms underlying these potential therapeutic effects. Perhaps unraveling how probiotics may contribute to CRC treatment will uncover new insights into cancer immunotherapy.

Moreover, despite their popularity, robust and reproducible data regarding probiotics efficacy is lacking and most studies out there remain unfortunately widely inconclusive to this day. Consequently, both the European Food Safety Authority and the US Food and Drug Administration have yet to approve any probiotic formulation available on the market as a valid therapeutic option. The current confusion stems mainly from the fact that none of the data about probiotics so far is the result of rigorously planned large-scale randomized and blinded clinical trials and that the results reported are often tinted with commercial interests. Another factor contributing to the variability and confusion in probiotics research is the inconsistency in the studied strains. As most formulations contain microorganisms belonging to the Lactobacillus and Bifidobacterium genera, as well as Lactococcus species, *Streptococcus thermophilus*, *E. coli Nissle* 1917 and the yeast *Saccharomyces boulardii*, it is important to note that not all strains belonging to these genera and species have the same mechanisms of action. Moreover, humans being largely different in terms of diet, age range, genetic background, and gut microbiome configuration, it is challenging to market a "one size fits all" probiotic pill that benefits everyone without any discrimination. Another critically important point regarding probiotics supplementation is whether these microorganisms actually "stick around" in the gut long enough to exert a therapeutic effect and here too, patient and microbiome factors seem to play a role, stressing the importance of tailoring probiotics to individuals.

Safety-wise, some reports have shown that the use of probiotics could be associated with a higher risk of infection and disease in young infants and immuno-compromised patients. Moreover, as Lactobacillus species can directly inhibit human microbiome growth, two studies have shown that supplementation with probiotics following antibiotic treatment was associated with the failure of the microbiome to recover its pre-antibiotic composition, which could predispose to long-term health effects such as type 1 and 2 diabetes, obesity, asthma, arthritis, allergies, and inflammatory bowel disease.

Given these observations and in light of the unfortunate historical lack of sufficient medical regulation for currently available probiotics, one cannot stress enough the importance of a formal regulatory approval process for probiotics, as it is the case for any other human medical intervention. It is also very important for researchers to deepen their understanding of the interactions between the human host and its resident microbes and probiotic microorganisms. Moreover, it is crucial for well-planned large-scale randomized and blinded clinical trials, preferentially devoid of commercial interests, to be the reference for any data-driven decision-making in the clinical settings. Clinical outcomes from these trials should be objectively assessed and analyzed to account for inter-individual differences. Adverse reactions should be better studied, reported and published, particularly in children, and immunosuppressed and critically ill individuals. Ideally, data should also be shared and made accessible in order to allow for a global collaborative effort to reproduce positive results before any guidelines are drafted.

Fecal Microbiota Transplantation: Fecal microbiota transplantation (FMT) refers to the idea of administering stool from a healthy person termed as the "donor" to a patient with a disease believed to be related to an unhealthy gut microbiome, most often in the form of an orally ingestible capsule. The donor may be a family member, a partner, or a friend or even an unrelated volunteer. It can also be the patient himself, from freezing his or her own stool prior to receiving an intensive therapy that might disrupt the microbiome, such as antibiotics treatments. This peculiar idea is not new. In fact, it goes back to 4th century China where human fecal suspensions were orally administered to victims of food poisoning and severe diarrhea. Lately, it has re-appeared as a successfully used treatment of an emerging healthcare problem in the world known as *Clostridium difficile* infection. Illness from the bacteria *C. difficile* most commonly affects patients in hospitals and typically occurs after the use of antibiotic medications. It manifests itself as a severe type of diarrhea. The treatment of choice of *C. difficile* infection is the oral administration of an antibiotic called vancomycin but in recent years, *C. difficile* infections have become more frequent, severe and difficult to treat. In that regard, FMT constitutes a successful therapeutic alternative that is currently being used in hospitals. In addition to that specific indication, FMT is being actively investigated as a therapeutic strategy for many other diseases such as Inflammatory Bowel Disease, Obesity, Irritable bowel syndrome but also metabolic syndrome, Type 2 diabetes mellitus, fatty liver disease, multi-drug resistant organism eradication, hepatic encephalopathy, and

pediatric allergies. Despite some promising preliminary results in certain cases, large randomized controlled trials as well as long-term clinical follow-ups have yet to be conducted before FMT can be considered as an efficient and safe therapeutic option in health conditions besides *C. difficile* infection. Safety-wise, it is crucial for the donor stool and resulting FMT to be screened for pathogens in order to avoid the transmission of potentially dangerous microorganisms that could result in serious health consequences, if not death, especially in immunocompromised patients.

Finally, it is important to keep in mind that the gut microbiota is a very complex ecosystem with many components that still escape our understanding. But despite the many challenges, the rapid technological and computational advances in human gut microbiome research offer great hope by allowing us to better analyze and understand defined microbial consortia and their effect on our gut microbiome and our health in order to hopefully target these consortia to the treatment of specific diseases someday. Significant clinical effectiveness of FMT has been demonstrated for recurrent *Clostridium difficile* infection and ongoing studies are investigating FMT for other diseases such as IBD[92]. Transplantation of autologous microbial populations to recipients likely triggers mucosal immune responses that, depending on the microbiota composition and the recipient's genotype, could range from pro- to anti-inflammatory. However, the impact of FMT on the recipient immune system is complex and unpredictable, and the ongoing discovery of commensal microbes and investigation of their impact on the host will lead to the development of new personalized probiotic agents and microbial consortia that will eventually replace FMT.

TLR Agonists and Antagonists: We are beginning to understand the therapeutic potential of individual TLRs in the fight against colon cancer through several recent clinical trials and through pre-clinical studies[93]. However, the development of optimal therapeutic strategies for targeting TLR signaling depends on a fuller characterization of the roles of TLR and their downstream adaptors under healthy and diseased conditions in the intestine, and through a more thorough understanding of the molecular mechanisms behind their role in the pathophysiology of CRC. TLRs in the gut are constantly exposed to commensal derived ligands, inducing downstream signaling pathways that contribute to gut mucosal immunity and homeostasis. The pleiotropic nature of TLRs and their ability to activate/modulate many important cellular mechanisms including intestinal permeability, inflammation, cell survival and death, regeneration and repair, autophagy, and tolerance makes them candidates for targeted intervention. Various strategies for

modulating TLR signaling exist, including administration of specific commensal species and/or ligands, administration of antibodies to block TLR signaling, and directly targeting the signaling of TLRs on tumor cells to trigger an anti-tumor innate and/or adaptive immune response.

As mentioned above, TLR9 signaling is not associated with carcinogenesis, but rather plays an important role in generating anti-tumor responses, making it a very promising therapeutic choice. Pre-treatment with the TLR9 ligand CpG DNA was recently shown to ameliorate both chemically induced and spontaneous colitis in mice by inhibiting pro-inflammatory cytokines and chemokines[78]. TLR9 agonists were also shown to suppress tumor growth, and metastasis of CRC cells to the lung and liver, thus increasing overall survival when administered alone or in combination with chemotherapy regimens in CRC xenografts[94,95,96]. The effects of TLR9 have been observed at the site of drug delivery and at systematic sites and are dependent upon the presence of MyD88[97,98]. TLR9 agonists are also being used to boost the anti-tumoral immune response of the host in the context of antigen-loaded DC immunotherapy. The combination of DC and TLR9 agonists have been used in CRC and have demonstrated a stronger immune response, both local and systemic anti-tumoral effects, and the induction of an antitumor memory response in the host[99].

Toll-like receptor 4 signaling has been shown to perpetuate carcinogenesis through the induction of prostaglandins and through the activation of β-catenin. Thus, patients with CRC may benefit from administration of antagonists that can inhibit TLR4 signaling. A recent study by Wei-Ting Kuo et al. showed that the administration of a synthetic analog of LPS that inhibits TLR4 signaling caused an increase in tumor cell apoptosis and decreased the tumor burden[100]. Like TLR9, TLR7 agonists seem to promote anti-tumor responses. For example, intravenous administration of the selective TLR7 agonist DSR-29133 led to enhanced anti-tumor effects in a solid tumor model[101].

Toll-like receptors can also be finely tuned to optimize tumor responsiveness to chemotherapy through their microbial recognition. Experimentally, polyI:C, a TLR3 agonist, was shown to enhance cycloheximide-induced apoptosis of tumor cells, whereas TLR9 agonist administered with chemotherapy regimens increased the overall survival of mice in a CRC xenograft model[102]. These studies are opening a new range of clinical applications for TLR agonists as an adjuvant for chemotherapy[103,104]. Table 8.1 summarizes information regarding the various human TLRs that have been implicated in colon cancer, their natural and synthetic ligands, their location throughout the large intestine, and their therapeutic potential.

Toll-Like Receptors (TLRs)	Ligands	Location in the large intestine	Effects on immunological functions associated with CRC
TLR2/1	Tri-acylated lipopeptides$_m$	Myofibroblasts	TLR1 I602S variant is associated with better response to CRC treatment and progression-free survival
TLR2	PG$_m$, LP$_m$, Zymosan$_m$, GPI$_m$, Tc52$_m$	LP macrophages, myofibroblast, circulating PMNs	TLR2KO had increased tumor number, IL6, IL17A, p-STAT3 in AOM/DSS CRC mouse model
TLR3	IMQ$_m$ Poly(I:C)$_m$	Epithelial cells (IECs) Dentritic cells (DCs)	Experimentally, polyI:C, was shown to enhance cycloheximide-induced apoptosis of tumor cells
TLR4	LPS$_m$, MPLA$_m$, Mannan$_m$	Endothelial cells	TLR4 signaling enhances carcinogenesis through the induction of STAT3, prostaglandins, and activation of β-catenin
	Glycoinositol phospholipids$_m$	LP macrophages, IECs, DCs	Patients with CRC may benefit from administration of antagonists that can inhibit TLR4 signaling
		Myofibroblasts Circulating PMNs	Phase II clinical trials using LPS for the treatment of CRC lead to tumor regression when directly injected into adoptively transferred tumors

TLR5	Flagellin$_m$	LP endothelial cells, LP macrophages, myofibroblasts, circulating PMNs	Lack of MyD88 or TLR5 expression dramatically enhanced tumor growth and inhibited tumor necrosis in mouse xenografts of human colon cancer (Rhee et al. 2008)
TLR2/6	Di-acylated lipopeptides$_m$	Myofibroblasts	No association with CRC
TLR7	RSQ 848$_s$	Myofibroblasts	IV administration of the selective TLR7 agonist DSR-29133 led to enhanced anti-tumor effects in a solid tumor model
TLR9	CpG DNA$_m$	Myofibroblasts	Anti-tumor immunity induced by TLR9 antagonists and inhibitory oligodeoxynucleotides
	Hemozoin$_m$	Circulating PMNs	TLR9 agonists suppress tumor growth, and metastasis of CRC, increasing overall survival when administered alone or in combination with chemotherapy regimens in CRC xenografts
	HMGB1$_h$		DC and TLR9 agonists strengthen local and systemic anti-tumoral effects, and induce anti-tumor memory response

Table 1.1 Toll-like receptor (TLR) biology and therapeutic potential in colorectal cancer (CRC)[105]

h: human, m: microbial, s: synthetic, PG: peptidoglycans, LP: lipoproteins, GPI: glycophosphatidylinositol, HSP: heat shock protein, HMGB1: high mobility group box protein 1, IMQ: imiquimod-R837, LPS: lipopolysaccharide, MPLA: monophosphorylate lipid A, RSQ 848: resiquimod, CpG: oligodeoxynucleotide, LP: lamina propria, PMNs: polymorphonuclear leukocytes.

In conclusion, the gut microbiota possesses a huge genomic and metabolic potential. It can also have an impact on the immune system through microbial PRR receptors that live near these commensal microorganisms in the gut. Thus, the gut microbiota constitutes a potential target for modulating colon cancer risk, through diagnostic or treatment strategies, offering a very exciting field of investigation and discoveries for scientists.

References:

1. Bardhan, K. & Liu, K. Epigenetics and Colorectal Cancer Pathogenesis. *Cancers (Basel)* **5**, 676–713 (2013).
2. Valle, L. Genetic predisposition to colorectal cancer: Where we stand and future perspectives. *World J Gastroenterol* **20**, 9828–9849 (2014).
3. Karin, M. NF-κB as a Critical Link Between Inflammation and Cancer. *Cold Spring Harb Perspect Biol* **1**, (2009).
4. Grivennikov, S. I., Greten, F. R. & Karin, M. Immunity, Inflammation, and Cancer. *Cell* **140**, 883–899 (2010).
5. Zackular, J. P. *et al.* The gut microbiome modulates colon tumorigenesis. *MBio* **4**, e00692-00613 (2013).
6. Belkaid, Y. & Hand, T. Role of the Microbiota in Immunity and inflammation. *Cell* **157**, 121–141 (2014).
7. Sobhani, I. *et al.* Microbial Dysbiosis in Colorectal Cancer (CRC) Patients. *PLOS ONE* **6**, e16393 (2011).
8. Chen, W., Liu, F., Ling, Z., Tong, X. & Xiang, C. Human Intestinal Lumen and Mucosa-Associated Microbiota in Patients with Colorectal Cancer. *PLOS ONE* **7**, e39743 (2012).
9. Geng, J., Fan, H., Tang, X., Zhai, H. & Zhang, Z. Diversified pattern of the human colorectal cancer microbiome. *Gut Pathogens* **5**, 2 (2013).

10. Sears, C. L. & Garrett, W. S. Microbes, microbiota, and colon cancer. *Cell Host Microbe* **15**, 317–328 (2014).

11. Qin, J. *et al.* A human gut microbial gene catalog established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).

12. Reuter, S., Gupta, S. C., Chaturvedi, M. M. & Aggarwal, B. B. Oxidative stress, inflammation, and cancer: How are they linked? *Free Radic Biol Med* **49**, 1603–1616 (2010).

13. Reynolds, J. G., Silva, E. & McCormack, W. M. Association of Streptococcus bovis Bacteremia with Bowel Disease. *J Clin Microbiol* **17**, 696–697 (1983).

14. Hoen, B. *et al.* Tumors of the colon increase the risk of developing Streptococcus bovis endocarditis: case-control study. *Clin. Infect. Dis.* **19**, 361–362 (1994).

15. Abdulamir, A. S., Hafidh, R. R. & Abu Bakar, F. The association of Streptococcus bovis/gallolyticus with colorectal tumors: the nature and the underlying mechanisms of its etiological role. *J. Exp. Clin. Cancer Res.* **30**, 11 (2011).

16. Boleij, A., Schaeps, R. M. J. & Tjalsma, H. Association between Streptococcus bovis and Colon Cancer. *J. Clin. Microbiol.* **47**, 516–516 (2009).

17. Wang, X. *et al.* 4-hydroxy-2-nonenal mediates genotoxicity and bystander effects caused by Enterococcus faecalis-infected macrophages. *Gastroenterology* **142**, 543-551.e7 (2012).

18. Wu, S. *et al.* A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat Med* **15**, 1016–1022 (2009).

19. Toprak, N. U. *et al.* A possible role of Bacteroides fragilis enterotoxin in the aetiology of colorectal cancer. *Clin. Microbiol. Infect.* **12**, 782–786 (2006).

20. Shen, X. J. *et al.* Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. *Gut Microbes* **1**, 138–147 (2010).

21. Apperloo-Renkema, H. Z., Van der Waaij, B. D. & Van der Waaij, D. Determination of colonization resistance of the digestive tract by biotyping of Enterobacteriaceae. *Epidemiol Infect* **105**, 355–361 (1990).

22. Mukhopadhya, I., Hansen, R., El-Omar, E. M. & Hold, G. L. IBD—what role do Proteobacteria play? *Nat Rev Gastroenterol Hepatol* **9**, 219–230 (2012).

23. Darfeuille-Michaud, A. *et al.* High prevalence of adherent-invasive Escherichia coli associated with ileal mucosa in Crohn's disease. *Gastroenterology* **127**, 412–421 (2004).

24. Nesić, D., Hsu, Y. & Stebbins, C. E. Assembly and function of a bacterial genotoxin. *Nature* **429**, 429–433 (2004).

25. Prorok-Hamon, M. *et al.* Colonic mucosa-associated diffusely adherent afaC+ Escherichia coli expressing lpfA and pks are increased in inflammatory bowel disease and colon cancer. *Gut* **63**, 761–770 (2014).
26. Arthur, J. C. *et al.* Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* **338**, 120–123 (2012).
27. Cuevas-Ramos, G. *et al.* Escherichia coli induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11537–11542 (2010).
28. Jin, Y. *et al.* Hemolytic E. coli Promotes Colonic Tumorigenesis in Females. *Cancer Res* **76**, 2891–2900 (2016).
29. Strauss, J. *et al.* Invasive potential of gut mucosa-derived Fusobacterium nucleatum positively correlates with IBD status of the host. *Inflamm. Bowel Dis.* **17**, 1971–1978 (2011).
30. McCoy, A. N. *et al.* Fusobacterium is associated with colorectal adenomas. *PLoS ONE* **8**, e53653 (2013).
31. Lee, P. & Tan, K. S. Fusobacterium nucleatum activates the immune response through retinoic acid-inducible gene I. *J. Dent. Res.* **93**, 162–168 (2014).
32. Rubinstein, M. R. *et al.* Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. *Cell Host Microbe* **14**, 195–206 (2013).
33. Abed, J. *et al.* Fap2 Mediates Fusobacterium nucleatum Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc. *Cell Host Microbe* **20**, 215–225 (2016).
34. Kostic, A. D. *et al.* Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Res.* **22**, 292–298 (2012).
35. Soler, A. P. *et al.* Increased tight junctional permeability is associated with the development of colon cancer. *Carcinogenesis* **20**, 1425–1431 (1999).
36. Grivennikov, S. I. *et al.* adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature* **491**, 254–258 (2012).
37. Loh, Y. H. *et al.* N-Nitroso compounds and cancer incidence: the European Prospective Investigation into Cancer and Nutrition (EPIC)-Norfolk Study. *Am. J. Clin. Nutr.* **93**, 1053–1061 (2011).
38. Roisin Hughes, I. R. R. Metabolic Activities of the Gut Microflora in Relation to Cancer. *Microbial Ecology in Health and Disease* **12**, 179–185 (2000).

39. Berg, J. W., Howell, M. A. & Silverman, S. J. Dietary hypotheses and diet-related research in the etiology of colon cancer. *Health Serv Rep* **88**, 915–924 (1973).

40. Devkota, S. *et al.* Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in IL-10-/- mice. *Nature* **487**, 104–108 (2012).

41. Barrasa, J. I., Olmo, N., Lizarbe, M. A. & Turnay, J. Bile acids in the colon, from healthy to cytotoxic molecules. *Toxicol In Vitro* **27**, 964–977 (2013).

42. Ajouz, H., Mukherji, D. & Shamseddine, A. Secondary bile acids: an underrecognized cause of colon cancer. *World J Surg Oncol* **12**, 164 (2014).

43. Carey, E. J. & Lindor, K. D. Chemoprevention of colorectal cancer with ursodeoxycholic acid: cons. *Clin Res Hepatol Gastroenterol* **36 Suppl 1**, S61-64 (2012).

44. Burmølle, M., Ren, D., Bjarnsholt, T. & Sørensen, S. J. Interactions in multispecies biofilms: do they actually matter? *Trends Microbiol.* **22**, 84–91 (2014).

45. Dejea, C. M. *et al.* Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18321–18326 (2014).

46. Swidsinski, A., Weber, J., Loening-Baucke, V., Hale, L. P. & Lochs, H. Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J. Clin. Microbiol.* **43**, 3380–3389 (2005).

47. Johnson, C. H. *et al.* Metabolism links bacterial biofilms and colon carcinogenesis. *Cell Metab* **21**, 891–897 (2015).

48. Gerner, E. W. & Meyskens, F. L. Polyamines and cancer: old molecules, new understanding. *Nat. Rev. Cancer* **4**, 781–792 (2004).

49. Selgrad, M. *et al.* The role of viral and bacterial pathogens in gastrointestinal cancer. *J. Cell. Physiol.* **216**, 378–388 (2008).

50. Rutter, M. D. *et al.* Cancer surveillance in longstanding ulcerative colitis: endoscopic appearances help predict cancer risk. *Gut* **53**, 1813–1816 (2004).

51. Gupta, R. B. *et al.* Histologic inflammation is a risk factor for progression to colorectal neoplasia in ulcerative colitis: a cohort study. *Gastroenterology* **133**, 1099–1105; quiz 1340–1341 (2007).

52. Pasare, C. & Medzhitov, R. Toll-like receptors: linking innate and adaptive immunity. *Microbes Infect.* **6**, 1382–1387 (2004).

53. Gay, N. J. & Gangloff, M. Structure and function of Toll receptors and their ligands. *Annu. Rev. Biochem.* **76**, 141–165 (2007).

54. Akira, S. & Hemmi, H. Recognition of pathogen-associated molecular patterns by TLR family. *Immunol. Lett.* **85**, 85–95 (2003).

55. Fitzgerald, K. A. *et al.* Mal (MyD88-adapter-like) is required for Toll-like receptor-4 signal transduction. *Nature* **413**, 78–83 (2001).

56. Kawai, T. & Akira, S. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. *Immunity* **34**, 637–650 (2011).

57. Melmed, G. *et al.* Human Intestinal Epithelial Cells Are Broadly Unresponsive to Toll-Like Receptor 2-Dependent Bacterial Ligands: Implications for Host-Microbial Interactions in the Gut. *The Journal of Immunology* **170**, 1406–1415 (2003).

58. Coley, W. B. The treatment of malignant tumors by repeated inoculations of erysipelas. With a report of ten original cases. 1893. *Clin. Orthop. Relat. Res.* 3–11 (1991).

59. Lundin, J. I. & Checkoway, H. Endotoxin and Cancer. *Environ Health Perspect* **117**, 1344–1350 (2009).

60. Couturier-Maillard, A. *et al.* NOD2-mediated dysbiosis predisposes mice to transmissible colitis and colorectal cancer. *J. Clin. Invest.* **123**, 700–711 (2013).

61. Klimesova, K. *et al.* Altered gut microbiota promotes colitis-associated cancer in IL-1 receptor-associated kinase M-deficient mice. *Inflamm. Bowel Dis.* **19**, 1266–1277 (2013).

62. Kamdar, K. *et al.* Genetic and Metabolic Signals during Acute Enteric Bacterial Infection Alter the Microbiota and Drive Progression to Chronic Inflammatory Disease. *Cell Host Microbe* **19**, 21–31 (2016).

63. Schumann, R. R. & Tapping, R. I. Genomic variants of TLR1--it takes (TLR-)two to tango. *Eur. J. Immunol.* **37**, 2059–2062 (2007).

64. Okazaki, S. *et al.* Clinical Significance of TLR1 I602S Polymorphism for Patients with Metastatic Colorectal Cancer Treated with FOLFIRI plus Bevacizumab. *Mol. Cancer Ther.* **15**, 1740–1745 (2016).

65. Salcedo, R. *et al.* MyD88-mediated signaling prevents development of adenocarcinomas of the colon: role of interleukin 18. *J. Exp. Med.* **207**, 1625–1636 (2010).

66. Lowe, E. L. *et al.* Toll-like receptor 2 signaling protects mice from tumor development in a mouse model of colitis-induced cancer. *PLoS ONE* **5**, e13027 (2010).

67. Depaolo, R. W. *et al.* Toll-like receptor 6 drives differentiation of tolerogenic dendritic cells and contributes to LcrV-mediated plague pathogenesis. *Cell Host Microbe* **4**, 350–361 (2008).

68. Troutman, T. D., Bazan, J. F. & Pasare, C. Toll-like receptors, signaling adapters and regulation of the pro-inflammatory response by PI3K. *Cell Cycle* **11**, 3559–3567 (2012).

69. Fukata, M. *et al.* Innate immune signaling by Toll-like receptor-4 (TLR4) shapes the inflammatory microenvironment in colitis-associated tumors. *Inflamm. Bowel Dis.* **15**, 997–1006 (2009).

70. Santaolalla, R. *et al.* TLR4 activates the β-catenin pathway to cause intestinal neoplasia. *PLoS ONE* **8**, e63298 (2013).

71. Li, Y. *et al.* Constitutive TLR4 signalling in intestinal epithelium reduces tumor load by increasing apoptosis in APC(Min/+) mice. *Oncogene* **33**, 369–377 (2014).

72. Apetoh, L. *et al.* Toll-like receptor 4-dependent contribution of the immune system to anticancer chemotherapy and radiotherapy. *Nat. Med.* **13**, 1050–1059 (2007).

73. Cario, E. The human TLR4 variant D299G mediates inflammation-associated cancer progression in the intestinal epithelium. *Oncoimmunology* **2**, e24890 (2013).

74. Otto, F. *et al.* Phase II trial of intravenous endotoxin in patients with colorectal and non-small cell lung cancer. *Eur. J. Cancer* **32A**, 1712–1718 (1996).

75. Ren, C. *et al.* Identification of TLR2/TLR6 signalling lactic acid bacteria for supporting immune regulation. *Scientific Reports* **6**, 34561 (2016).

76. Kanzler, H., Barrat, F. J., Hessel, E. M. & Coffman, R. L. Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists. *Nat. Med.* **13**, 552–559 (2007).

77. Katakura, K. *et al.* Toll-like receptor 9-induced type I IFN protects mice from experimental colitis. *J. Clin. Invest.* **115**, 695–702 (2005).

78. Rachmilewitz, D. *et al.* Immunostimulatory DNA ameliorates experimental and spontaneous murine colitis. *Gastroenterology* **122**, 1428–1441 (2002).

79. Rakoff-Nahoum, S. & Medzhitov, R. Regulation of spontaneous intestinal tumorigenesis through the adaptor protein MyD88. *Science* **317**, 124–127 (2007).

80. Lee, S. H. *et al.* ERK activation drives intestinal tumorigenesis in Apc(min/+) mice. *Nat. Med.* **16**, 665–670 (2010).

81. Orlando, A., Messa, C., Linsalata, M., Cavallini, A. & Russo, F. Effects of Lactobacillus rhamnosus GG on proliferation and polyamine metabolism in HGC-27 human gastric and DLD-1 colonic cancer cell lines. *Immunopharmacol Immunotoxicol* **31**, 108–116 (2009).

82. Russo, F., Linsalata, M. & Orlando, A. Probiotics against neoplastic transformation of gastric mucosa: Effects on cell proliferation and polyamine metabolism. *World J Gastroenterol* **20**, 13258–13272 (2014).
83. Resta-Lenert, S. & Barrett, K. E. Live probiotics protect intestinal epithelial cells from the effects of infection with enteroinvasive Escherichia coli (EIEC). *Gut* **52**, 988–997 (2003).
84. Qin, H., Zhang, Z., Hang, X. & Jiang, Y. L. plantarum prevents Enteroinvasive Escherichia coli-induced tight junction proteins changes in intestinal epithelial cells. *BMC Microbiology* **9**, 63 (2009).
85. Paolillo, R., Romano Carratelli, C., Sorrentino, S., Mazzola, N. & Rizzo, A. Immunomodulatory effects of Lactobacillus plantarum on human colon cancer cells. *Int. Immunopharmacol.* **9**, 1265–1271 (2009).
86. Urbanska, A. M., Bhathena, J., Cherif, S. & Prakash, S. Orally delivered microencapsulated probiotic formulation favorably impacts polyp formation in APC (Min/+) model of intestinal carcinogenesis. *Artif Cells Nanomed Biotechnol* **44**, 1–11 (2016).
87. Park, E., Jeon, G.-I., Park, J.-S. & Paik, H.-D. A probiotic strain of Bacillus polyfermenticus reduces DMH induced precancerous lesions in F344 male rat. *Biol. Pharm. Bull.* **30**, 569–574 (2007).
88. Ishikawa, H. *et al.* Randomized trial of dietary fiber and Lactobacillus casei administration for prevention of colorectal tumors. *Int. J. Cancer* **116**, 762–767 (2005).
89. Rafter, J. *et al.* Dietary synbiotics reduce cancer risk factors in polypectomized and colon cancer patients. *Am. J. Clin. Nutr.* **85**, 488–496 (2007).
90. Gianotti, L. *et al.* A randomized double-blind trial on perioperative administration of probiotics in colorectal cancer patients. *World J Gastroenterol* **16**, 167–175 (2010).
91. Pala, V. *et al.* Yogurt consumption and risk of colorectal cancer in the Italian European prospective investigation into cancer and nutrition cohort. *Int. J. Cancer* **129**, 2712–2719 (2011).
92. Pamer, E. G. Fecal microbiota transplantation: effectiveness, complexities, and lingering concerns. *Mucosal Immunol* **7**, 210–214 (2014).
93. Hedayat, M., Takeda, K. & Rezaei, N. Prophylactic and therapeutic implications of toll-like receptor ligands. *Med Res Rev* **32**, 294–325 (2012).
94. Fűri, I. *et al.* Epithelial toll-like receptor 9 signaling in colorectal inflammation and cancer: Clinico-pathogenic aspects. *World J Gastroenterol* **19**, 4119–4126 (2013).

95. Kim, S. *et al.* Berberine suppresses the TPA-induced MMP-1 and MMP-9 expressions through the inhibition of PKC-α in breast cancer cells. *J. Surg. Res.* **176**, e21-29 (2012).
96. Zoglmeier, C. *et al.* CpG blocks immunosuppression by myeloid-derived suppressor cells in tumor-bearing mice. *Clin. Cancer Res.* **17**, 1765–1775 (2011).
97. Heckelsmiller, K. *et al.* Combined dendritic cell- and CpG oligonucleotide-based immune therapy cures large murine tumors that resist chemotherapy. *Eur. J. Immunol.* **32**, 3235–3245 (2002).
98. Westwood, J. A. *et al.* Toll-Like Receptor Triggering and T-Cell Costimulation Induce Potent Antitumor Immunity in Mice. *Clin. Cancer Res.* **15**, 7624–7633 (2009).
99. Heckelsmiller, K. *et al.* Peritumoral CpG DNA elicits a coordinated response of CD8 T cells and innate effectors to cure established tumors in a murine colon carcinoma model. *J. Immunol.* **169**, 3892–3899 (2002).
100. Kuo, W.-T., Lee, T.-C. & Yu, L. C.-H. Eritoran Suppresses Colon Cancer by Altering a Functional Balance in Toll-like Receptors That Bind Lipopolysaccharide. *Cancer Res.* **76**, 4684–4695 (2016).
101. Dovedi, S. J. *et al.* Intravenous administration of the selective toll-like receptor 7 agonist DSR-29133 leads to anti-tumor efficacy in murine solid tumor models which can be potentiated by combination with fractionated radiotherapy. *Oncotarget* **7**, 17035–17046 (2016).
102. Jiang, Q., Wei, H. & Tian, Z. Poly I:C enhances cycloheximide-induced apoptosis of tumor cells through TLR3 pathway. *BMC Cancer* **8**, 12 (2008).
103. Adams, S. Toll-like receptor agonists in cancer therapy. *Immunotherapy* **1**, 949–964 (2009).
104. Bhardwaj, N., Gnjatic, S. & Sawhney, N. B. TLR AGONISTS: Are They Good Adjuvants? *Cancer J* **16**, 382–391 (2010).
105. Uematsu, S. & Akira, S. Toll-Like receptors (TLRs) and their ligands. *Handb Exp Pharmacol* 1–20 (2008) doi:10.1007/978-3-540-72167-3_1.

2. Characterization of mucosal *Bacteroides fragilis* in patients with colorectal adenoma

Melissa Kordahi[1], Marion Avril[2], Denise Chac[2], Ben Ross[3], Christian Diener[4], Sumita Jain[5], Anika Parker[2], Jenny Huang[2], Wynn Burke[2], Sean Gibbons[4], Amy Willis[6], Richard Darveau[5], Bill Grady[2], Cynthia Ko[2] and R. William DePaolo[1,2,6]*

[1] Department of Pathology, University of Washington, Seattle, WA 98195
[2] Department of Medicine, University of Washington, Seattle WA 98195
[3] Department of Microbiology and Immunology at Dartmouth College, Hanover, NH 03755
[4] e-Science Institute, University of Washington, Seattle, WA 98195
[5] Department of Dentistry, University of Washington, Seattle WA 98195
[6] Department of Biostatistics, University of Washington, Seattle WA 98195
[7] Center for Microbiome Sciences & Therapeutics, University of Washington, Seattle WA 98195

* Corresponding author:
R. William DePaolo, PhD
Department of Medicine
University of Washington,
1959 NE Pacific Avenue,
Seattle, WA 98195
(312) 371-7828
Email: wdepaolo@medicine.washington.edu

KEYWORDS: Microbiome, colorectal adenoma, inflammation, innate immunity

Grant Support: Supported by UW start-up funds (RWD)

2.1. ABSTRACT

CRC is the third most common cancer worldwide, with approximately 1.5 million people diagnosed with the disease each year. More than 90% of CRC cases are sporadic and believed to be driven by environmental factors. Recently, research into the human gut microbiome suggests a role for "driver" bacteria in the development of CRC. One such bacterium is specific strains of *Bacteroides fragilis*, which are considered capable of remodeling the mucosal immune response and colonic bacterial community that promote oncogenic changes in the colon epithelium. In this study, we were interested in identifying microbiome signatures of early CRC development in patients undergoing routine CRC screening. Using a combination of culturomics, matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) and 16S rRNA gene sequencing, we noticed an enrichment of mucosal *Bacteroides fragilis* isolation in patients with proximal tubular adenoma and sessile serrated adenoma compared to patients without adenoma. High *Bacteroides* recovery was associated with larger lesions and higher levels of mucosal inflammation in patients with tubular adenoma. We then proceeded to show that *Bacteroides fragilis* isolates recovered from patients with adenoma displayed a TLR2-mediated pro-inflammatory immunogenic phenotype *in vitro*. At the genomic level, *Bacteroides fragilis* isolates from patients with tubular adenoma and sessile serrated adenoma contain a higher proportion of pro-inflammatory LPS-O antigen genes while *Bacteroides fragilis* isolates from patients with sessile serrated adenoma and Polyp-free individuals contained a higher proportion of tolerogenic capsular polysaccharides. *Bacteroides fragilis* isolates from patients with tubular adenoma and sessile serrated adenoma were also significantly proteomically distinct from the isolates recovered from Polyp Free individuals. Taken together, the results gathered through this study explore the relationship between *Bacteroides fragilis* and early-stage colorectal neoplasia and provide a biological framework for biomarkers and offer insights into microbiota-based therapies. (Add clearer take home message)

2.2. INTRODUCTION

Colorectal cancer (CRC) is the third most common type of cancer and a leading cause of cancer-related death in both men and women[1,2]. CRC has been firmly linked to chronic intestinal inflammation[3,4]. During intestinal inflammation, infiltrating immune cells are important sources of pro-tumorigenic cytokines and chemokines that can induce DNA damage and mutation[5]. CRC

is a disease that develops over the span of several years, starting with precursor lesions called polyps or adenomas[6]. Adenomas result from a disruption in the normal proliferation and apoptosis cycles of the epithelial lining of the colon[7]. The two main types of colorectal adenomas with a relatively high malignant potential are tubular adenoma and sessile serrated adenoma[8]. They differ in their appearance in the sense that tubular adenoma are protruding fleshy lesions whereas sessile serrated adenoma are flat hyperproliferative lesions[9]. These two types of adenoma also emerge through different molecular pathways and are associated with distinct molecular signatures[10]. However, we still ignore to this day what drives the development of tubular adenoma compared to sessile serrated adenoma in early stage CRC. We also ignore why less than 10% of patients with adenoma end up developing CRC whereas the rest of the population with adenoma does not[11]. CRC is also grouped into proximal and distal cancer[12]. Proximal colorectal cancers tend to present at a more advanced stage and have worse prognosis[13]. Thus, there is an additional need to understand factors associated with the development of proximal cancer in particular.

The intestinal tract is heavily colonized by about 10^{14} bacteria, referred to as the microbiota[14]. The intestinal microbiota is fundamental to human health and development and altered microbiomes have been associated with chronic diseases such as inflammatory bowel disease and CRC[15]. The microbiota is usually kept at a safe distance from the host intestinal epithelium by a physical barrier consisting of a thick layer of mucus[16]. CRC features microbiota penetrating the mucus layer and forming biofilms that adhere to the intestinal mucosa[17]. With *Helicobacter pylori* as the best example of an infectious bacteria driving gastric cancer development[18], recent publications have highlighted the importance played by select commensal bacteria in carcinogenesis[19]. For example, the enterotoxigenic strain of *Bacteroides fragilis* (ETBF) is a bacterium that has been linked to CRC development[20]. In mouse studies, ETBF has the ability to penetrate the intestinal mucus layer and form biofilms at the surface of the intestinal epithelium[21]. It can also express a toxin called *fragilysin* that binds to an unknown receptor of the intestinal epithelium and compromises the integrity of the epithelial tight junctions through the cleavage of E-cadherin[22]. This results in bacterial translocation, allowing the microorganism and its metabolites and secreted factors such as fragilysin to come in contact with immune cells and trigger an inflammatory response. Subacute chronic inflammation in the colon can lead to onco-transformation over the span of several years, especially in susceptible individuals .

Simultaneously, *Bacteroides fragilis* is a ubiquitous and important Gram-negative anaerobe that colonizes the human lower gastrointestinal tract[23]. *Bacteroides* species are among the earliest

colonizers of the gut microflora. *B. fragilis* features eight distinct surface capsular polysaccharides (PSA) that are believed to play a role in immune homeostasis in the gut[24]. Indeed, *B. fragilis* PSA normally mediates the establishment of a proper Pro-inflammatory/Anti-inflammatory balance in the host, which is a fundamental aspect of a healthy immunologic function[25]. *B. fragilis* PSA exerts its immunomodulatory role through activation of the TLR pathway, inducing mucosal tolerance[26]. Therefore, unlike TLR ligands of pathogens which lead to clearance of infections, PSA actually promote microbial colonization through TLR signaling.

In this study we looked for specific microbial signatures associated with tubular adenoma and sessile serrated adenoma in the proximal colon in a group of patients undergoing CRC screening and found a high relative abundance of mucosal *Bacteroides fragilis* recovered *in vitro* from patients with adenoma. These mucosal *Bacteroides fragilis* isolates were correlated to larger adenoma in patients and to mucosal inflammation. We then characterized *Bacteroides fragilis* isolates in order to better understand how they can benefit the host as a gut commensal in healthy mucosa and contribute to an inflammatory phenotype in patients with pre-cancerous lesions in the colon.

2.3. MATERIALS AND METHODS

Patients recruitment and Microbiome culture:

At the Gastroenterology department of the University of Washington, in Seattle Washington, 40 patients were enrolled after obtaining informed consent between May 2017 and September 2018 (IRB # 34095A). Exclusion criteria included age <50, BMI > 30 kg/m$_2$ use of antibiotics within the past 3 months, active inflammation, and inability to sign informed consent. All colon mucosa samples were collected by designated gastroenterologist Dr. Cynthia Ko. Polyp associated mucosa (PA) and healthy non-polyp associated tissue (NPA) 10 cm further away from where the polyp was found were obtained from the same subject collected by endoscopy biopsy forceps, generating a total of 240 samples (mean age of subjects: 64 ± 9 years old). All collected biopsies were placed in cryovials (Thermo Fisher Scientific Inc., USA) filled with anaerobic media (Anaerobic systems, USA). Following immediate and thorough homogenization with a homogenizing pestle (Thermo Fisher Scientific Inc., USA) in an anaerobic chamber and under sterile conditions, samples were plated on non-selective Tryptic Soy Agar containing 5% sheep blood (Anaerobic systems, USA) for bacterial growth. The plates were kept under anaerobic conditions (90% N2, 5% CO2, 5% H2)

at 37°C for 2 days. Colonies cultured for 48 hours were expanded in TSB + 5% sheep blood after identification, then transferred to skim milk stock (Thermo Fisher Scientific Inc., USA), and preserved at -80°C until analyzed.

MALDI-TOF

Bacterial Test Standard (BTS) was used for external calibration of MALDI-TOF MS. For sample analysis, individual colonies of clinical isolates were individually transferred from agar plate cultures to a MALDI target plate (Bruker Daltonics, Germany), using a sterile applicator (Puritan, USA) then air dried. Samples were then treated with 1 μL of a 70% formic acid solution (Sigma) for protein extraction, then air dried again. Samples were also treated with 1 μL of a-cyano-4-hydroxycinnamic acid matrix solution (a-CHCA, Bruker Daltonics, Germany), air dried, and analyzed in Microflex LT MALDI-TOF mass spectrometer (Bruker Daltonics, Germany). Linear positive-ion mode, N 2 laser, λ = 355 nm, pulse duration: 150 ns, laser frequency: 200 Hz. Spectra were recorded over the range m/z 2000-20,000.

16S rRNA Sequencing on select biopsy homogenates samples

Library preparation and sequencing 16S sequencing libraries prepared by QIAseq 16S/ITS screening panel (QIAGEN). Extracted DNA were diluted into 1ng/uL. Seven 16S regions (V1V2, V2V3, V3V4, V4V5, V5V7, V7V9 and ITS) were captured and amplified in three different primer panel pools for each sample, according to manufacturer's protocol. For each of the primer panel pool, 2uL of the diluted DNA was used as input and incubated at 95°C for 2 minutes, then amplified for 20 PCR cycles of 95°C for 30s, 50°C for 30s and 72°C for 2 minutes, with a final extension at 72°C for 7 minutes using an UCP Multiplex PCR kit (QIAGEN). Reactions from the same samples were pooled and the pooled intermediate products was cleaned up twice using 1.1X QIAseq beads (QIAGEN) to remove unused PCR primers. A final PCR reaction was done using UCP Multiplex PCR kit (QIAGEN) and QIAseq 16S/ITS indices (QIAGEN) to incorporate sample indices and sequencing adapters. The reaction mix was incubated at 95°C for 2 minutes and 14 cycles of 95°C for 30s, 50°C for 30s and 72°C for 2 minutes, with a final extension at 72°C for 7 minutes. The final PCR product was cleaned up using 0.9X QIAseq beads (QIAGEN), and the cleaned-up libraries were inspected on a Agilent Tapestation (Agilent). The 16S libraries were sequenced on an Illumina MiSeq for 1.5 M 2x276bp reads per sample

Data Analysis 16S:

16S screening panel sequencing data analysis was carried out using the QIAGEN CLC Genomics Workbench (version 11.0) and the Data QC and OTU Clustering workflow from the Microbial Genomics Pro Suite Module (version 4.0). Briefly, raw reads were demultiplexed and grouped into different 16S/ITS regions (e.g., V1V2, V3V4, and ITS) using the QIAseq 16S/ITS Demultiplexer tool which classified reads into different regions using the phased 16S primer sequences. Operational taxonomic clustering for each 16S/ITS region was performed using the OTU Clustering tool, where the demultiplexed reads are aligned against the SILVA 16S database using a similarity percentage parameter at 97%, to create an alignment score for each OTU.

Next Generation Sequencing:

5mL TSB + 5% sheep blood liquid cultures (Fisher scientific, USA) from stock of clinical isolates were prepared and incubated under anaerobic conditions (90% N2, 5% CO2, 5% H2) at 37°C for 48 hrs. 1 mL of each isolate's liquid culture was transferred to an Eppendorf tube and centrifuged at 13000 rpm for 60 seconds to collect bacteria. Total bacterial genome DNA of *Bacteroides fragilis* isolates was extracted, referring to steps in the instructions of the bacterial DNA extraction kit (Puregene yeast/Bact kit B, Qiagen, USA). Extracted DNA samples were then quantified using a Qubit fluorometer (Thermo Fisher Scientific Inc., USA) and library prepped using Illumina's Nextera DNA Flex Library Prep kits and protocol. Only samples with a minimum of 100 ng DNA inputs were included. Pooled libraries were then quantified using the Qubit and diluted with Resuspension Buffer (Illumina, USA) to 2 nM, then loaded on a NextSeq for NGS.

Preprocessing and assembly of isolate sequencing data

Preprocessing and assembly were performed using existing Anvi'o pipelines (Li et al. 2015; Eren et al. 2015). In brief, the raw FASTQ files received from sequencing were first subjected to quality filtering using the "illumina-utils" package (Murat Eren et al. 2013). This was followed by de novo assembly of filtered reads into contigs for each separate sample using MEGAHIT with a minimum contig size of 1250 bps (Li et al. 2015). For each sample the filtered reads were then realigned to the assembled contigs to yield coverage profiles which were used together with the assembled contigs to create the Anvi'o contig and profile databases. Gene-calling was performed by identifying open reading frames with Prodigal (Hyatt et al. 2010) and basic functional annotation was performed by using the hidden markov models included with Anvi'o's "run-hmms" function.

Additionally, taxonomy was assigned to individual genes using Centrifuge (Kim et al. 2016). All of the processed data was then used to manually refine the assembled contigs in the Anvi'o interactive interface ("anvi-interactive") by selecting contigs with homogeneous coverage and taxonomy. In particular, we removed contigs which were classified as having bovine or ovine origin as those likely came from the growth media. Completeness and redundancy for all assemblies was quantified by the presence of taxon-specific single copy core genes (SCGs). All manual assemblies showed larger than 99% completeness and less than 5% of redundancy/contamination. The resulting curated genomes and plasmids were stored in the Anvio'o contig databases.

Protein extraction method of *B. fragilis* isolates

Single *B. fragilis* isolates colonies grown on TSA + 5% sheep blood plates were selected and added to 300 ul of HPLC-grade water in a 1.5 ml Eppendorf tube and then mixed thoroughly with 900 µl of 100% Ethanol. After centrifugation at 13000 rpm for 2 min twice, pellets were dried at room temperature for 5 minutes then they were directly mixed with equal volumes of 70% formic acid and acetonitrile (20-40 ul, depending on pellet size) for protein extraction. After centrifugation at 13000 rpm for 2 minutes, 1 µl of protein extract was spotted on a 96-target polished steel plate (Bruker Daltonics, Bremen, Germany) in four technical replicates, air-dried, and overlaid with 1 µl of matrix solution (Bruker Daltonics, Bremen, Germany). Target plate was placed in the MALDI TOF Biotyper for microbial identification (Bruker, Germany).

MALDI-ToF Data analysis

Raw spectra text files were analyzed using the R package, MALDIquant [https://www.ncbi.nlm.nih.gov/pubmed/22796955]. The raw data were trimmed to a spectra range of 3,000 to 15,000 m/z. The spectra intensities were then square-root transformed and smoothed using the Savitzky-Golay algorithm. Baseline noise was removed using the statistics-sensitive non-linear iterative peak clipping (or SNIP) algorithm with 100 iterations. The data were then normalized using total ion current (or TIC) calibration, which sets the total intensity to 1. Multiple spectra within the same analysis were aligned to the same x-axis using the Lowess warping method, a signal-to-noise ratio of 3, and a tolerance of 0.001. Peaks were detected from the average of at least 4 technical replicates using median absolute deviation. Principal components analyses

and hierarchical clustering were also performed in R using the base stats package. Hierarchical clustering was performed on a calculated Euclidean distance matrix using Ward's method.

Biofilm formation using crystal violet formation assay

5mL TSB + 5% sheep blood liquid cultures from stock of clinical isolates were prepared and incubated under anaerobic conditions (90% N2, 5% CO2, 5% H2) at 37°C for 48 hrs.

1:100 dilutions were prepared, with a total volume of 1 mL, with TSB + 5% sheep blood for each liquid culture. 100 µL of each dilution was plated in sets of 4 wells in a round-bottom 96 well plate. The 96-well plate was covered and incubated anaerobically at 37°C for 48 hours. Then, the plate was shaken out over a tray to remove all planktonic bacteria.

The 96-well plate was rinsed in a large beaker of water and the water shaken out over the tray. All wells were stained with 125 µL of 0.1% crystal violet for 10 minutes then rinsed out then left to dry out overnight. 200 µL of 30% acetic acid was added to all wells that were stained to solubilize the crystal violet and allowed to sit for 10 minutes, then 125 uL of the acetic acid/crystal violet solution was transferred from each well into a well in a flat-bottom 96-well plate and read at OD 450 nm.

Preparation of biopsy samples for cytokines analysis:

Extraction of proteins from mechanically homogenized biopsy samples was achieved by adding RIPA lysis buffer and protease-inhibitors. The samples were then incubated on ice for 30 min with occasional shaking. The insoluble components were removed via centrifugation at 15 000 g for 30 min at 4C.The protein concentration was measured using the Bradford assay. Samples were normalized accordingly before cytokine quantification.

Co-culture cell assays

5mL of chopped meat cultures from skim milk stocks of clinical isolates were prepared and incubated under anaerobic conditions (90% N2, 5% CO2, 5% H2) at 37°C for 48 hrs. 1 mL of each isolate liquid culture was transferred to an Eppendorf tube and centrifuged at 13000 rpm for 30 min to pellet bacteria and collect secreted factors in bacterial supernatant. 100 000 cells from THP1 human cell lines (ATCC, USA) were seeded in a 96 well plate then co-cultured with resuspended and normalized isolates cultures at a concentration of 1:5. The cells were then incubated at 37°C for 24 hrs for the THP1 cell line. When removed from the incubator, the plates were centrifuged and transferred at -20°C until they were used for ELISA assays.

Cytokine quantification by ELISA

50 μL ELISA Diluent was added to each well followed by 100 μL of standard or sample and incubated 2 hours at room temperature. The content of the wells was the aspired and washed 5 times. 100 μL of prepared Working Detector was added to each well and the wells were incubated 1 hour at room temperature. The content of the wells was the aspired and washed 7 times. Then 100 μL TMB One-Step Substrate Reagent was added to each well and incubated for 30 minutes at room temperature. 50 μL Stop Solution was finally added to each well and the plate was read at OD 450 nm. (IL6, IL-8, TNF alpha, IL-12P40 Human ELISA kits BD OptEIA, USA).

TLR2 AND TLR4 Assays

HEK293 cells were plated in 96-well plates and transfected the following day with plasmids encoding human TLRs, NF-kB-dependent firefly luciferase reporter, and B-actin promoter-dependent *Renilla* luciferase reporter. In the case of human TLR4, 0.002 ug plasmid encoding human TLR4 was co-transfected with 0.0025 ug plasmid encoding human MD-2. In the case of human TLR2, 0.001 ug plasmid encoding TLR2 was co-transfected with 0.002 ug plasmid encoding human mCD14. At 18 to 20 h post-transfection, test wells were stimulated in duplicatesfor 4 h at 37°C with *B. fragilis* supernatants, which were suspended in Dulbecco's modified Eagle medium (DMEM) containing 10% human serum. Luciferase activity was assayed using a dual-luciferase assay reporter system (Promega, Madison, WI). NF-B activity was measured as the ratio of NF-B-dependent firefly luciferase activity to B-actin promoter-dependent *Renilla* luciferase activity, which served as an internal standard. The data were plotted as the fold difference between the NF-kB activity of the sample and that of the unstimulated control.

Statistical analyses

Data are expressed either as the mean value + standard error of the mean (s.e.m.) or as individual values. Specific statistical tests used for each experiment are described in the figure legends and performed in Graphpad PRISM (Graphpad Software LLC, San Diego CA). $p < 0.05$ was considered significantly different. For functional assays, One-way Anova with multiple comparisons was performed with the Mean of each group compared to the Mean of every other group. All functional analysis was a repeat of 2-4 experiments with pooled data.

Data Deposition

The 16S rRNA profiling data and WGS data sequenced in this study will be deposited in online databases under accession number.

2.4. RESULTS

Patients and Pathological data

A total of 40 patients were included in this study. From those, histological analyses revealed that 18 patients had tubular adenoma (TA), 7 patients had sessile serrated adenoma (SSA), 4 patients had hyperplastic polyps (HP) and 2 patients had lesions with no diagnostic alterations. The remaining 9 patients recruited had no lesions in the colon at time of the colonoscopy and served as a healthy polyp-free control group (PF). Of the patients with polyps, two biopsies were collected from each patient, one corresponding to the polyp-associated mucosa (PA) and one to healthy mucosa located 10 cm further away and termed non-polyp associated mucosa (NPA). Patient characteristics and details of colonoscopy treatment performed as well as information about colonic lesions are outlined in Table 2.1 and Table 2.2.

	Polyp Free	Polyp/Adenoma	p-value
Patient sample size *	9 (22.5)	31 (77.5)	
Age in years (Mean ± SD)	61 ± 8	64 ± 9	0.341
Gender (%)			
Female	6 (67)	19 (61)	
Male	3 (33)	12 (38.7)	
BMI (Mean ± SD)	24 ± 4	25 ± 3	0.307
Formal smoker (%)	2 (22.2)	10 (32.3)	0.544
Positive family history (%)	3 (33.3)	10 (32.3)	1
Anemia (%)	4 (44.4)	16 (51.6)	0.714

*Exclusion criteria: Age < 50, BMI > 30, antibiotic use in the last 3 months, IBD, Diabetes Mellitus.

Table 2.1. Study participants demographics

	N (%)	Size of lesions (cm) Mean (range)
Tubular Adenoma (TA)	16 (51.6)	0.55 (0.10-1.00)
Sessile Serrated Adenoma (SSA)	7 (22.6)	0.59 (0.20-0.90)
Hyperplastic Polyp (HP)	4 (12.9)	0.32 (0.2-1.0)
No Diagnostic Alterations **	4 (12.9)	0.28 (0.20-0.40)

Table 2.2. Lesions characteristics

Cataloguing of different polyp-associated microbiota and comparison to healthy mucosa microbiota

Since 80% of the most aggressive lesions (TA and SSA) were located in the proximal colon (Supplemental Diagram 2.1), we decided to focus on the patients with lesions found in the proximal colon. According to the NCI Dictionary of Cancer Terms, the proximal colon refers to the first and middle parts of the colon. The proximal colon includes the cecum (a pouch that connects the small intestine to the colon), the ascending colon (the right side of the colon), and the transverse colon (the part of the colon that goes across the body between the right and left sides of the colon). Consequently, thirteen Polyp Associated mucosa (PA) and corresponding healthy Non-Polyp Associated tissue (NPA) samples from patients with tubular adenoma (TA), five Polyp Associated mucosa (PA) and corresponding healthy Non-Polyp Associated tissue (NPA) samples from patients with sessile serrated adenoma (SSA), as well as nine healthy Polyp-free tissue (PF) (i.e. total of 45 samples), retrieved from 27 subjects undergoing colonoscopy for routine colorectal cancer screening were used to evaluate the microbiota taxonomic profile by means of culturomics.

Enrichment for Bacteroides fragilis *in vitro* in patients with colorectal adenoma

A total of 1550 colonies were tested from the proximal colon and identified by MALDI-TOF MS with a mean ± SD of 34 ± 40 (ranging from 1 to 215) per patient and 83 different bacterial species were identified (Figure S1). The bacterial species belonged to Proteobacteria (n = 619 different species, 37.3%), followed by Bacteroidetes (n = 536 , 32.3%), Firmicutes (n= 312, 18.8%), Actinobacteria (n = 82, 4.9%), and Fusobacteria (n = 2, 0.1%). These isolates belonged to 32 different genera: *Escherichia* was the most common (n = 597, 38.5%), followed by *Bacteroides* (n = 525 , 33.9%), *Clostridia* (n = 136, 8.8%), *Collinsella* (n = 70, 4.5%), *Ruminococcus* (n= 51,

3.3%), and *Flavonifractor* (n=43, 2.8%). All the bacterial species isolated and identified were bacteria previously described in the human gut.

When evaluating the microbiota taxonomic profiles according to the different types of mucosal tissues the bacteria were isolated from, we noticed that the average proportion of bacterial species isolated from PA and corresponding NPA samples from patients with TA was compositionally identical with a predominance of Bacteroidetes (average of 60% ± 31.3% and 59% ± 32% respectively), followed by Proteobacteria (21.4% ± 34.7% and 22.0% ± 33.2% respectively), Firmicutes (11.2% ± 15.3% and 12.0% ± 21.4%), and Actinobacteria (6.8% ± 16.5% and 8.0% ± 15.4%), with a small proportion of Fusobacteria in the PA samples only (0.5% ± 1.7%). The average proportion of bacterial species isolated from PA and corresponding NPA samples from patients with SSA was also compositionally similar with an almost equal predominance of Bacteroidetes (38.4% ± 26.7% and 32.2% ± 30.4% respectively) and Proteobacteria (37.4% ± 38.3% and 35.4% ± 35.3%) followed by Firmicutes (19.6% ± 15.0% and 17.0 ± 11.2 %) and Actinobacteria (4.6% ± 5.3% and 15.2% ± 12.3%).When looking at the average proportion of bacterial species isolated from PF healthy mucosal biopsies, we noticed that they resembled more the average proportion of bacterial species isolated from PA and corresponding NPA samples from patients with SSA than the average proportion of bacterial species isolated from PA and corresponding NPA samples from patients with TA with a Predominance of Proteobacteria (36.6% ± 42.6%) very closely followed by Bacteroidetes (31.0% ± 30.3%), Firmicutes (27.4% ± 29.1%) and Actinobacteria (4.8% ± 7.2%) (Figure 2.1 A). When performing a statistical analysis using AMOVA at the phylum level of the relative abundances of the bacterial species recovered from the various types of biopsies, there were no statistically significant differences between the bacterial populations recovered from the healthy PF mucosa and the PA and corresponding NPA samples from patients with SSA (p-value: 0.86). However, the bacterial species recovered from the PA and corresponding NPA samples from patients with TA were significantly different from the bacterial populations recovered from the healthy PF mucosa and from both the PA and corresponding NPA samples from patients with SSA with a significant increase in Bacteroides (relative percentage difference of 28.5%) and a significant decrease in Proteobacteria (relative percentage difference of 14.8%) between the averaged relative abundances for each Phylum of bacterial species isolated from PA and NPA TA samples compared to PF healthy mucosal samples (p-value: 0.001). The bacterial composition of bacterial species isolated from averaged PA and NPA TA samples was also significantly distinct from the bacterial species isolated from averaged

PA and NPA SSA samples with another significant increase in Bacteroides (relative percentage difference of 24.2%) and a decrease in Proteobacteria (relative percentage difference of 14.6%) (p-value: 0.014). (Figure 2.1 B).

Furthermore, at the genus level, we identified a total of 23 genera of which Bacteroides and Escherichia were the most predominant, representing on average 43.7% ± 32.1% and 26% ± 34.5% of the total genera abundance respectively. When trying to identify the differences in microbiota recovery between TA, SSA, and PF patients, we noticed that TA patients had on average a significantly larger predominance of Bacteroides (53.8% ± 32.1%) compared to patients with SSA (31.5% ± 28.1%) and PF patients (27.9% ± 27.3%). On the other hand, the most predominant genera in both SSA patients and PF patients was Escherichia with a comparable prevalence of 32.4% ± 35.9% and 36.3% ± 42.5% respectively vs a prevalence of 20.0% ± 30.9% only in TA patients. This resulted in a significantly inverted correlative relationship between Bacteroides and Escherichia across our sample population suggesting a potential competition between the two genera when cultured together (Fig). The next most prevalent genera in PF patients were Clostridia (10.0% ± 23.1%) which are known as leading commensal bacteria in the maintenance of gut homeostasis and the promotion of gut health (REF). On the other hand, the next leading genera recovered from TA and SSA patients were Collinsella (7.2% ± 15.6% and 8.3% ± 9.3% respectively) which have been associated with low dietary fiber intake in the general population and an alteration of the overall fermentation pattern in gut microbiota (Figure 2.1 C).

In order to validate whether what we identified by means of Culturomics was reflective of the microbiota composition in our samples, we selected a subset of homogenized biopsies for 16S rRNA profiling. We selected four homogenized Polyp Associated mucosa (PA) and corresponding healthy Non-Polyp Associated tissue (NPA) samples from patients with tubular adenoma (TA), three Polyp Associated mucosa (PA) and corresponding healthy Non-Polyp Associated tissue (NPA) samples from patients with sessile serrated adenoma (SSA), as well as 5 healthy Polyp-free tissue (PF) (i.e. total of 23 samples), retrieved from 12 subjects (Supplemental figure 2.1)

The bacterial species identified by 16S belonged to Firmicutes (56.8% ± 21.5%), followed by Proteobacteria (22.6% ± 26.0%), Bacteroides (16.7% ± 13.3%), Actinobacteria (2.2% ± 1.8%), and Fusobacteria (1.6% ± 3.2%) (Table). These isolates belonged to 35 different genera of which 20 were also identified by culturomics. Blautia was the most prevalent (15.6% ± 17.9%), followed by Peptoniphilus (12.4% ± 9.5%), Bacteroides (9.8% ± 13.0%), and Escherichia (7.3% ± 18.6%).

When evaluating the microbiota taxonomic profiles according to the different types of mucosal tissues the bacteria were isolated from, we noticed that the average proportion of bacterial species isolated from PA and corresponding NPA samples from patients with TA was compositionally comparable with a predominance of Firmicutes (average of 54.9% ± 25.1% and 50.5% ± 17.3% respectively), followed by Proteobacteria (30.9% ± 28.2% and 22.7% ± 25.8% respectively). Bacteroides was less prevalent in the TA PA samples compared to the TA NPA samples (6.9% ± 2.5% and 24.5% ± 18.8%). On the other hand, Actinobacteria and Fusobacteria were more prevalent in the TA PA samples compared to the TA NPA samples (3.0% ± 1.9% and 1.5% ± 1.7% respectively for Actinobacteria vs 4.2% ± 6.4% and 0.9% ± 1.0% for Fusobacteria).

The average proportion of bacterial species isolated from PA and corresponding NPA samples from patients with SSA was also compositionally similar with a comparable predominance of Firmicutes (51.7% ± 22.3% and 44.9 ± 30.0 %), Proteobacteria (31.1% ± 24.5% and 33.8% ± 44.0%), and Bacteroidetes (14.3% ± 3.4% and 19.0% ± 16.4% respectively) followed by Actinobacteria (2.4% ± 1.5% and 2.2% ± 2.4%) and Fusobacteria (0.5% ± 0.2% and 0.2% ± 0.3%).

When looking at the average proportion of bacterial species isolated from PF healthy mucosal biopsies, we noticed that they were distinct from the average proportion of bacterial species isolated from PA and corresponding NPA samples from patients with SSA and TA with a significantly less Proteobacteria (5.8% ± 5.9%) and more Firmicutes (76.3% ± 10.0%) in the PF samples. A shift in Proteobacteria/Firmicutes could be indicative of a diseased microbiome in the patients with TA and SSA lesions in the colon (Figure 2.1 D). When performing a statistical analysis using AMOVA at the phylum level of the relative abundances of the bacterial species recovered from the various types of biopsies, there were no statistically significant differences between the bacterial populations recovered from the PA and corresponding NPA samples from patients with TA and SSA (p-value: 0.252). However, the bacterial species recovered from the healthy PF mucosa were significantly different from the PA and corresponding NPA samples from patients with both TA and SSA. There was a significant increase in Firmicutes (relative percentage difference of 23.6%) and a significant decrease in Proteobacteria (relative percentage difference of 21.0%) between the averaged relative abundances for each Phylum of bacterial species isolated from PA and NPA TA samples compared to PF healthy mucosal samples (p-value: 0.015 HOMOVA). There was another significant increase in Firmicutes (relative percentage difference 30.2%) and a significant decrease in Proteobacteria (relative percentage difference of 25.3%) between the averaged relative abundances for each Phylum of bacterial species isolated

from PA and NPA SSA samples compared to PF healthy mucosal samples (p-value: 0.003 HOMOVA) (Figure 2.1 E).

When comparing the ratio of bacterial species recovered by Culturomics to 16S profiling we notice a clear selection for *Bacteroides in vitro* (~4-fold difference) while the other phyla were relatively comparable in predominance. We decided then to focus on the *Bacteroides* isolated *in vitro* in order to see if a particular species predominated (Figure 2.1 F). Interestingly, we found that Bacteroides species isolated from PA and corresponding NPA samples from patients with TA were largely predominated by *Bacteroides fragilis* with proportions of 81% and 80% respectively. Bacteroides species isolated from PA and corresponding NPA samples from patients with SSA were also largely predominated by *Bacteroides fragilis* but to a lesser extent than in patients with TA with proportions of 56% and 66% respectively. On the other hand, Bacteroides recovered from patients who were Polyp-free were not dominated by a particular species. Indeed, there a had a relatively more balanced recovery of *Bacteroides fragilis* (27%), *Bacteroides ovatus* (25%), and *Bacteroides uniformis* (20%) (Figure 2.1 G)

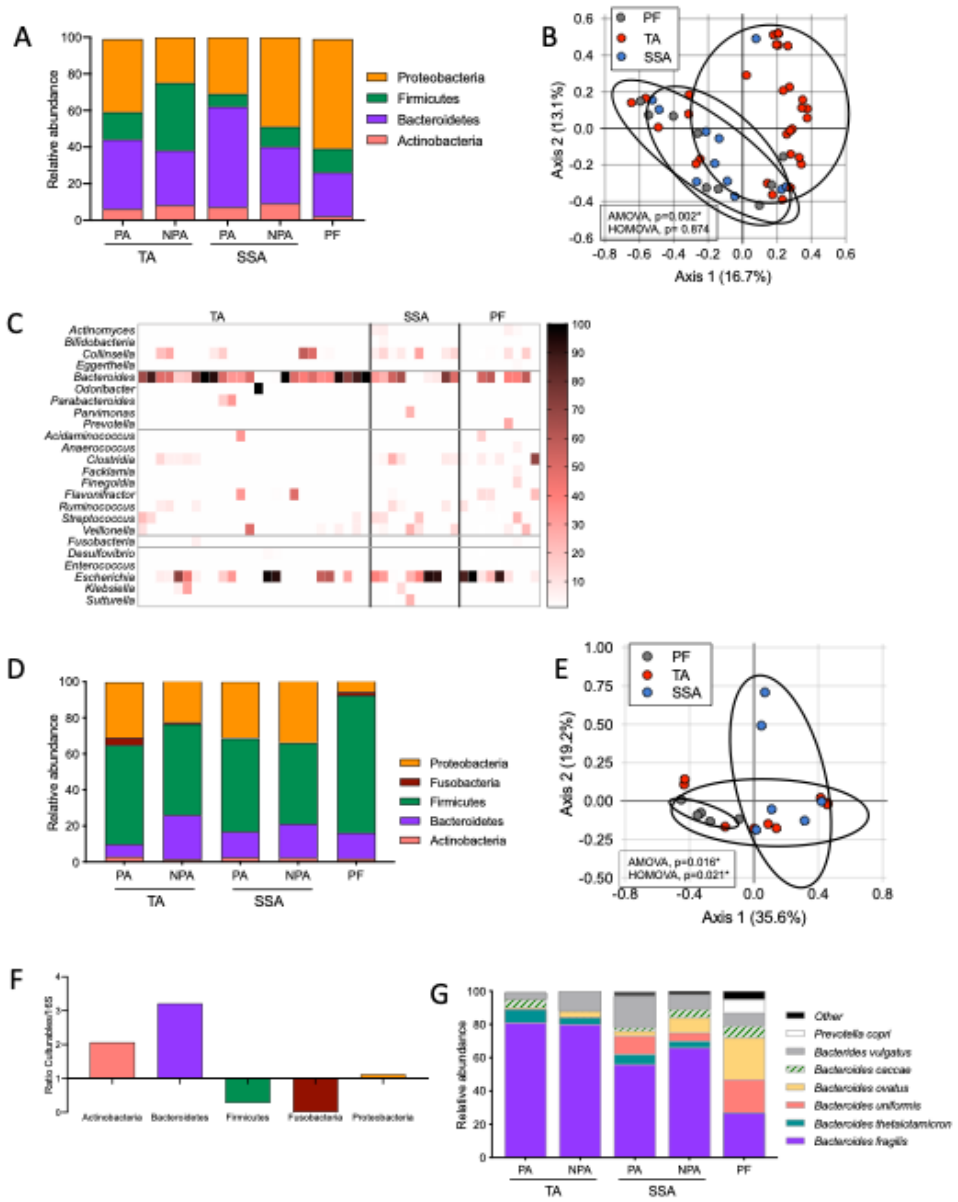


Figure 2.1. Enrichment for Bacteroides fragilis in vitro in patients with colorectal adenoma

Microbiota sampling of mucosal biopsies in TA PA/NPA, SSA PA/NPA, and PF biopsies. (A) Phylum-level abundance in culturomics. (B) Principle Coordinate of Analysis (PCoA) of microbiota composition by culturomics at phylum level. (C) Heatmap representation of genus level abundance (D) Phylum-level abundance in 16S rRNA sequencing (E) Principle Coordinate of Analysis (PCoA) of microbiota composition by 16S at phylum level (F) Ratio of total phyla recovered in culturomics relative to total phyla identified by 16S (G) Species level abundance of Bacteroides species in culturomics. Statistics in (B) and (E) are analysis of molecular variation (AMOVA) and homogeneity of molecular variance (HOMOVA).

NTBF *B. fragilis* recovery correlates with increased polyp size and host tissue inflammation in patients with tubular adenoma

After having characterized the microbiota composition of the biopsy samples by culturomics and 16S and because of the relationship of *Bacteroides fragilis* with CRC in the literature, we decided to assess the inflammatory status of the host tissue and try to correlate it to the important recovery of Bacteroides found in certain samples.

We first found that in patients with tubular adenoma, the abundance of *Bacteroides* recovered from the PA and NPA biopsy samples was positively correlated to an increase in adenoma size ($p=0.0191$) (Figure 2.2 A). adenoma size being a measure of severity of the lesion; the larger a polyp, the higher its propensity to become malignant. However, this correlation was not significant in patient with sessile serrated adenoma.

In the TA biopsy samples (both PA and NPA), we also found a positive correlation between the abundance of Bacteroides recovered and the inflammatory cytokines IL-12P40 ($p=0.0324$) and IL-8 ($p=0.0233$) (Figure 2.2 B and C). On the other hand, there was a significant negative correlation between the abundance of Bacteroides in the TA biopsy samples and the anti-inflammatory cytokine IL-10 (Figure 2.2 D). These correlations were not found to be significant in the SSA biopsy samples neither (Supplemental Figure 2.2).

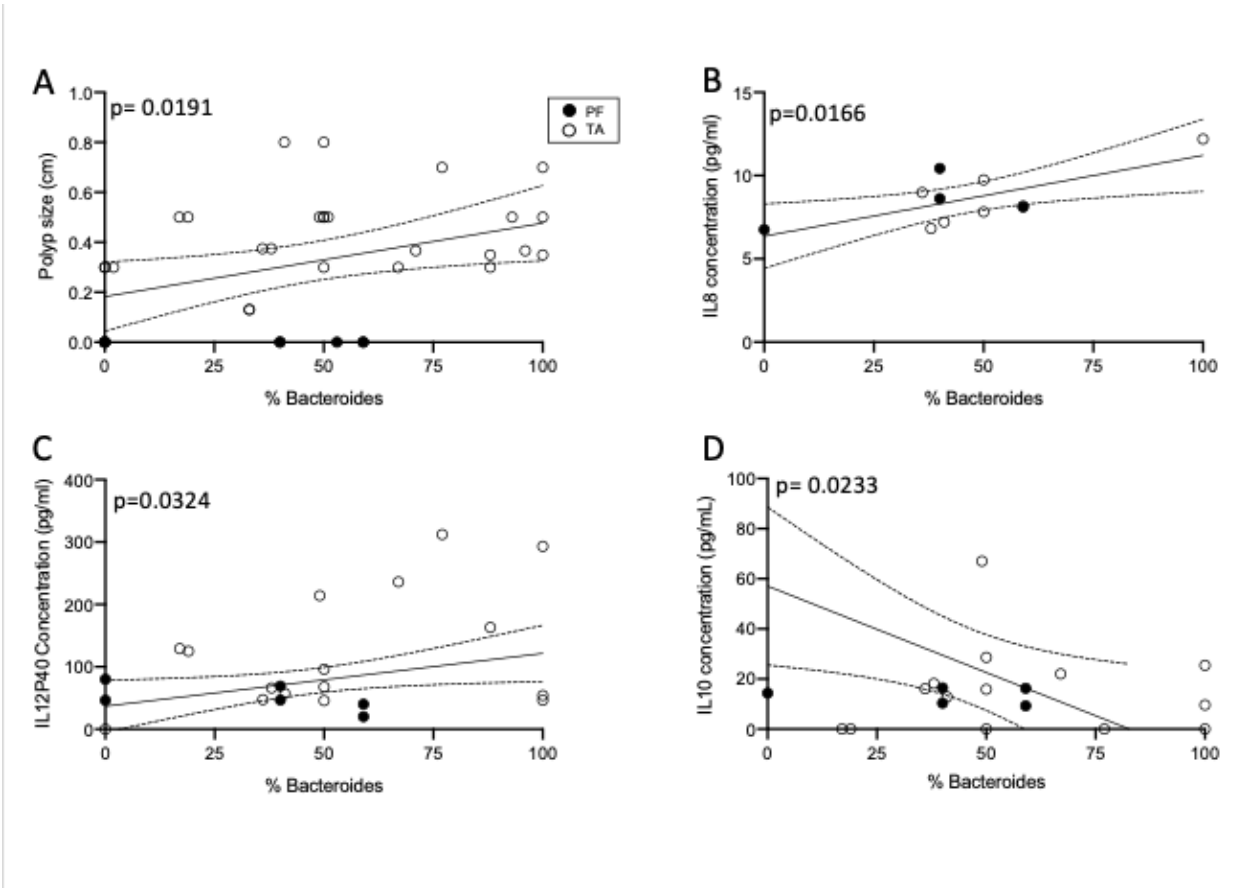


Figure 2.2. NTBF B. fragilis recovery correlates with increased polyp size and host tissue inflammation in patients with tubular adenoma

Correlation between (A) Polyp size in cm and abundance of *Bacteroides* isolated *in vitro* (B) Mucosal IL-8 concentration in pg/mL and abundance of *Bacteroides* isolated *in vitro* (C) Mucosal IL-12p40 concentration in pg/mL and abundance of *Bacteroides* isolated *in vitro* (D) Mucosal IL-10 concentration in pg/mL and abundance of *Bacteroides* isolated *in vitro* Data is mean ± s.e.m of 3 independent experiments. Statistics are Pearson Correlation Coefficient of all data points.

NTBF *B. fragilis* isolates from patients with tubular adenoma and sessile serrated adenoma are immunogenetically distinct from isolates from Polyp Free individuals

Next we decided to characterize the immunogenic potential of our *Bacteroides fragilis* in order to see whether it confirmed the correlations found earlier in the host tissue between *Bacteroides* abundance and inflammatory and anti-inflammatory cytokines. In order to assess that, we co-cocultured our isolates with THP1 monocytic cells in order to measure IL-12P40, IL-8, and IL-10 production. IL-1ß response of THP1 cells was greater with ETBF isolates(221.6 ± 28.8 pg ILB/mL at 24 hours, n = 56) and NTBF isolates from patients with TA PA and NPA (506.3 ± 52.4 pg ILB/mL at 24 hours, n = 56) and (359.8 ± 37.0 pg ILB/mL at 24 hours, n = 56) and SSA PA and NPA isolates (418.1 ± 86.6 pg/mL, n = 26) and (603.3 ± 68.5 pg/mL, n = 28) than with PF

isolates (15.6 ± 9.9 pg/mL, n = 14), P < 0.0001) (Figure 2.3 A). IL-8 response of THP1 cells was also greater with ETBF isolates (25667.1 ± 1499.2] pg IL-8/mL at 24 hours, n = 56) and NTBF isolates from patients with TA PA (25789.4 ± 1164.9] pg IL-8/mL at 24 hours, n = 62) and TA NPA (32287.5 ± 1232.9] pg IL-8/mL at 24 hours, n = 50) as well as in SSA PA and NPA isolates (17197.3 ± 1242.9 pg/mL, n = 26) and (27570.0 ± 1403.6 pg/mL, n = 26) respectively than with PF isolates (4555.1 ± 530.4 pg/mL, n = 13), P < 0.0001) (Figure 2.3 B). IL-12p40 response of THP1 cells was greater with ETBF isolates([257.6 ± 72.1] pg IL-12p40/mL at 24 hours, n = 48) and NTBF isolates from patients with TA PA([357.3 ± 62.5] pg IL-12p40/mL at 24 hours, n = 48) and SSA PA isolates (550.5 ± 91.9 pg/mL, n = 25) than with PF isolates (0 ± 0 pg/mL, n = 12), P < 0.0001) (Figure 2.3 C). IL-12P40 is a cytokine which correlates with pro-inflammatory TH17 T cell immune response. IL-10 response of THP1 cells was also greater with ETBF isolates(104.7 ± 53.0] pg IL-10/mL at 24 hours, n = 50) and NTBF isolates from patients with TA PA and NPA (139.9 ± 16.8] pg IL-10/mL at 24 hours, n = 72) and (116.5 ± 12.5] pg IL-10/mL at 24 hours, n = 72) and SSA NPA isolates (149.2 ± 25.8 pg/mL, n = 24) compared with PF isolates (35. 4 ± 13.5 pg/mL, n = 36), P < 0.0001). However, PF *B. fragilis* were at least able to induce IL-10 formation unlike any of the other pro-inflammatory cytokines, which is an expected tolerogenic response from a gut commensal (Figure 2. 3 D).

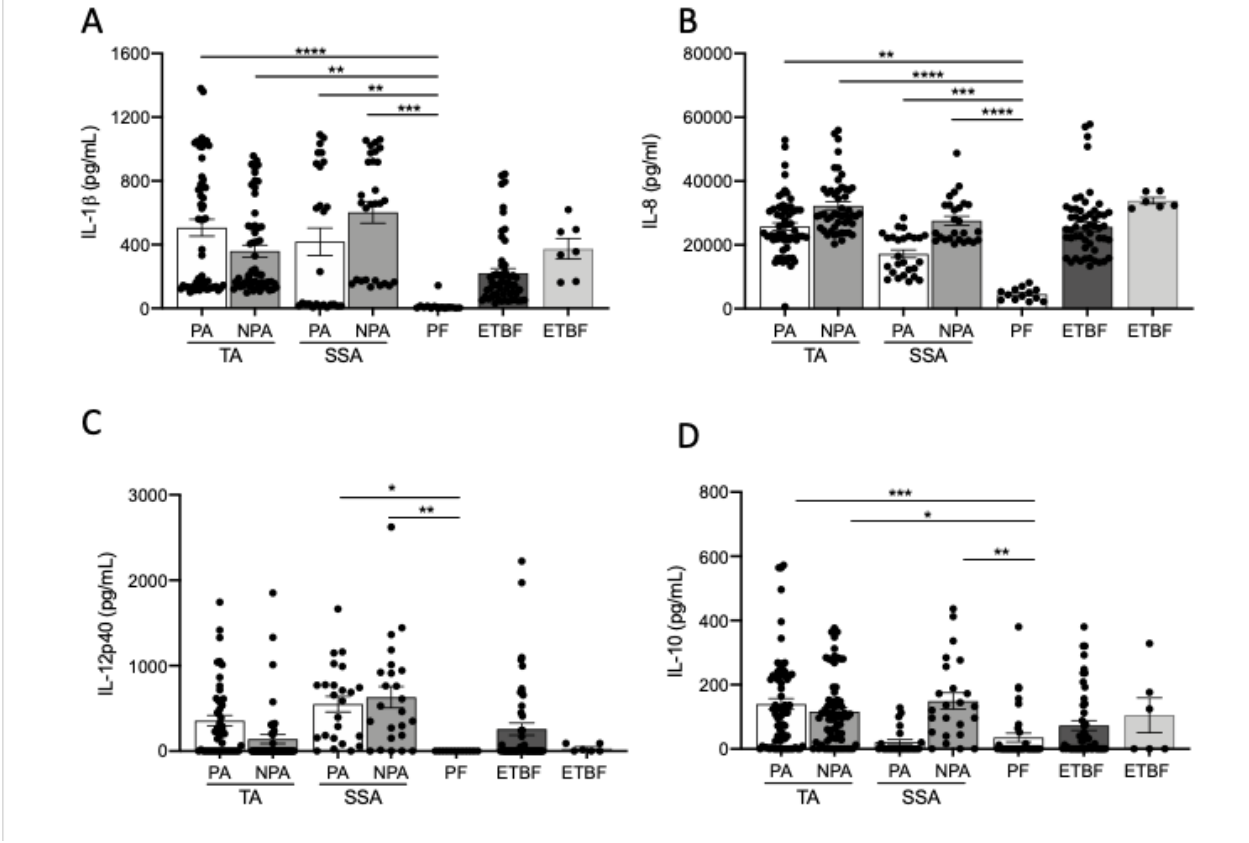


Figure 2.3 NTBF B. fragilis isolates from patients with tubular adenoma and sessile serrated adenoma are immunogenetically distinct from isolates from Polyp Free individuals

Levels of cytokines detected from co-culturing *Bacteroides fragilis* supernatants with monocytic cell line by ELISA. Data is represented as the mean ± s.e.m of 3-4 independent experiments. A-D, *, $p < 0.05$. **, $p < 0.001$. One-way ANOVA with multiple comparisons.

B. fragilis isolates from patients with tubular adenoma and sessile serrated adenoma contain a higher proportion of LPS-O antigen genes while *B. fragilis* isolates from patients with sessile serrated adenoma and Polyp-free individuals contained a higher proportion of capsular polysaccharides.

We performed Whole Genome Sequencing of the *Bacteroides fragilis* isolates in this study, in order to further characterize them. The enterotoxigenic strain of *Bacteroides fragilis* was also correlated to colon cancer development in various studies so we initially wondered whether the immunogenic *Bacteroides fragilis* isolates recovered from patients with adenoma belonged to this particular strain. We selected 36 different *Bacteroides fragilis* isolates recovered from 9 different patients for whole genome sequencing. Out of these 9 patients, 3 had sessile serrated adenoma in

the proximal colon, 4 had tubular adenoma and 2 patients were polyp-free. We selected 2-4 PA and 2-4 NPA isolates from each patient. After performing a pangenome analysis, we noticed that all the isolates shared an important core genome. Moreover, isolates from the same individuals were highly similar in gene content, whereas some variations were seen in the isolates across patients. We also found that all our *B. fragilis* samples had type 6 secretion system in their genome, which likely conferred them with a fitness advantage, allowing them to establish an ecological niche in the gut and explaining the high recovery rates *in vitro*. (Figure 2.4.A and B).

B. fragilis also usually features eight distinct surface capsular polysaccharides (PSA) that are believed to play a role in immune homeostasis in the gut. Blasting the genomes of our *B. fragilis* isolates further confirmed that they all had capsular polysaccharides in their genome, which likely allowed their persistence on mucosal surfaces. Our *B. fragilis* isolates also comprised other genes implicated in virulence such as hemolysins, genes associated with biofilm formation, and the *nanH* gene encoding neuraminidase, which has the capacity to cleave the mucin's polysaccharides and enhance the growth of the microorganism by generating available glucose. Other virulence factors identified in *B. fragilis*'s genome were genes encoding for antibiotic resistance as well as several virulence factors of foreign origin, suggesting a considerable "sharing" of genes within the crowded neighborhood of the gut flora. Moreover, when blasting the isolates' genomes for the B. fragilis enterotoxin (BFT), we discovered that all of our patients with adenoma were in fact colonized with Non-toxigenic strains of *B. fragilis* (NTBF) except for one patient with tubular adenoma and two patients which were adenoma-free. (Figure 2.4.A and B).

Next, we decided to characterize whether some of these virulence genes exerted a phenotype *in vitro*. Thus, we decided to first characterize the potential of these isolates to form biofilms *in vitro*. *Bacteroides fragilis* rich biofilms adhering to the surface of the colon's epithelial mucosa have been strongly correlated to CRC development. We noticed that all our NTBF isolates were similar in their potential to form biofilm while the ETBF isolates were better biofilm former (Supplemental Figure 2.3). This finding confirmed previous results found in the literature. However, B. fragilis isolates from patients with tubular adenoma and sessile serrated adenoma contain a higher proportion of LPS-O antigen genes while B. fragilis isolates from patients with sessile serrated adenoma and Polyp-free individuals contained a higher proportion of capsular polysaccharides (Figure 2.4. B).

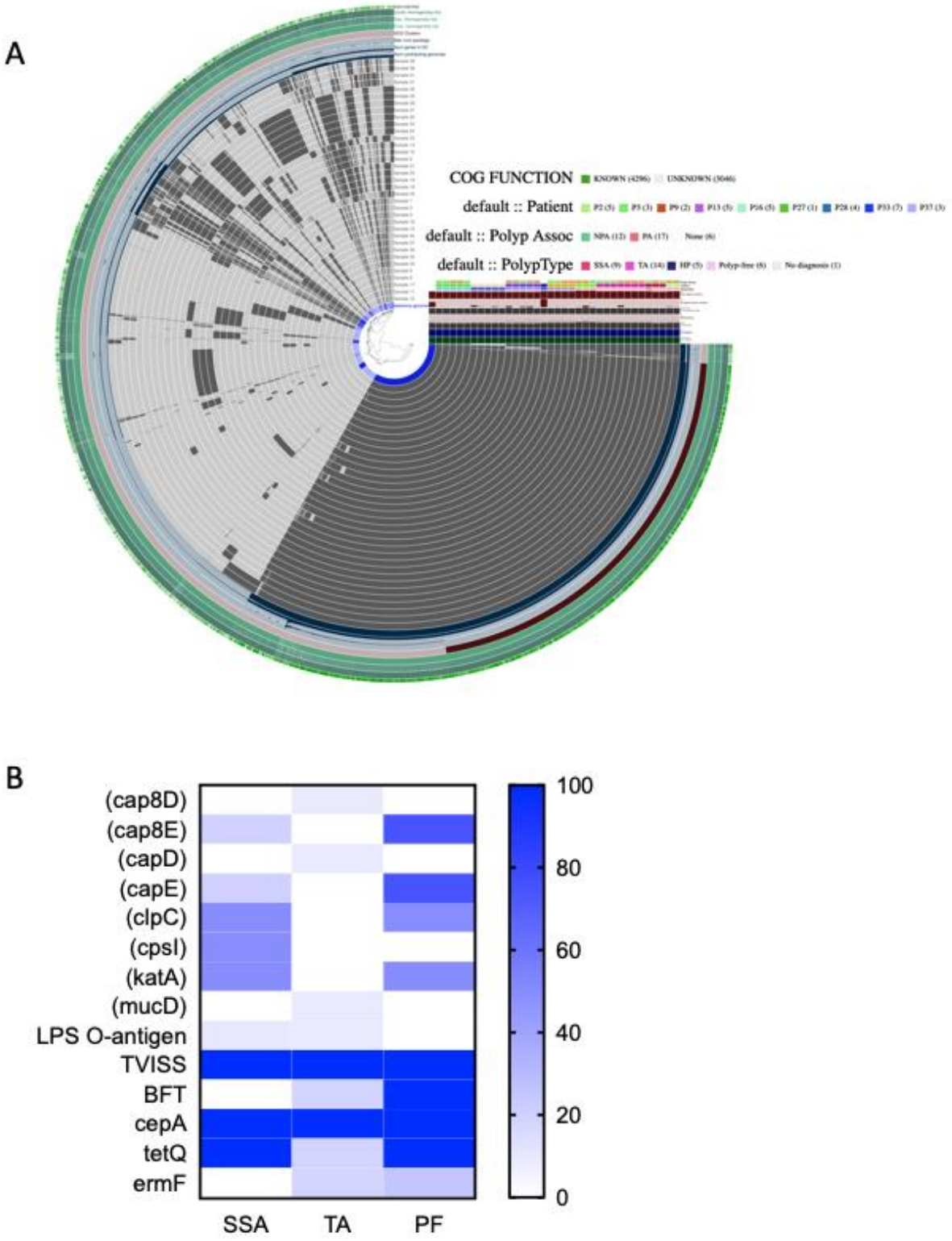


Figure 2.4. tubular adenoma and sessile serrated adenoma contained a higher proportion of LPS-O antigen genes while B. fragilis isolates from patients with sessile serrated adenoma and Polyp-free individuals contained a higher proportion of capsular polysaccharides.

Genome annotations and comparative genome analysis of *B. fragilis* isolates. (A) Aligned B. fragilis genomes made in Anvio'o (B) Heatmap representing the proportion of virulence factors found in each different type of isolates (TA, SSA, and PF).

NTBF *B. fragilis* isolates from patients with tubular adenoma and sessile serrated adenoma are significantly proteomically distinct from the isolates recovered from Polyp Free individuals.

We also had an insight into the proteome of these same isolates which we could obtain from their Peptide Mass Fingerprints (PMF) when running them on the MALDI TOF. A Principle component analysis of the PMFs of our *B. fragilis* isolates showed a discrimination between the isolates according to the type of tissue the microorganisms were isolated from. The TA PA and TA NPA isolates from our 4 different patients clustered similarly and distinctly from the SSA PA and NPA isolates. Moreover, both TA and SSA isolates were significantly proteomically distinct from the isolates recovered from PF individuals (Figure 2. 5).

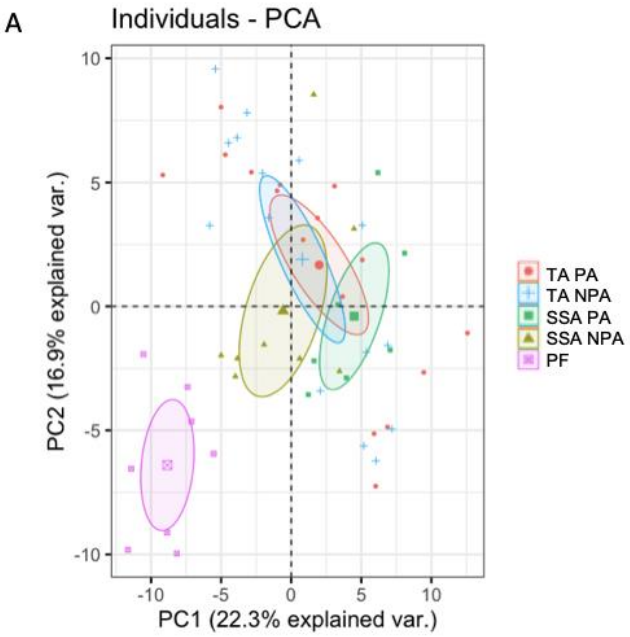


Figure 2.5. NTBF B. fragilis isolates from patients with tubular adenoma and sessile serrated adenoma are significantly proteomically distinct from the isolates recovered from Polyp Free individuals.

Principle component analysis based on peptide mass fingerprint profile of various NTBF *B. fragilis* isolates. Data is mean +/- of 2-3 independent experiments with n=4-6. Each dot represents the average of technical replicates.

NTBF *B. fragilis* isolates from patients with tubular adenoma and sessile serrated adenoma isolates stimulate TLR2-dependent signaling.

We investigated whether the proinflammatory response triggered by the TA PA NPA and SSA PA and NPA isolates was mediated by activation of either TLR2 or TLR4 innate immune receptors by utilizing a HEK293 luciferase reporter assay (REF). Nonimmune HEK293 cells were transfected with either TLR2-expressing plasmids, followed by infection with TA PA NPA and SSA PA and NPA as well as PF *B. fragilis* cell-free supernatants. As seen in Figure 2.6 A and B, all *B. fragilis* isolates stimulated TLR2 but did not increase levels of TLR4 stimulation. Since this assay measures fold change in activation, TLR2 activation by *B. fragilis* supernatants is highly significant.

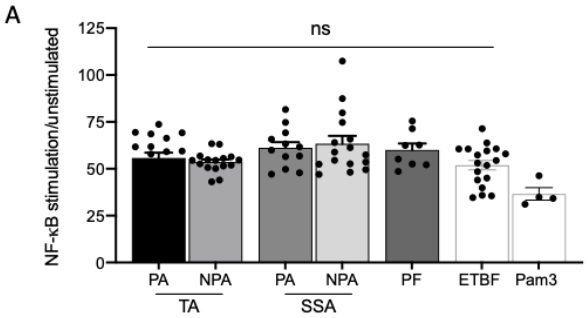


Figure 2.6. NTBF B. fragilis isolates from patients with tubular adenoma and sessile serrated adenoma isolates stimulate TLR2-dependent signaling

Fold NF-kB stimulation of HEK293 cells infected with B. fragilis isolates supernatants relative to unstimulated control. Results are means ± s.e.m from 203 independent experiments. (* $p < 0.05$ by One-way ANOVA with multiple comparisons.).

2.5. DISCUSSION

Recent studies have shown that certain symbiotic bacteria can tightly adhere to the intestinal mucosa[27] and that the colon microbiota composition in patients with colorectal adenoma is distinct from the microbiota composition in healthy patients[28,29,30,31]. This was further confirmed by our study when assessing the microbiota composition by 16S rRNA sequencing of a subset of patients in our study. Our study further showed that the microbiota composition of patients with tubular adenoma is distinct from the microbiota composition of patients with a different type of polyp: the

sessile serrated adenoma. However, while we didn't see statistically significant compositional differences between SSA biopsy samples and Polyp free biopsy samples in culture, the composition of the mucosal microbiome was significantly distinct when performing 16S between the same two groups. This was likely due to the selection processes that tend to occur in culture, clearly favoring *Bacteroides* species in the case of this study and potentially dwarfing other compositional differences. This can also further be explained by the smaller number of fastidious bacterial species that we can successfully isolate and grow *in vitro* compared to what can be simply detected by 16S sequencing.

While we recovered a higher relative abundance of commensal *Bacteroides* species *in vitro*, validation by 16S rRNA sequencing revealed an overall higher relative abundance of firmicutes in the mucosal biopsies. This was likely due to the fact that we only used Tryptic Soy Agar supplemented with 5% sheep blood in order to isolate the mucosal microbiota from our biopsy samples. This type of media, while being non-selective, doesn't particularly favor the growth of Firmicutes which prefer to grow on media such as MRS agar plates[32,33].

Bacteroides fragilis was recovered in the highest abundance in patients with tubular adenoma followed with patients with sessile serrated adenoma compared to patients who were polyp-free. It is usually a ubiquitous gut commensal and it has been reported that its capsular PSA plays a role during the homeostatic intestinal colonization by the microbiota. *B. fragilis* PSA induces the anti-inflammatory function of Tregs through TLR2 signaling. This function contrasts with the pro-inflammatory role of TLR2 ligands for pathogens and hence allows the persistence of *B. fragilis* on mucosal surfaces of the gut. Moreover, Livanis et al. showed in 2016 that the gut species *B. fragilis* produced T6SSs that deploy toxins able to antagonize other species in the gut[34]. These systems are synthesized in the mammalian gut and allow the producing *B. fragilis* strain to create a protected niche in the human colon .

Multiple virulence factors of "foreign" origin were also found in the genomes of our B. fragilis isolates. Acquiring genes that code for improved adhesion and antibiotic resistance for instance gives these *B. fragilis* isolates an edge in further establishing a niche for themselves.

Both capsular PSA and T6SSs genes were found in the genomes of our *B. fragilis* isolates along with virulence factors found in other bacterial species as well, and antibiotic resistance genes. All these features explain why *B. fragilis* was recovered at such high abundance in our study. The arsenal of adhesins, biofilm forming genes, secretions systems, antibiotic resistance genes, PSA capsules will consequently allow *B. fragilis* to robustly survive on the colon's mucosal surface.

Enterotoxigenic *B. fragilis* (ETBF) is a toxin-producing bacterium thought to possibly promote colorectal carcinogenesis by modulating the mucosal immune response and inducing epithelial cell changes. While certain studies have shown an association of ETBF positivity and increased abundance with early-stage carcinogenic lesions[35], our study only featured 3 patients out of 40 with ETBF positive *Bacteroides fragilis*, two of which were adenoma-free at the time of the colonoscopy. The literature reports less than 10% of patients with colorectal adenoma end up developing CRC[11]. In our study, one ETBF-positive patient out of 31 patients with adenoma represents 3.2% of our patient population and could signify that this particular patient is at a higher risk of developing CRC compared to all the others. However, this does not negate the fact that non-toxigenic *B. fragilis* (NTBF) strains isolated from our patients with adenoma turned out to have a highly pro-inflammatory profile when assayed *in vitro* with cells lines expressing TLR2. Moreover, the fact that less than 10% of patients with colorectal adenoma end up developing CRC can likely be due to the fact that adenoma are usually removed during a colonoscopy and thus rarely get the chance to evolve to malignancy with proper screening. While NTBF entertains a mutualistic relationship with the human host in a healthy gut, this commensal bacterium can cause significant pathology, including bacteremia and abscess formation, when translocated to a new environment[23]. Our NTBF isolates from patients with adenoma, while inducing *IL-10* in vitro, were also able to induce pro-inflammatory cytokines IL-12P40, IL-8, and IL-1ß. Moreover, patients with adenoma had higher levels of IL-12p40 and lower levels of IL-10 compared to the healthy control group. IL-12P40 is a subunit part of another heterodimeric cytokine IL-23 (p40 and p19)[36]. Some studies have reported that early stages of CRC tend to present with high levels of IL-12P40 and lower levels of IL-10[37]. The subunit IL-12p40 presents in both IL-12 and IL-23 and is believed to be an important immunostimulatory agent through the promotion of DC migration, macrophage inflammation, and the production of Interferon-gamma in CD8+ T cells[38]. IL-12p40 can also bind another subunit -p19 in order to form IL-23, a biologically active cytokine that has been intimately related to inflammatory bowel disease and CRC. In fact, IL-23 signaling was shown to promote tumor growth and progression, and development of a tumoral IL-17 response in a mouse model of CRC[36]. Moreover, T-helper IL-17 production in early stages CRC (stage I/II) was associated with a drastic decrease in disease-free survival in the course of the disease.

Our isolates were also shown to produce very important levels of IL-8 when isolated from patients with adenoma. Multiple studies have suggested IL-8 as a promising biomarker for CRC

detection[39]. IL-8, also known as neutrophil-activating factor (NAF), was the first chemokine identified. IL-8 is believed to control leukocyte trafficking during homeostasis and inflammation[40]. However, it was also shown to be promote CRC cell proliferation, invasion, migration, and angiogenesis[41]. In that regard, blocking the IL-8 receptors CXCR1 and CXCR2 could constitute an interesting anti-tumor therapeutic strategy for CRC prevention and management.

Another cytokine believed to play a key role in pro-tumor inflammation and that was robustly produced by NTBF isolates from patients with adenoma is ILB. IL-1ß is also accountable for the overexpression of cyclooxygenase-2 (COX-2) in CRC and its antiapoptotic role in tumor cells[42] (REF). Thus, we believe that a dysbiotic, inflamed, pre-cancerous gut environment harbors NTBF displaying pathogenic characteristics through the regulation of their genes. Proteomic profile of our NTBF isolates clustered according to the type of mucosal environment that the isolates were recovered from despite highly similar gene contents in all the isolates. Indeed, the microbial proteins content of isolates from adenoma-free patients were significantly distinct from those of isolates from patients with adenoma. Thus, a pre-cancerous mucosal environment could cause *B. fragilis* to express new or different surface epitopes, such as lipid-O antigens, leading them to be recognized as pathogens by TLR2 receptors on Antigen-presenting cells. Signaling of *B. fragilis* through TLR2 activates the nuclear factor K B (NF-KB), potentially leading to the production of several proinflammatory cytokines involved in promoting the growth of colon tumors.

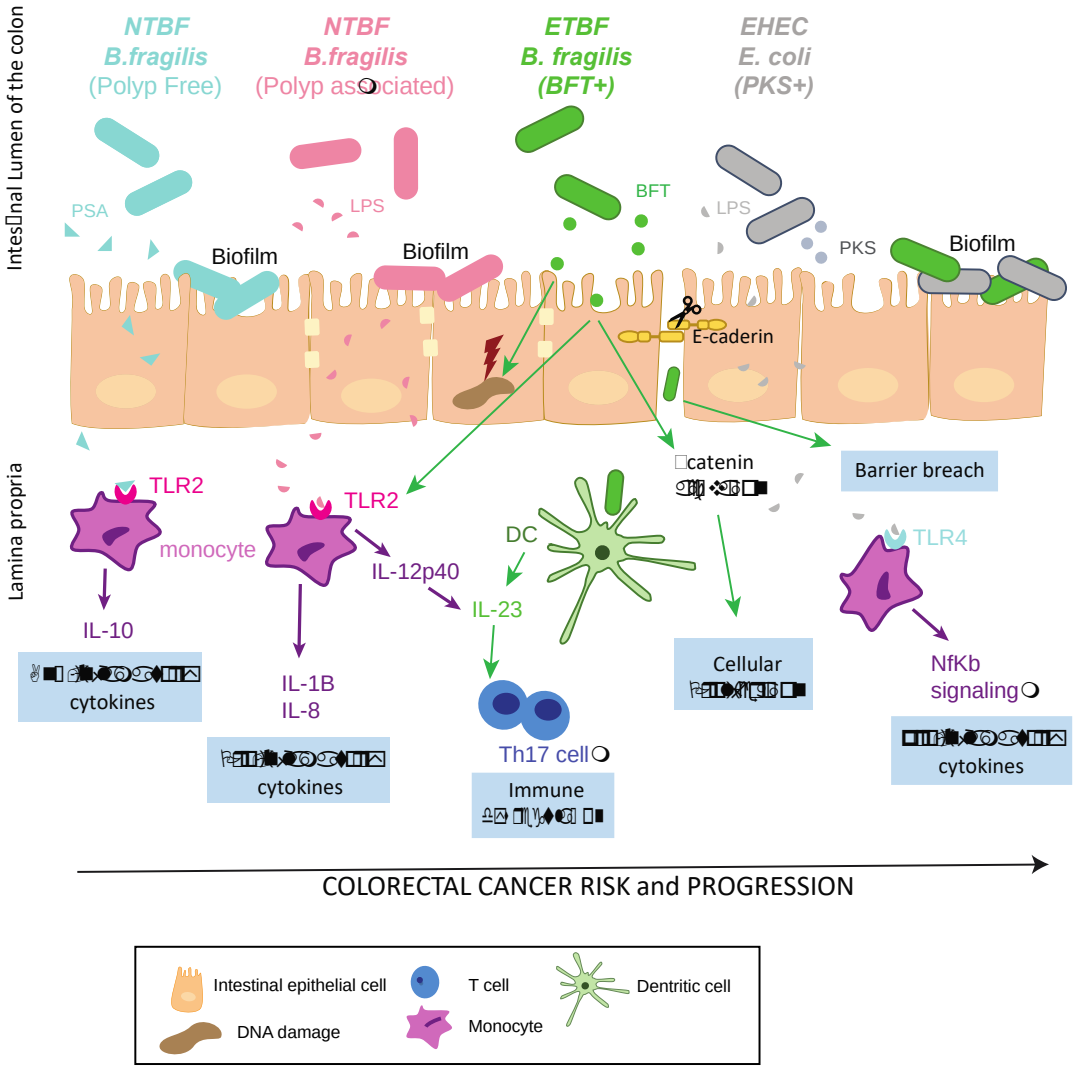
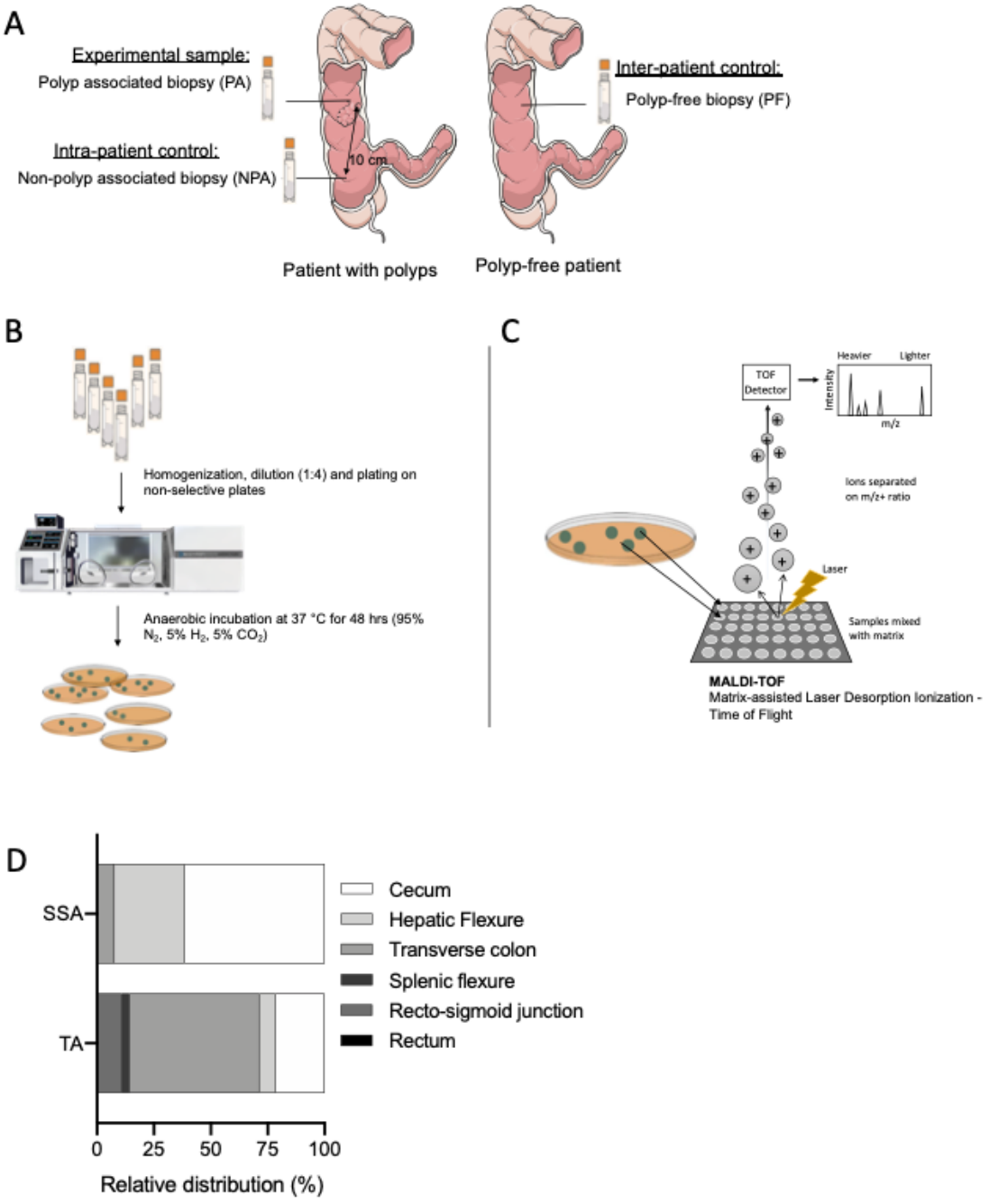


Figure 2. 7. Mechanisms of Bacteroides fragilis in CRC progression

2.6. CONCLUSION

A pathological imbalance of the gut microbiome is present in patients with colorectal adenoma. *Bacteroides fragilis* is a microorganism that is found in the gut mucosa and that can affect CRC development by invading tissues and modulating the host immune response. Interesting features of early colorectal cancer development identified in this study were high recovery rates of *Bacteroides fragilis*, associated with high levels of IL-8, IL12-P40, and IL-1ß, which are all believed to play a role in CRC pathogenesis. The production of these pro-inflammatory cytokines was also associated with NF-kB stimulation by *B. fragilis* through TLR2. Taken together, these features allow the identification of potential microbiome-based diagnostic markers. Blocking the receptors for IL-8 and IL-23 may also offer important strategies to interfere with the development and progression of CRC. So would the inhibition of NF-KB through the identification of a *B. fragilis*-associated ligand that is responsible for the pro-inflammatory TLR2 signaling phenotype that we observed. The information revealed in this study also stresses the importance of pairing genomic analysis with *ex* and *in vivo* characterization of gut commensals in order to better understand their relationship with the human host and their involvement in CRC pathogenesis.

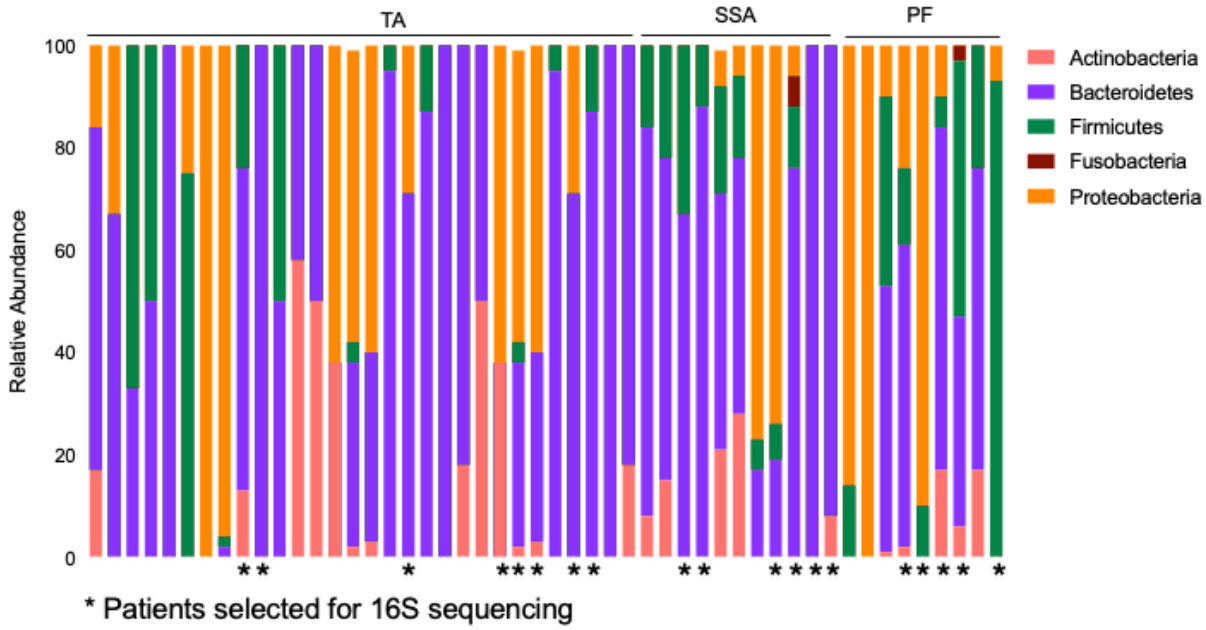
2.7. SUPPLEMENTAL MATERIAL



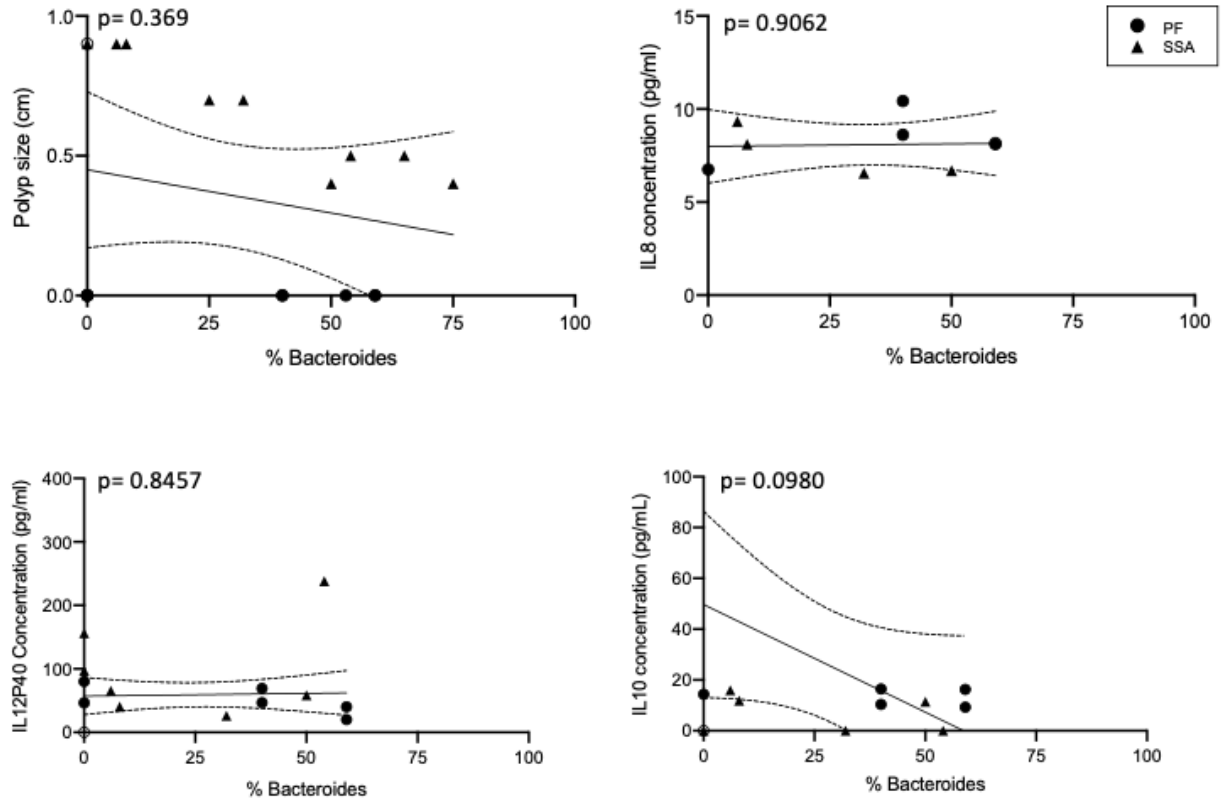
Supplemental Diagram 2.1. Biopsy sampling and polyp types cataloguing

Species	Strains	Characteristics	Origin
B. fragilis	86-5443-2-2	BFT + (intestinal)	Cynthia Sears Lab
B. fragilis	Δ86-5443-2-2	BFT - (intestinal)	Cynthia Sears Lab
E. coli	DH5 alpha	bfp – (intestinal)	ATCC
B. fragilis	See phylogenetic tree	BFT + (intestinal)	This study
B. fragilis	See phylogenetic tree	BFT - (intestinal)	This study

Supplemental Table 2.1 Bacterial strains used in this study.

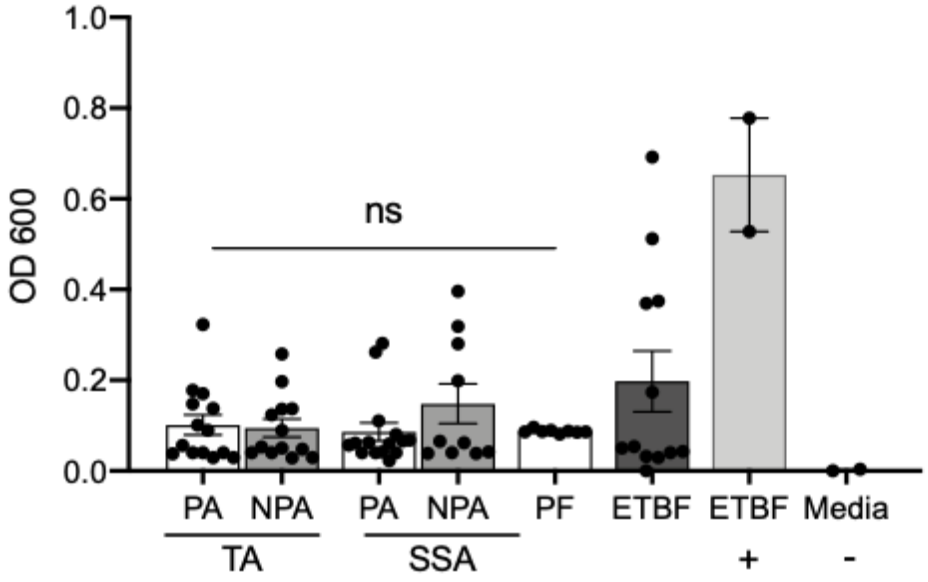


Supplemental Figure 2.1 Phylum level abundance of mucosal isolates at the individual patient level.



Supplemental Figure 2.2 Correlation in Patients with sessile serrated adenoma

Correlation in patients with SSA between (A) Polyp size in cm and abundance of *Bacteroides* isolated *in vitro* (B) Mucosal IL-8 concentration in pg/mL and abundance of *Bacteroides* isolated *in vitro* (C) Mucosal IL-12p40 concentration in pg/mL and abundance of *Bacteroides* isolated *in vitro* (D) Mucosal IL-10 concentration in pg/mL and abundance of *Bacteroides* isolated *in vitro* Data is mean ± s.e.m of 3 independent experiments. Statistics are Pearson Correlation Coefficient of all data points



Supplemental Figure 2.3 Biofilm formation of B. fragilis isolates

In vitro Biofilm formation quantified by Crystal violet staining. Data is represented as the mean ± s.e.m of 3 independent experiments. (*, $p < 0.05$ One-way ANOVA with multiple comparisons)

2.8. REFERENCES

1. Siegel, R. L. *et al.* Colorectal cancer statistics, 2017. *CA: A Cancer Journal for Clinicians* **67**, 177–193 (2017).
2. Key statistics for colorectal cancer. http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-key-statistics.
3. Long, A. G., Lundsmith, E. T. & Hamilton, K. E. Inflammation and Colorectal Cancer. *Curr Colorectal Cancer Rep* **13**, 341–351 (2017).
4. Shalapour, S. & Karin, M. Immunity, inflammation, and cancer: an eternal fight between good and evil. *J Clin Invest* **125**, 3347–3355 (2015).

5. Grivennikov, S. I., Greten, F. R. & Karin, M. Immunity, Inflammation, and Cancer. *Cell* **140**, 883–899 (2010).
6. Sandouk, F., Al Jerf, F. & Al-Halabi, M. H. D. B. Precancerous Lesions in Colorectal Cancer. *Gastroenterol Res Pract* **2013**, (2013).
7. Testa, U., Pelosi, E. & Castelli, G. Colorectal Cancer: Genetic Abnormalities, Tumor Progression, Tumor Heterogeneity, Clonal Evolution and Tumor-Initiating Cells. *Med Sci (Basel)* **6**, (2018).
8. Lieberman, D. A. *et al.* Guidelines for Colonoscopy Surveillance After Screening and Polypectomy: A Consensus Update by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology* **143**, 844–857 (2012).
9. Strum, W. B. Colorectal adenomas. *N. Engl. J. Med.* **374**, 1065–1075 (2016).
10. Mak, T., Lalloo, F., Evans, D. G. R. & Hill, J. Molecular stool screening for colorectal cancer. *Br J Surg* **91**, 790–800 (2004).
11. Bonnington, S. N. & Rutter, M. D. Surveillance of colonic polyps: Are we getting it right? *World J Gastroenterol* **22**, 1925–1934 (2016).
12. Yang, J. *et al.* Characteristics of Differently Located Colorectal Cancers Support Proximal and Distal Classification: A Population-Based Study of 57,847 Patients. *PLoS One* **11**, (2016).
13. Wong, R. Proximal Tumors Are Associated with Greater Mortality in Colon Cancer. *J Gen Intern Med* **25**, 1157–1163 (2010).
14. Thursby, E. & Juge, N. Introduction to the human gut microbiota. *Biochem J* **474**, 1823–1836 (2017).

15. Guinane, C. M. & Cotter, P. D. Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. *Therap Adv Gastroenterol* **6**, 295–308 (2013).
16. Hansson, G. C. Role of mucus layers in gut infection and inflammation. *Curr Opin Microbiol* **15**, 57–62 (2012).
17. Ellermann, M. & Sartor, R. B. Intestinal bacterial biofilms modulate mucosal immune responses. *J Immunol Sci* **2**, 13–18 (2018).
18. Mentis, A.-F. A., Boziki, M., Grigoriadis, N. & Papavassiliou, A. G. Helicobacter pylori infection and gastric cancer biology: tempering a double-edged sword. *Cell. Mol. Life Sci.* **76**, 2477–2486 (2019).
19. Kordahi, M. C. & William DePaolo, R. The Influence of the Microbiota on the Etiology of Colorectal Cancer. in *Mechanisms Underlying Host-Microbiome Interactions in Pathophysiology of Human Diseases* (eds. Sun, J. & Dudeja, P. K.) 167–193 (Springer US, 2018). doi:10.1007/978-1-4939-7534-1_8.
20. Zamani, S. *et al.* Enterotoxigenic Bacteroides fragilis: A Possible Etiological Candidate for Bacterially-Induced Colorectal Precancerous and Cancerous Lesions. *Front. Cell. Infect. Microbiol.* **9**, (2020).
21. Dejea, C. M. *et al.* Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* **359**, 592–597 (2018).
22. Sears, C. L. Enterotoxigenic Bacteroides fragilis: a rogue among symbiotes. *Clin. Microbiol. Rev.* **22**, 349–369, Table of Contents (2009).
23. Wexler, H. M. Bacteroides: the good, the bad, and the nitty-gritty. *Clin. Microbiol. Rev.* **20**, 593–621 (2007).

24. Mazmanian, S. K., Liu, C. H., Tzianabos, A. O. & Kasper, D. L. An Immunomodulatory Molecule of Symbiotic Bacteria Directs Maturation of the Host Immune System. *Cell* **122**, 107–118 (2005).
25. Mazmanian, S. K., Round, J. L. & Kasper, D. L. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* **453**, 620–625 (2008).
26. Round, J. L. *et al.* The Toll-like receptor pathway establishes commensal gut colonization. *Science* **332**, 974–977 (2011).
27. Chassaing, B. & Gewirtz, A. T. Identification of Inner Mucus-Associated Bacteria by Laser Capture Microdissection. *Cell Mol Gastroenterol Hepatol* **7**, 157–160 (2018).
28. Shen, X. J. *et al.* Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. *Gut Microbes* **1**, 138–147 (2010).
29. Sanapareddy, N. *et al.* Increased rectal microbial richness is associated with the presence of colorectal adenomas in humans. *ISME J* **6**, 1858–1868 (2012).
30. Brim, H. *et al.* Microbiome analysis of stool samples from African Americans with colon polyps. *PLoS ONE* **8**, e81352 (2013).
31. Wei, H. *et al.* Structural shifts of gut microbiota as surrogate endpoints for monitoring host health changes induced by carcinogen exposure. *FEMS Microbiol. Ecol.* **73**, 577–586 (2010).
32. Stackebrandt, E. The Family Lachnospiraceae. in *The Prokaryotes: Firmicutes and Tenericutes* (eds. Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E. & Thompson, F.) 197–201 (Springer, 2014). doi:10.1007/978-3-642-30120-9_363.
33. Dicks, L. & Endo, A. The Family Lactobacillaceae: Genera Other than Lactobacillus. in *The Prokaryotes: Firmicutes and Tenericutes* (eds. Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E. & Thompson, F.) 203–212 (Springer, 2014). doi:10.1007/978-3-642-30120-9_207.

34. Chatzidaki-Livanis, M., Geva-Zatorsky, N. & Comstock, L. E. Bacteroides fragilis type VI secretion systems use novel effector and immunity proteins to antagonize human gut Bacteroidales species. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3627–3632 (2016).

35. Purcell, R. V. *et al.* Colonization with enterotoxigenic Bacteroides fragilis is associated with early-stage colorectal neoplasia. *PLoS ONE* **12**, e0171602 (2017).

36. Grivennikov, S. I. *et al.* adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature* **491**, 254–258 (2012).

37. Mager, L. F., Wasmer, M.-H., Rau, T. T. & Krebs, P. Cytokine-Induced Modulation of Colorectal Cancer. *Front Oncol* **6**, (2016).

38. Cooper, A. M. & Khader, S. A. IL-12p40: an inherently agonistic cytokine. *Trends Immunol.* **28**, 33–38 (2007).

39. Jin, W.-J., Xu, J.-M., Xu, W.-L., Gu, D.-H. & Li, P.-W. Diagnostic value of interleukin-8 in colorectal cancer: A case-control study and meta-analysis. *World J Gastroenterol* **20**, 16334–16342 (2014).

40. Ha, H., Debnath, B. & Neamati, N. Role of the CXCL8-CXCR1/2 Axis in Cancer and Inflammatory Diseases. *Theranostics* **7**, 1543–1588 (2017).

41. Ning, Y. *et al.* Interleukin-8 is associated with proliferation, migration, angiogenesis and chemosensitivity in vitro and in vivo in colon cancer cell line models. *Int J Cancer* **128**, 2038–2049 (2011).

42. Andersen, V., Holst, R., Kopp, T. I., Tjønneland, A. & Vogel, U. Interactions between Diet, Lifestyle and IL-10, IL-1ß, and PTGS2/COX-2 Gene Polymorphisms in Relation to Risk of Colorectal Cancer in a Prospective Danish Case-Cohort Study. *PLOS ONE* **8**, e78366 (2013).

3. Detection of Physiological Changes from Environmental Conditions through Matrix-Assisted Laser Desorption/Ionization – Time of Flight (MALDI-TOF) Mass Spectrometry

Adapted from:
Proteomic changes in bacteria caused by exposure to environmental conditions can be detected by Matrix-Assisted Laser Desorption/Ionization – Time of Flight (MALDI-ToF) Mass Spectrometry

Melissa Kordahi[1], **Denise Chac**[1], Leandra Brettner[2], Arushi Verma[3], Paul McCleary[1], Kelly Crebs[1], Cara Yee[1], and R. William DePaolo[4,5]*

Authors in bold contributed equally to this manuscript

[1] Department of Pathology, University of Washington, Seattle, WA 98195
[2] Department of Bioengineering, University of Washington, Seattle WA 98195
[3] Division of Endocrinology, Seattle Children's Hospital | University of Washington, Seattle WA 98115
[4] Department of Medicine, University of Washington, Seattle WA 98195
[5] Center for Microbiome Sciences & Therapeutics, University of Washington, Seattle WA 98195

* Corresponding author:
R. William DePaolo, PhD
Department of Medicine
University of Washington,
1959 NE Pacific Avenue,
Seattle, WA 98195
(312) 371-7828
Email: wdepaolo@medicine.washington.edu

KEYWORDS: MALDI-TOF MS, proteotyping, microbiology

Grant Support: Supported by UW start-up funds (RWD)

3.1. ABSTRACT

In the past decade, matrix-assisted laser desorption/ionization time-of-flight (MALDI-ToF) mass spectrometry (MS) has become a timely and cost-effective alternative to bacterial identification. The MALDI-ToF MS technique analyzes the total protein of culturable microorganisms at the species level and produces a mass spectrum based on peptides which is compared to a database of identified profiles. Consequently, unique signatures of each microorganism are produced allowing identification at the species and, more importantly, strain level. Our present study proposes that the MALDI-ToF MS can be further used to screen functional and metabolic differences between the same species of bacteria. While other studies applied the MALDI-ToF technique to identify subgroups within species, we investigated how various environmental factors could alter the unique bacterial signatures. We found that genetic and phenotypic differences between microorganisms belonging to the same species can be differentiated using the peptide-mass fingerprints generated by MALDI-ToF MS. Importantly, our results indicate that proteomic differences caused by exposure to different environmental stimuli leave behind a proteomic imprint that can still be detected by MALDI-ToF even when the bacteria are subsequently cultured on the same nutrient rich agar. Together, our results demonstrate that MALDI-ToF MS is a cost-effective and efficient screen that can be used to discriminate proteomic signatures within the same species of bacteria and can be paired with other -Omic approaches to provide a holistic view of microbial transcriptional regulation.

3.2. INTRODUCTION

The MALDI-ToF (matrix-assisted desorption/ionization time-of-flight) mass spectrometry (MS) technology offers a time- and cost-effective method of identifying microorganisms. Compared to previous time-consuming and expensive methods to identify microorganisms based on 16s rRNA or whole genome sequencing, the MALDI-ToF MS provides rapid, accurate and inexpensive identification within minutes via proteotyping[2]. While the MALDI-ToF is limited to culturable microorganisms and public databases, it has been quickly incorporated in the clinical setting and

used for diagnostic purposes. Studies have shown that MALDI-ToF MS can equally, or more efficiently, better identify sources of systemic[3,4], urinary[5–7], respiratory[8] and intestinal[9,10] infections. The MALDI-ToF MS technology allows for identification down to the strain level and has been shown to discriminate between strains of methicillin-resistant *Staphylococcus aureus*[11–14], shiga-toxigenic *Escherichia coli*[15], clinically relevant strains of Aspergillus[16] and many others[2].

Although many existing methods allow rapid identification of microorganisms to the species level, identification to the more specific ‘‘strain’’ taxon tends to be more challenging, as strains within a single microbial species are often very genotypically and phenotypically similar, despite having different functions. Higher resolution approaches such as molecular genetics are thus commonly employed to identify and characterize strains within a microbial species[17]. These include pulsed field gel electrophoresis (PFGE)[18], multilocus sequence typing[19] (MLST), repetitive extragenic palindromic PCR[20] (rep-PCR), housekeeping gene (e.g., PheS) sequence analysis[21], and whole genome sequencing[22,23]. Each of these approaches has been shown to have adequately high resolution to distinguish microbial strains from one another; however, they are costly, labor- and time-intensive, and lack the required rapid high-throughput nature of MALDI-ToF. A big advantage of the MALDI-ToF MS strain typing application would be epidemiologic investigations that require rapid identification of a single strain within a single species to determine the origin and spread of an outbreak in order to mitigate risks to public safety posed by microbial food and water contamination or potential acts of bioterrorism[17,24]. Moreover, in an era where microbiome-based diagnostics and therapeutic targets are envisioned[25], this technology could provide a framework to link disease phenotypes to peptides translated by microorganisms isolated from various biological samples.

The MALDI-ToF MS is an ionization process that became commercially available in the early 1990s. The ionization process involves the mixture of an analyte with a “matrix” solution and the co-crystallization of both matrix and analyte (**Diagram 3.1A-B**). During ionization, a laser of a determined UV nm wavelength is fired at the mixture and causes its desorption into a gaseous phase. MALDI ion sources are combined with time of flight (ToF) tubes for ion separation[16]. A peptide mass fingerprint (PMF) is generated per sample and compared to a database of known bacterial species PMFs (**Diagram 3.1C**). MALDI-ToF looks at whole cell differences as opposed to a method that target a handful of molecules such as qPCR, protein blots, etc. which allows

identification of species and strain differences that other molecular techniques would miss. While the technique lacks discriminatory power with some species such as *Enterococcus faecium* and *Staphylococcus aureus*[26], the MALDI-ToF MS can robustly identify bacterial species in various culturing conditions[27,28].

In contrast to studies that evaluate the ability of MALDI-ToF MS ability to identify bacterial species regardless of environmental conditions, we propose that the variation of MALDI-ToF MS-produced peaks due to culturing can detect environmental imprints. By comparing PMF peaks and observing the differences in dendrogram and principal component of analysis (PCA) format (**Diagram 3.1 D and E**), we report that the MALDI-ToF can reliably differentiate between various conditions including oxygen presence, nutrient availability and temperature stress. The MALDI-ToF MS is also able to detect previous environmental conditions after culturing out of the initial environment. Our study shows that the MALDI-ToF can be applied to screening bacterial isolates from various environmental conditions and biological samples in order to see potential functional differences within the same microorganism.

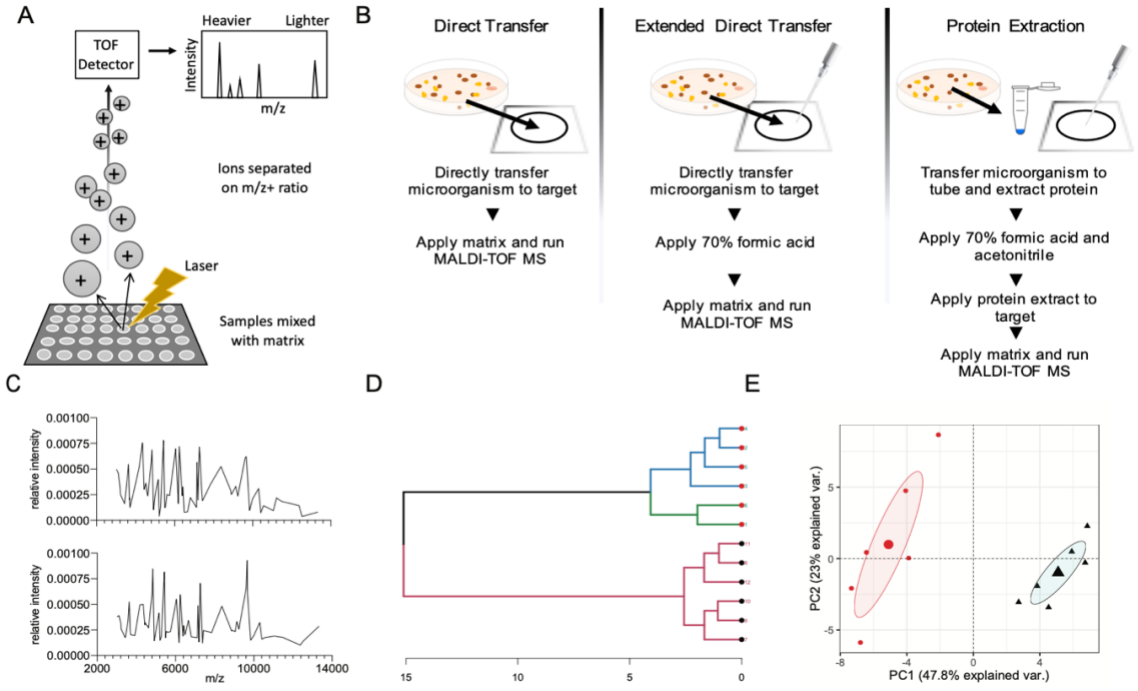


Diagram 1: MALDI-ToF MS methodology

(A) Schematic of MALDI-ToF MS technology and B) methods of processing as described by manufacturer Bruker. Analysis of peptide-mass fingerprints (PMF) by C) spectra, D) dendrogram and E) principal component analysis (PCA).

3.3. MATERIALS AND METHODS

Bacterial strains and growth conditions:
DH5α *Escherichia coli* and enteropathogenic *E. coli* were grown on tryptic soy agar (TSA) plates and incubated at 37°C for 24 hours. *Yersinia enterocolitica* 8081 and *Y. enterocolitica inv, yadA,* and *yscA* knock-out strains were grown on TSA and incubated at room temperature for 48 hours. Enterotoxigenic *Bacteroides fragilis* (ETBF) and non-enterotoxigenic *B. fragilis* (NTBF) were grown on TSA with 5% defibrinated sheep blood (TSBA) from skim milk stocks and incubated at 37°C anaerobically for 5 days. *Lactobacillus reuteri* was grown on TSA and incubated at 37°C for 24 hours aerobically and anaerobically. All bacterial strains used are listed in Supplemental Table 1. For testing of various growth agar, *E. coli* DH5α was plated on TSA, TSBA, MacConkey and eosin methylene blue agars. The plates were then incubated aerobically at 37°C for 24 hours. For testing of temperature stress environment, overnight cultures of DH5α in Lysogeny broth (LB) were diluted and incubated at 37°C or 50°C for 15 and 30 minutes. For testing of residual effects of environment, overnight cultures of DH5α in tryptic soy broth (TSB) were subcultured in various media and then plated on TSA. Briefly, DH5α overnight culture in TSB was centrifuged at 3,000 x g for 5 minutes and washed in sterile PBS twice. The pellet was then resuspended in PBS and cultured in LB at 37°C or 50°C for 1 hour, shaking. Samples were then plated on TSA plates and incubated at 37°C for 24 hours.

Isolation and growth of clinical samples:
Stool samples were collected from pre-diabetic adolescents aged 13-19 years as part of an ongoing study after institutional IRB approval. Participants were given a sterile stool collection kit (stool collection container [Precision (Covidien, Fisher Scientific)]), stool collector hat and sterile gloves to self-collect stool at home and mail to the University of Washington laboratory within 24-48 hours of collection at room temperature. When received, the sample was immediately aliquoted, and one tube was used for the analysis of 'Fresh' sample on the same day it was received. The other aliquots were stored at -80°C for 2-3 months until analyzed. For analysis of the 'frozen sample', the sample was thawed on ice for 30 mins and then room temperature for 30 mins. Forty-eighty mg of each sample was suspended in 1ml PBS and centrifuged at 300 rpm for 2 mins. Serial

dilutions were made and plated on TSA and TSA blood and incubated in aerobic and anaerobic conditions at 37°C for 48 hours. Both 'Fresh' and 'Frozen' samples were analyzed using the direct protein method.

Protein Extraction Method for MALDI-ToF MS

Bacterial colonies on agar were picked and added to 300 µl of HPLC-grade water in a 1.5 ml Eppendorf tube and then mixed thoroughly with 900 µl of 100% ethanol. After centrifugation at 13000 rpm, pellets were dried at room temperature and directly mixed with equal volumes of 70% formic acid and acetonitrile (20-40 µl, depending on pellet size). After centrifugation at 13000 rpm, 1 µl of protein extract was spotted on a 96-target polished steel plate (Bruker Daltonics, Bremen, Germany) in four replicates, air-dried, and overlaid with 1 µl of alpha-cyano-4-hydroxycinnamic acid matrix solution (Bruker Daltonics, Bremen, Germany). Target plate was placed in the MALDI-ToF MS MicroFlex Biotyper for microbial identification (Bruker, Germany).

Extended Direct Method for MALDI-ToF MS

Bacterial colonies on agar were picked and placed directly on a target plate for identification. Samples were overlaid with 1 µL of 70% formic acid and alpha-cyano-4-hydroxycinnamic acid matrix solution (Bruker Daltonics, Bremen, Germany). Target plate was placed in the MALDI-ToF MicroFlex Biotyper for microbial identification (Bruker, Germany).

Cellular adhesion and invasion assays with *Y. enterocolitica*

SW620 (ATCC; CCL-227) human intestinal epithelial adenoma cells were used. SW620 cells were maintained in DMEM supplemented with 10% fetal bovine serum and 200uM L-glutamine. Prior to cellular assays, SW620 cells were seeded into 24-well plates and grown to near confluency over 3-4 days. SW620 were rinsed with PBS and infected with *Y. enterocolitica* at a MOI of 10:1 with each starting dose plated on TSA to enumerate CFUs. Briefly, overnight *Y. enterocolitica* cultures were diluted 1:10 in TSB. Subcultures were incubated for 2 hours at 37°C while shaking at 150 rpm. Cultures were washed with PBS and diluted to the appropriate infection dose in cell media. *Y. enterocolitica* were introduced to cells and infection initiated by centrifuging the plates for 5 minutes at 500 x g. For adhesion assays, Y. *enterocolitica* was co-cultured with SW620 cells for 1 hour at 37°C in a 5% CO_2 incubator. For intracellular invasion, *Y. enterocolitica* was co-cultured with SW620 cells for 1 hour, washed with PBS and treated with 100 µg/ml gentamicin (Corning,

Corning, NY). To elute the *Y. enterocolitica* and lyse the SW620s, each well was washed three times with warm PBS and treated with 500 ul PBS with 1% Triton X-100 (AMRESCO, VWR, Randor PA) for 5 minutes. Samples were plated on TSA in serial dilutions to enumerate *Y. enterocolitica* colonization. In parallel, wells containing media only were infected with an equal quantity of bacteria. Percentage of colonization was calculated by dividing the number of *Y. enterocolitica* CFU recovered from co-cultures by the number of *Y. enterocolitica* CFU initially applied to the wells.

Data analysis and Statistics

Raw spectra text files were analyzed using the R package, MALDIquant [https://www.ncbi.nlm.nih.gov/pubmed/22796955]. The raw data were trimmed to a spectra range of 3,000 to 15,000 m/z. The spectra intensities were then square-root transformed and smoothed using the Savitzky-Golay algorithm. Baseline noise was removed using the statistics-sensitive non-linear iterative peak clipping (or SNIP) algorithm with 100 iterations. The data were then normalized using total ion current (or TIC) calibration, which sets the total intensity to 1. Multiple spectra within the same analysis were aligned to the same x-axis using the Lowess warping method, a signal-to-noise ratio of 3, and a tolerance of 0.001. Peaks were detected from the average of at least 4 technical replicates using median absolute deviation. Principal components analyses and hierarchical clustering were also performed in R using the base stats package. Hierarchical clustering was performed on a calculated Euclidean distance matrix using Ward's method.

3.4. RESULTS AND DISCUSSION

Peptide mass fingerprints distinguish genetic differences within microbial species

MALDI-ToF based bacterial profiling at the genera and species levels has provided results that are superior to, less expensive, and less time consuming than those obtained from more conventional approaches such as 16S rRNA sequencing. Such successes can be found across diverse areas of research and across many disciplines such as clinical microbiology, biodefense, food safety and environmental health. Although many studies have reported strain-specific peaks generated by MALDI-ToF MS[11,15,29,30], identification of reliable peaks as strain-specific biomarkers has been hindered by poor profile reproducibility. Further, it appears that the limits of the taxonomic resolution of MALDI-ToF MS profiling at the strain level has been determined in large part by the

nature of the particular bacterium profiled. The more genetically indistinguishable bacteria are, the more challenging their profiling has been[31]. In this study, we show that MALDI-ToF MS offers the possibility to discriminate between genetic mutants or between different strains of the same bacterial species. To investigate the capabilities of MALDI-ToF to distinguish strain-level differences, we analyzed three different gram-negative bacteria.

Two sample preparation methods have been mainly reported for MALDI-ToF MS-based identification and are known as the "Direct Transfer Method" and the "Protein extraction Method"[32] (supplemental table 3.2). The "Direct Transfer Method" consists of picking bacterial material from single colony forming units on a culture plate with a sterile transfer device and applying the material as a thin layer onto the MALDI-ToF target plate. The sample is overlaid with a matrix for MALDI-ToF MS analysis[32]. While this method is simple and rapid, it is inferior in accuracy to the "Protein Extraction Method" because of insufficient cell wall disruption[32]. The "Protein Extraction Method" uses formic acid and acetonitrile to disrupt the cells prior to applying protein extracts directly to the target plate. In our experiments, it was proven that the "Protein Extraction Method" was necessary for the best performance in order to reliably and reproducibly detect strain level differences between the various tested microorganisms (supplemental figure 1.). Moreover, using the protein extraction method allowed for tighter clustering within groups (Supplemental figure 2) and over time with different MALDI-TOF runs / days (supplemental figure 1).

Bacteroides fragilis is a gram-negative, obligate anaerobic, rod-shaped bacterium. It is part of the normal microbiota of the human colon and is generally commensal, but can cause infection if displaced into the bloodstream or surrounding tissue following surgery, disease, or trauma[33]. Enterotoxigenic *B. fragilis* (ETBF) are strains that secrete a 20-kDa heat-labile zinc-dependent metalloprotease toxin termed the *B. fragilis* toxin (BFT). ETBF strains are associated with inflammatory diarrheal disease in children older than 1 year of age and in adults, as well as in inflammatory bowel disease flare-ups and colorectal cancer[29,34–36]. In this experiment, the *bft* positive WT strain of *B. fragilis* (*bft*+) was grown alongside its mutant Δ*bft* on TSA supplemented with 5% defibrinated sheep blood for 48 hours at 37°C anaerobically. Multiple colonies of the WT strain and mutant were tested in four technical replicates each on a MALDI-ToF target following protein extraction and a Principle Component Analysis (PCA) was generated after analyzing the

peptide mass fingerprint profiles (PMF) of each protein extraction. PMFs were generated using Bruker's FlexAnalysis software and the PCA was performed with the help of the MALDIquant package in R. The mean PMF of the biological replicates from each cluster is shown by the larger point in the middle of each ellipse (**Figure 3.1A**). We observed statistically significant differences between the proteomic profiles of the WT strain and Δ*bft* strain in the PCA (**Figure 3.1A**). This experiment clearly demonstrates that the MALDI-ToF MS is a sensitive enough technology to detect genetic differences in highly similar microorganisms, as it could reliably discriminate between an enterotoxigenic *B. fragilis* strain and its mutant (**Figure 3.1A**).

Y. enterocolitica is a gram-negative bacterium that causes food-borne illness when ingested in contaminated meat and water. *Y. enterocolitica* causes a transient infection in immunocompetent adults, but can be lethal to immunocompromised individuals especially young children under the age of five[37,38]. Similar to the *B. fragilis* studies above, Y. *enterocolitica* 8081 and two mutant strains were grown on TSA for 48 hours aerobically at room temperature prior to protein extraction and analysis by MALDI-ToF MS. The MALDI-ToF was able to discriminate three distinct clusters among wild type *Y. enterocolitica* strain (Ye8081), a *yadA* and *yscA* (**Figure 3.1B**).

Escherichia coli is a gram-negative, rod-shaped bacterium commonly found in the lower intestines. While most *E. coli* strains are harmless, pathogenic varieties can cause serious gastroenteritis, urinary tract infections, meningitis or septic shock in humans[30]. Enteropathogenic *E. coli* (EPEC) is a strain of *E. coli* that harbors a virulence factor known as intimin. Intimin is an adhesin that binds host intestinal cells, causing watery diarrhea in those afflicted[39]. In this experiment, EPEC was grown aerobically on TSA for 24 hours at 37°C alongside a non-pathogenic strain of *E. coli*, DH5α. Our data from *B. fragilis* and *Y. enterocolitica*, demonstrated that the MALDI-ToF was able to discriminate the proteomic profile of a bacterial species in which a single gene was ablated. Here, we sought to understand whether the MALDI-ToF could differentiate proteomic signatures of a pathogenic E. coli harboring a number of virulence plasmids and genes from a nonpathogenic laboratory strain. The PCA comparing the PMF of EPEC and DH5α clearly show two distinct populations that cluster by pathogenicity (**Figure 3.1C).** Taken together, these experiments clearly show that MALDI-ToF MS is sensitive enough to detect strain-level differences and pathogenicity, but also to discriminate between bacteria that differ in only in the expression of a single genetic deletion.

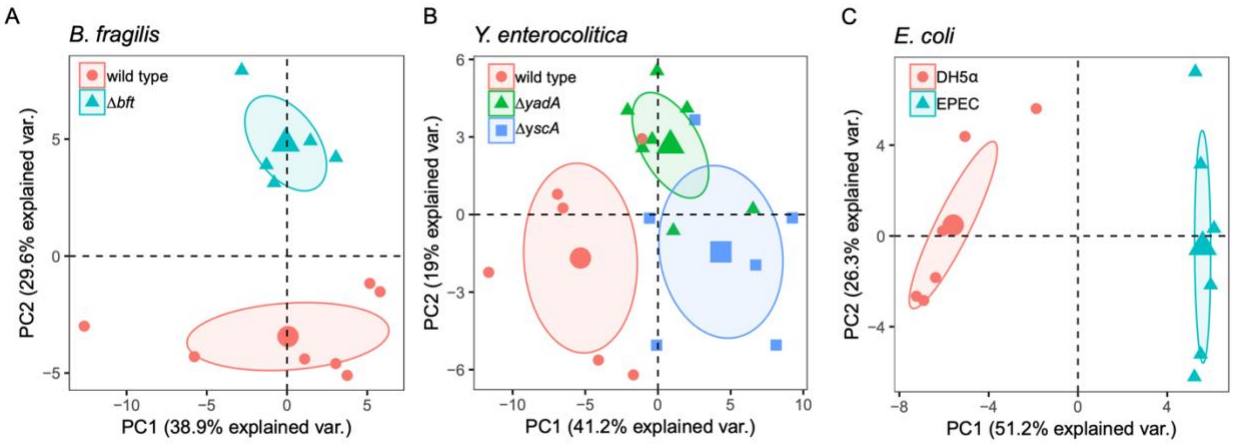


Figure 3.1. Differentiation of bacterial genetic mutants on MALDI-ToF MS

Principal component analysis (PCA) based on peptide mass fingerprint (PMF) profile of various bacterial strains and mutants. A) *Bacteroides fragilis* with and without the enterotoxigenic gene (*bft*) are compared. B) *Yersinia enterocolitica* strain 8081 WT (Ye8081) and mutants with genetic differences in genes *yadA* and *yscA* are compared. C) Enteropathogenic *Escherichia coli* and DH5α strains are compared. Each dot represents the average of 4 technical replicates; the larger point is the mean of the experimental group with a 95% confidence interval drawn on mean. Data for each experiment are from two independent experiments with n=5-7.

PMFs distinguish phenotypic differences within genetically identical strains

The reproducibility of the mass spectra generated by MALDI-ToF MS analysis of proteins from bacterial extracts can be impacted by several experimental factors. Other researchers have discovered that differences in incubation and culturing conditions can alter peak intensities[40] and identification rates[41] (**Table 3.1**). In order to test whether environmental exposures could reproducibly impact the proteome, we decided to look at how various factors such as exposure to oxygen, different nutrient availabilities and temperature impact the PMF of several bacteria.

Condition	Species	Reference
Media type	E. coli	This study
		Arnold et al. 1999[50]
		Valentine et al., 2005[28]

		Reich et al., 2013[51]
		Veenemans et al., 2016[40]
	Y. enterocolitica	Valentine et al., 2005[28]
	B. subtilis	Valentine et al., 2005[28]
	Mycobacteria	Balážová et al., 2015 [52]
	S. aureus	Reich et al., 2013[51]
	Lactobacillus spp.	Šedo et al., 2013[53]
	Acetic Acid bacteria	Wieme et al., 2013[54]
	Nocardia spp.	McTaggart et al., 2018[55]
Growth rate / Age	E. coli	Arnold et al. 1999[50]
		Wunschel et al., 2005[27]
		Reich et al., 2013[51]
		Veenemans et al., 2016[40]
	Mycobacteria	Balážová et al., 2015 [52]
	Beverage spoiling yeasts	Usbeck et al., 2013[56]
	S. aureus	Reich et al., 2013[51]
	Lactobacillus	Šedo et al., 2013[53]
	Rhizobium	Mandal et al., 2007[57]
	Nocardia spp.	McTaggart et al., 2018[55]
Oxygen	L. reuteri	This study
	Beverage spoiling yeasts	Usbeck et al., 2013[56]
Temperature	E. coli	This study
		Wunschel et al., 2005[27]
	Y. enterocolitica	This study
		Wunschel et al., 2005[27]
	Bacillus spp.	Shu and Yang 2017[58]
	Lactobacillus	Šedo et al., 2013[53]
Broth vs agar	Beverage spoiling yeasts	Usbeck et al., 2013[56]
	E. coli	Reich et al., 2013[51]
	S. aureus	Reich et al., 2013[51]

pH	E. coli	Wunschel et al., 2005[27]

Table 3.1. Environmental influences on MALDI-ToF MS analyses

Using the microorganism *Lactobacillus reuteri*, we investigated if MALDI-ToF MS analysis could detect differences in PMF following overnight incubation at 37°C in aerobic and anaerobic environments. *Lactobacillus spp.* are facultative anaerobes and several genes are differently expressed with the presence of oxygen[42]. The data generated by the MALDI-ToF confirmed these previous studies by revealing that if the same starting culture was cultured under aerobic or anaerobic conditions there are significant alterations in the PMF profile (**Figure 3.2A)**. We observed that the two conditions clearly separate along the x-axis with over 50% of the variation explained by PC1 (**Figure 3.2A**). Interestingly, the clustering of *L. reuteri* grown aerobically exhibit more variability in PMFs compared to *L. reuteri* grown anaerobically. These differences may be indicative of the impact of anaerobic stress exerting a more distinct response on the bacteria.

As others have noted, culturing on various media does not alter species identification[28]. For this study, we cultured *E. coli* DH5α on four different nutrient-rich media: TSA, TSA supplemented with 5% blood (TSBA), MacConkey agar (MC), and eosin methylene blue agar (EMB) to test whether the MALDI-ToF could discriminate proteomic differences due to nutrient availability in the E. coli. While *E. coli* grown on MC and EMB agar had overlapping PMF profiles and clustered together on the PCA, there were clearly different clusters for *E. coli* grown on TSA and TSBA. (**Figure 3.2B**). The overlapping of MC and EMB-grown *E. coli* is likely due to the induction of lactose fermentation genes by *E. coli* when it is grown on these two types of agars. While the differences between MC and EMB agars was not enough to induce differential proteomic profiles of *E. coli*, the presence of 5% defibrinated sheep blood in the TSA was able to trigger the expression of a different proteomic profile in *E. coli* grown on TSA without blood (**Figure 3.2B**). Thus, MALDI-ToF MS has the capacity to significantly discriminate between bacteria grown on different media, if these media provide nutrients that promote the upregulation of distinct metabolic pathways necessary for nutrient utilization.

Next, we tested whether temperature affects the PMF profiles of microorganisms. *Y. enterocolitica* is well adapted for survival and proliferation at room temperature (RT) but induces its virulence factors when exposed to host temperatures of 37°C[43]. These metabolic adaptations are clearly captured when we analyze *Y. enterocolitica* grown at two different temperatures on the MALDI-ToF MS, with *Y. enterocolitica* grown at RT clustering distinctly from the 37°C treated group (**Figure 3.2C**). The two clusters for *Y. enterocolitica* are clearly separated on the x-axis with over 60% of the PMF variation explained on PC1, demonstrating a clear distinction in proteome expressions due to temperature conditions. The individual PMFs of *Y. enterocolitica* grown at room temperature are also more tightly clustered than the *Y. enterocolitica* grown at 37°C. This variable clustering could be indicative of differences in protein expression by *Y. enterocolitica* when grown in virulence-inducing temperatures. Previous studies using MALDI-ToF MS on *Y. enterocolitica* point out that temperature and media both play a role in PMF intensities with inference that temperature may play a bigger role[27,28]. Whereas Wunschel and collegues[27] identified that *Y. enterocolitica* grown at 37°C has similar PMFs despite the media it is grown on, we demonstrate that the MALDI-ToF can reliably distinguish between distinct temperature conditions.

Next we tested whether MALDI ToF MS could discriminate between bacteria =heat shocked and those that were not. Here, liquid cultures of *E. coli* were incubated at 37°C or 50°C for 15 minutes or 30 minutes. As expected, *E. coli* exposed to 50°C had a different PMF profile and different clustering than *E. coli* maintained at 37°C, likely due to the induction and translation of bacterial heat shock proteins[44]. Importantly, these differences became even more significant when the exposure time was increased from 15 to 30 minutes indicating that MALDI-Tof can even kinetic differences in proteomic signatures (**Figure 3.2D**).

All together these results demonstrate that using PMF profiles generated by the MALDI-ToF allows the reproducible discrimination of different metabolic states in the same species of bacteria exposed to different environmental conditions. While other studies have shown that different culturing conditions impact the identification of bacteria by MALDI-ToF, they failed to assess whether different environmental exposures could be identified within the same species of bacteria. The work presented here builds upon these previous studies and has identified a novel and important function for this technology. The ability to easily differentiate the direct effects of

environmental conditions on a species of bacteria using MALDI-ToF presents a new way to screen proteomic profiles of bacteria from different biological samples and to identify specific metabolic programs.

To further understand the extent of the MALDI-ToF MS technique on discerning environmental influences on bacteria, we examined whether the same bacteria exposed to two different environmental stressors, but subsequently plated on the same nutrient agar would leave a proteomic imprint that could be detected by MALDI ToF. For this experiment, *E. coli* was grown in liquid culture at 37°C or 50°C for 2 hours and both samples were collected and plated on TSA for 24 hours. Our previous experiment measured the direct environmental impact of temperature on *E. coli* (**Figure 3.2D**), but this experiment was to determine whether the MALDI-ToF could discriminate *E. coli* exposed to different temperatures when all the samples were collected, plated on the same nutrient rich agar and cultured for 24 hours at 37°C. Surprisingly, the PMFs generated by the MALDI-ToF produced two distinct clusters on the PCA corresponding to the 37°C and 50°C group and indicate that the MALDI-ToF was able to detect proteomic imprints of temperature on *E. coli* despite being harvested and plated on the same nutrient agar (**Figure 3.2E-F).** These data demonstrate that the MALDI-ToF can identify environmental signatures despite the removal of stressor or condition and subsequent growth on the same nutrient agar. These results are contrary to current dogma in the field that assume functional, proteomic or metabolic differences of bacterial isolates from both environmental and human tissue are abolished once the bacteria are plated on the same nutrient agar. Instead, our data indicates that prior environmental exposures leave a proteomic imprint that is maintained despite re-culturing. Our data reveal that this proteomic imprint can be discerned using the MALDI-ToF.

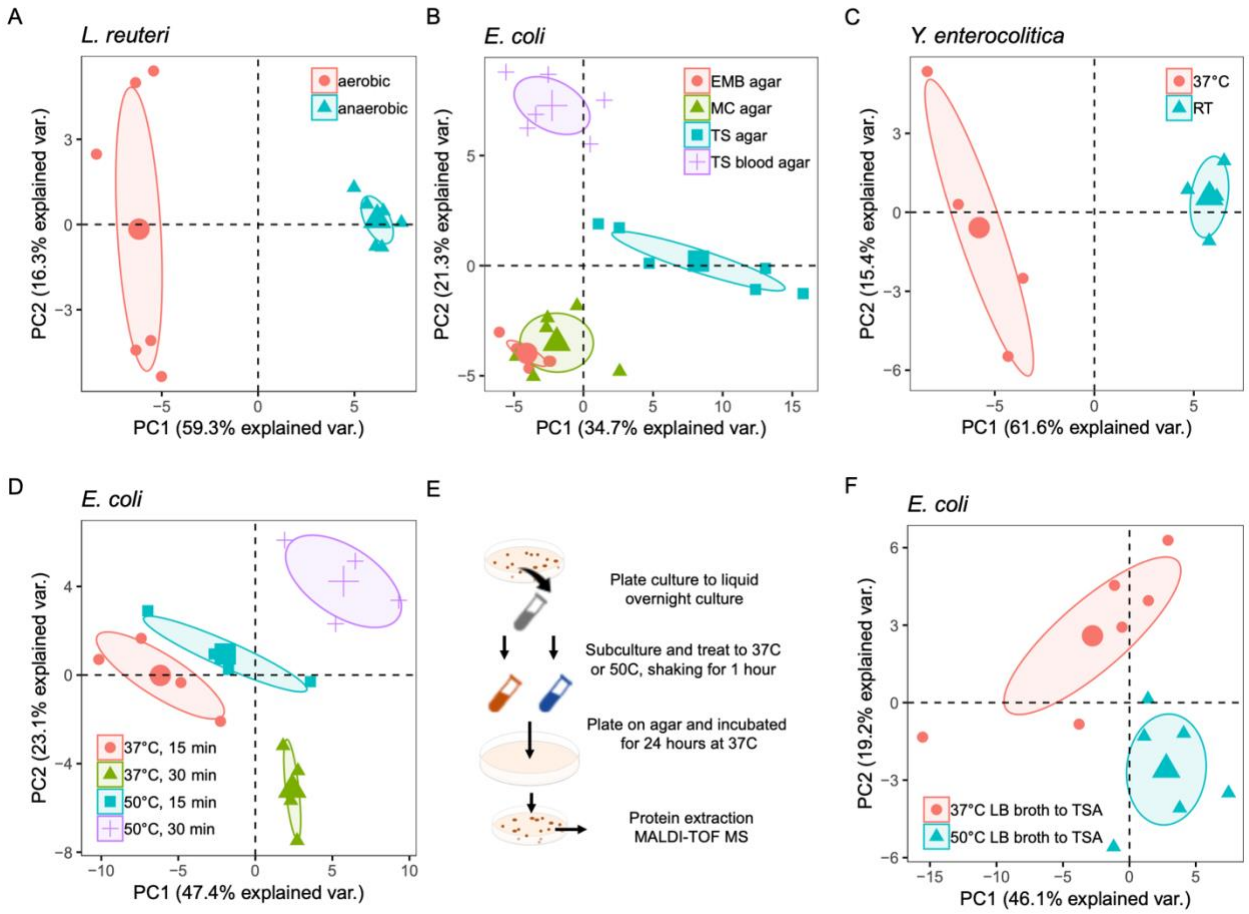


Figure 3.2. Differentiation of bacterial biological environment on MALDI-ToF MS

PCA based on PMF profile of various bacterial species cultured in different growth conditions. A) *Lactobacillus reuteri* plated on TSA and incubated overnight at 37°C in aerobic and anaerobic conditions is compared. B) DH5α *Escherichia coli* strain grown on 4 different types of Media: TSA (TS agar), TSA with 5% sheep blood (TS blood agar), MacConkey agar (MC agar), and eosin methylene blue agar (EMB agar) is compared. C) *Y. enterocolitica* grown at room temperature (RT) or 37°C for 48 hours on TSA. D) DH5α grown in liquid culture at 37°C and 50°C for 15 and 30 min while shaking. E) Diagram and F) PCA of DH5α grown in liquid culture at 37°C and 50°C for 1 hour and then plated to observe residual effects of different environmental conditions. Each dot represents the average of 4 technical replicates; the larger point is the mean of the experimental group with a 95% confidence interval drawn on mean. Data for each experiment are from two independent experiments with n=5-7.

Our data show that MALDI-ToF analysis was able to distinguish between two growth states of *Y. enterocolitica*, the more virulent state induced by temperatures of 37°C and the less virulent state which is maintained at room temperature. These two states differ dramatically due to the induction of the virulence plasmid and expression of the type III secretion system making detection by MALDI-ToF fairly straightforward. This led us to question whether MALDI-ToF could detect differences in *Y. enterocolitica* which may have already been expressing its virulence genes but subsequently exposed to different subcellular compartments. During the course of a *Y. enterocolitica* infection the bacteria will pass from the lumen of the small intestine into the Peyer's patches of the lamina propria and the mesenteric lymph nodes[45]. In order for *Y. enterocolitica* to survive in these three very different environments, different genes and metabolic states are necessary for *Y. enterocolitica's* survival. To test the ability of the MALDI-ToF to distinguish *Y. enterocolitica* isolated from different environments, we inoculated a human intestinal epithelial cell line, SW620, with *Y. enterocolitica* grown at 37C to induce its virulence plasmid. At the end of the *in vitro* stimulation both extracellular and intracellular *Y. enterocolitica* were harvested on onto the same nutrient agar (TSA) for 48 hours. Bacterial cells were proteomically prepared as described above and the PMFs generated by the MALDI-ToF were compared. PMFs of untreated *Y. enterocolitica*, *Y. enterocolitica* applied to SW620s for 1 hour, or applied to SW620s for 1 hour and treated with gentamicin to remove extracellular bacteria were compared (**Figure 3.3A**). After samples were collected, *Y. enterocolitica* was plated on TSA and incubated at room temperature for 48 hours. The analysis of the PMF's of intracellular and extracellular *Y. enterocolitica* generated by the MALDI show three distinct clusters. Intracellular *Y. enterocolitica* clustered together on the left of the PCA significantly distinct from the extracellular *Y. enterocolitica* which clustered on the right of the PCA. (Figure 3B). The control samples containing *Y. enterocolitica* grown for 1 and 2 hours in the absence of SW620 clustered intermediately, but distinctly, from the intra- and extra- cellular *Y. enterocolitica* (**Figure 3.3B**). Using the PMFs and PCA a dendrogram was created showing the hierarchical clustering of the samples. Both the intra- and extra- cellular *Y. enterocolitica* samples clustered by group but remained distinct from each other indicating clearly different proteomic signatures between the samples run in each experimental group (Figure 3C). These data also confirm the experiments above and illustrate that the MALDI-ToF can detect distinct proteomic imprints of samples isolated from two different environmental niches despite being re-cultured on TSA plates (Figure 3B-C). These data indicate that the proteomes of *Y. enterocolitica* differ greatly between intra- and extra-cellular lifestyles and the PMFs provide a

preliminary blueprint of distinct and shared peaks that correspond to specific proteins that can be used to generate more specific hypothesis.

Taken together, these data using *E. coli* and *Y. enterocolitica* demonstrate that the MALDI-ToF can detect differences in metabolic and virulence states, even after the bacteria are removed from the initial environment and re-plated on the same nutrient agar. This is especially important for trying to assess phenotypic differences of clinical isolates or samples that cannot be readily grown in its natural environment. These methods could be applied to situations in which metabolic or phenotypic difference of experimental groups can be compared against known culturing conditions. It can also be used as a method of screening of known bacterial phenotypes.

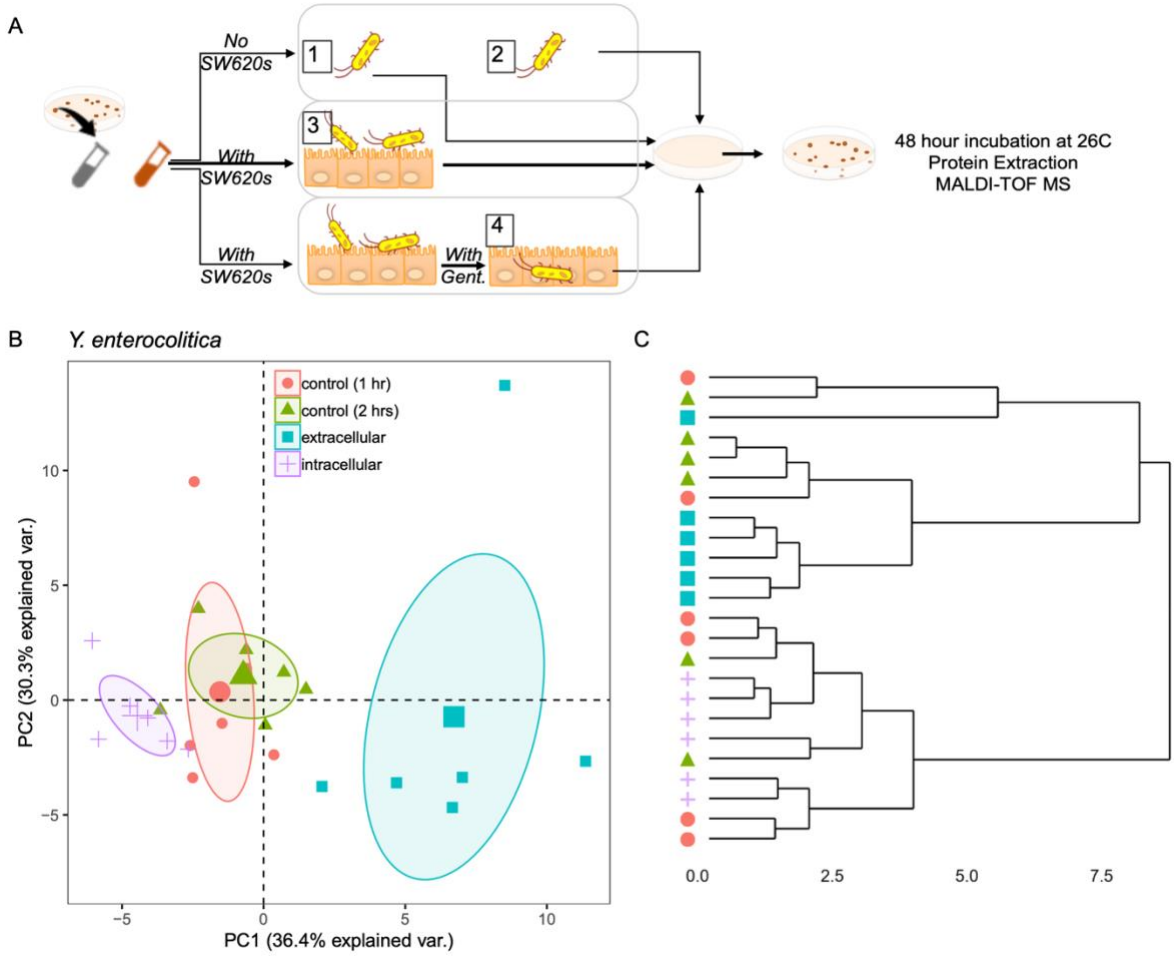


Figure 3.3. PMF analysis of Y. enterocolitica virulence states

A) Schematic of protocol. B) PCA and C) dendrogram of *Y. enterocolitica* cultured 48 hours at room temperature on TSA following adhesion and invasion assay with SW620 cells. Each dot

represents the average of 4 technical replicates; the larger point is the mean of the experimental group with a 95% confidence interval drawn on mean. Data for each experiment are from two independent experiments with n=6.

PMFs distinguish between fresh and frozen clinical isolates of the same species

A very common use of the MALDI-ToF in clinical studies is to identify isolates removed from biological samples for diagnostic purposes. The experiments described above used the protein extraction method in a well-controlled environment with well characterized strains, the use of MALDI-ToF in clinical settings is less likely to resort to this time-consuming method. In a clinical setting, using the protein extraction method is time consuming and not all samples are easily processed immediately. Additionally, freezing samples prior to culturing can reduce the number of bacteria recovered, skew the diversity of bacteria, and alter community composition[46,47]. Therefore, these next experiments evaluated how frozen versus fresh conditions altered the PMFs of clinical samples using the "Extended Direct Transfer" method as described above.

Using fresh human fecal samples, we tested whether culturing fresh or frozen microbiota samples altered the phenotype of clinical isolates (**Figure 3.4A**). In these experiments, whole fecal samples were plated aerobically on TSBA for 48 hours. Single colonies were then processed using the extended direct colony method. Bacteria identified as the same species were then compiled and examined retrospectively.

Within our clinical samples, the MALDI-ToF identified *Citrobacter sedlakii*, *E. coli* and *Bacteroides ovatus* from both fresh and frozen cultures. Despite distinct clustering between fresh and frozen isolates of *C. sedlakii* and *E. coli*, the clustering was not statistically significant as there was overlap seen in the PCA and dendrogram (**Figure 3.4B-C**). Freezing seemed to have a more profound impact on *C. sedlakii* compared to *E. coli* as these isolates appeared to have little overlap in PMFs (**Figure 3.4B**). In contrast, *E. coli* isolates from fresh and frozen stool had more similarities suggesting that the response to freezing may be bacteria dependent (**Figure 3.4C**). These experiments provide a rationale for why sample collection during microbiome studies must be standardized.

We next compared the MALDI-ToF analysis of fresh and frozen *B. ovatus* taken from two different patients. As expected, *B. ovatus* clustering was patient dependent, likely due to strain variation of *B. ovatus* or exposure of *B. ovatus* to microenvironments that contribute to genetic diversity via phage and horizontal gene transfer (**Figure 3.4D**). Interestingly, fresh or frozen status had little impact on the clustering of *B. ovatus* from patient 2, while patient 4 had high variability of PMFs within each of the fresh and frozen groups. The different sensitivity between patient 2 and patient 4 *B. ovatus* isolates from fresh and frozen stool may be due to the presence of more temperature sensitive genes in the isolates from patient 2 and indicate that freezing stool may alter the PMF generated by the MALDI-ToF in a bacteria-dependent manner.

Our results using clinical isolates derived from a mixed microbiota sample reveal that fresh versus frozen culturing techniques may alter the proteotyping of microorganisms. Further study is needed to analyze the freezing effects. This clinical experiment was limited by MALDI-ToF processing, lack of technical replicates, and difference in MALDI-ToF sampling days. These results exemplify a need for rigorous sample preparation and analysis. Given our ability to proteotype bacteria by various environmental factors, there is huge potential in analyzing phenotypic differences amongst clinical samples. Should this methodology be adapted in the clinic there are numerous applications, such as distinguishing proteotypes from inflammatory, nutrient-rich, nutrient-depleted, or competitive microenvironments.

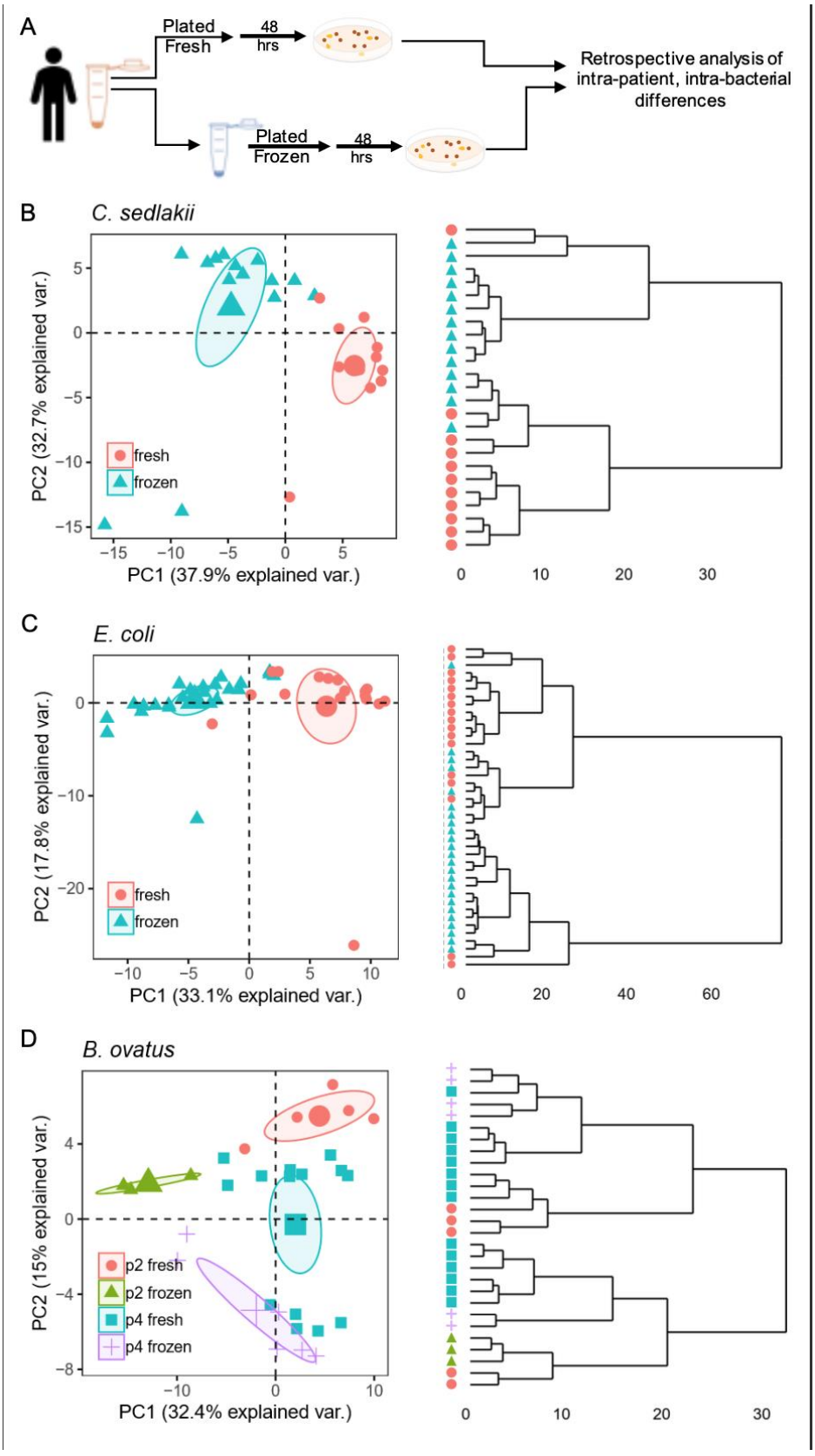


Figure 3.4. PMF analysis of fresh and frozen clinical isolates

A) Schematic of clinical sample analysis. PCA and dendrogram of fresh and frozen PMFs from B) *C. sedlakii*, C) *E. coli* and D) *B. ovatus* isolates identified on aerobic TSA with 5% sheep blood. Isolates proceeded via extended direct transfer. Each dot resembles a biological replicate. Ellipse drawn on the 95% confidence interval of the mean.

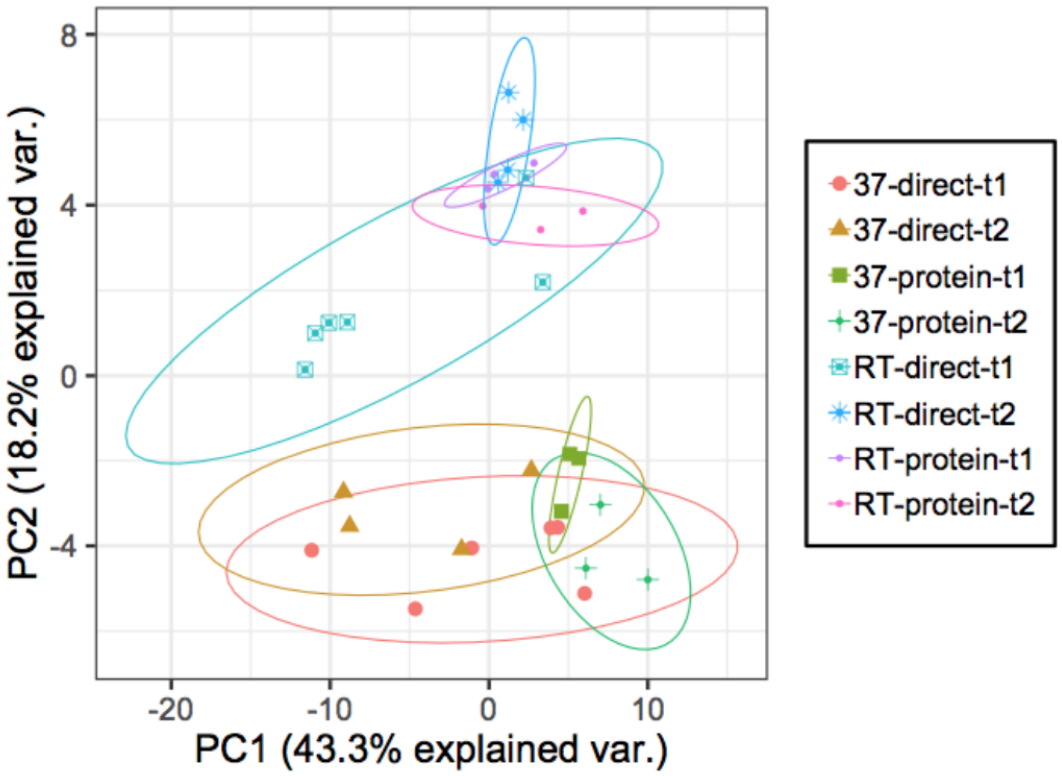
3.5. CONCLUSION

The emergence of MALDI-ToF MS has reinvigorated microbial identification over the past decade. By creating proteomic signatures through PMFs of microorganisms, species and subspecies identification is possible via proteotyping. Routine use of MALDI-ToF MS as a diagnostic tool is favorable due to its ease of use, time and cost efficiency, and robust sensitivity. Our current study demonstrates a new function by exploiting the unique proteomic signatures and comparative analysis afforded by the MALDI-ToF MS bioinformatic packages. While proteotyping of different physiological and metabolic states of microorganisms through MALDI-ToF has been previously suggested by others[48,49] and variations in culturing conditions have been extensively tested for impact on species identification (Table 3.1), we test the ability to distinguish between multiple environmental conditions within the same microorganism. Environmental effects of oxygen, temperature and nutrients can all be easily distinguished. As a proof-of-principle, we also demonstrate that PMF signatures can differentiate by stressors even when microorganisms are removed from the initial environment. We further exemplify how the MALDI-ToF MS technique can separate different virulent states of *Y. enterocolitica* with the SW620 adhesion assay.

One limitation of proteotyping of metabolic and functional states include lack of PMF peak to protein identification readily available through the Bruker software. While publicly accessible databases for peptide/protein identification and assignment exist, such as Mascot (www.matrixscience.com), there still remains a need to develop a database of peak information that can be more readily incorporated with MALDI-ToF MS data. Another limitation is addressed with our clinical data in which rigorous protein extraction and technical replicates may not be easily incorporated in a clinical setting. While our data inconclusively delineate fresh versus frozen isolates, these experiments need to be repeated with a wide variety of microorganisms.

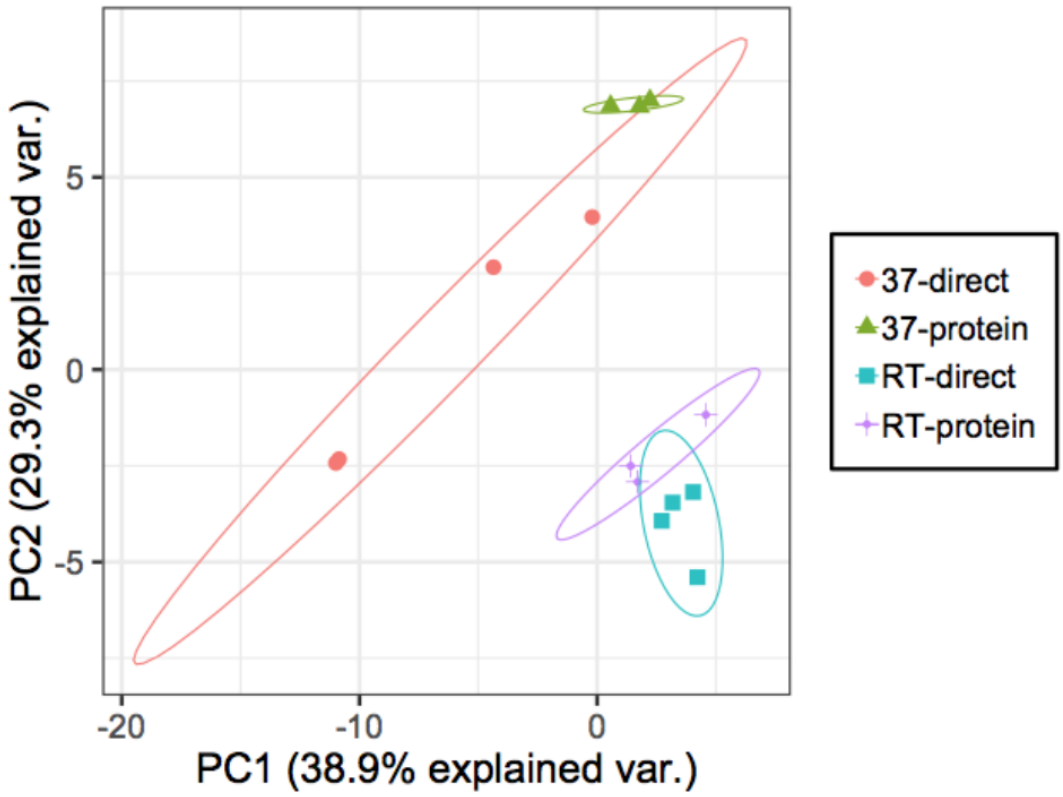
Altogether, we demonstrate a new and powerful use for MALDI-ToF MS to screen for physiological changes from direct and indirect environmental conditions. In conjunction with transcriptomic data and validation through PCR or LC/MS, this technique could provide a rapid and efficient way to screen for metabolic functions and physiological states of microorganisms. This methodology can be especially useful for analyzing phenotypic changes of bacteria cultured from previously unique environments such as an inflamed gastrointestinal tract or progression of different virulent states. In addition to the original function of species identification, the MALDI-ToF MS provides a plethora of data that has yet to be fully explored and exploited. Pairing these data with peptide/protein identity assignment could ultimately constitute a strong framework for the identification of microbiome-based biomarkers and therapeutic targets in various disease states.

3.6. SUPPLEMENTAL MATERIAL



Supplemental Figure 3.1. PMF analysis of Y. enterocolitica with extended direct method of processing vs the protein extraction method

Ye8081 grown on TSA and incubated at either room temperature or at 37°C. Colonies were then processed by the extended direct method (direct) or the protein extraction method (protein). Protein samples (n=3/group) were done with 8 technical replicates. Direct samples (n=4-6/group) had no technical replicates. Ellipse drawn on the 95% confidence interval of the cluster. t1 and t2 indicate two different independent experiments on different dates and different MALDI-ToF targets.



Supplemental Figure 3.2. PMF analysis of Y. enterocolitica grown at different temperatures with extended direct method of processing vs the protein extraction method

Ye8081 grown on TSA and incubated at either room temperature or at 37°C. Colonies were then processed by the extended direct method (direct) or the protein extraction method (protein). Protein samples (n=3/group) were done with 8 technical replicates. Direct samples (n=4/group) had no technical replicates. Ellipse drawn on the 95% confidence interval of the cluster.

Species	Strain	Source	Reference
B. fragilis	86-5443-2-2	piglet (*bft-2)*	Lyle Myers
B. fragilis	∆86-5443-2-2	piglet (∆ *bft-2)*	Lyle Myers
Y. enterocolitica	Ye8081 biotype 1B serotype O:8	Blood, Human	ATCC® 27729™/PB Carter, CF Varga
Y. enterocolitica	Δ8081	Human (ΔYadA)	ATCC® 27729™/PB Carter, CF Varga
Y. enterocolitica	Δ8081	Human (ΔYscA)	ATCC® 27729™/PB Carter, CF Varga
E. coli	CDC B170-Serotype O111	Feces, Human	ATCC® 43887™/CDC
E. coli	DH5 alpha	N/A	N/A
L. reuteri	DSM 20016	Feces, Human	ATCC® 23272™/ PA Hansen
B. ovatus	N/A	Feces, Human	This study
C. sedlakii	N/A	Feces, Human	This study
E. coli	N/A	Feces, Human	This study

Supplemental Table 3.1. Bacterial species and strains used in this study

	Pros	Cons
Extended direct transfer method	• Reliable identification • Fast • Less reagents consumption	• More variations in PMFs in the same sample • Less detailed peptide profiling
Protein extraction method	• Reliable identification • Less variation in PMFs in the same sample • More detailed peptide profiling • Identifies phenotypic differences	• More time consuming

Supplemental Table 3.2. Pros and cons of extended direct transfer method vs protein extraction method

3.7. REFERENCES

1. Bardhan, K. & Liu, K. Epigenetics and Colorectal Cancer Pathogenesis. *Cancers (Basel)* **5**, 676–713 (2013).
2. Singhal, N., Kumar, M., Kanaujia, P. K. & Virdi, J. S. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front. Microbiol.* **6**, (2015).

3. Bhavsar, S. M., Dingle, T. C. & Hamula, C. L. The impact of blood culture identification by MALDI-TOF MS on the antimicrobial management of pediatric patients. *Diagnostic Microbiology and Infectious Disease* **92**, 220–225 (2018).
4. Khan, S. *et al.* MALDI-TOF MS in Adult Inpatients with Bloodstream Infections: Pre- and Post-intervention Study. *Open Forum Infectious Diseases* **4**, S626–S626 (2017).
5. Haiko, J., Savolainen, L. E., Hilla, R. & Pätäri-Sampo, A. Identification of urinary tract pathogens after 3-hours urine culture by MALDI-TOF mass spectrometry. *Journal of Microbiological Methods* **129**, 81–84 (2016).
6. Ferreira, L. *et al.* Direct Identification of Urinary Tract Pathogens from Urine Samples by Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. *Journal of Clinical Microbiology* **48**, 2110–2115 (2010).
7. Íñigo, M. *et al.* Direct Identification of Urinary Tract Pathogens from Urine Samples, Combining Urine Screening Methods and Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry. *J. Clin. Microbiol.* **54**, 988–993 (2016).
8. Svarrer, C. W. & Uldum, S. A. The occurrence of Legionella species other than Legionella pneumophila in clinical and environmental samples in Denmark identified by mip gene sequencing and matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clinical Microbiology and Infection* **18**, 1004–1009 (2012).
9. Samb-Ba, B. *et al.* MALDI-TOF Identification of the Human Gut Microbiome in People with and without Diarrhea in Senegal. *PLoS ONE* **9**, e87419 (2014).
10. Rizzardi, K. & Åkerlund, T. High Molecular Weight Typing with MALDI-TOF MS - A Novel Method for Rapid Typing of Clostridium difficile. *PLoS ONE* **10**, e0122457 (2015).

11. Wang, H.-Y. *et al.* A new scheme for strain typing of methicillin-resistant Staphylococcus aureus on the basis of matrix-assisted laser desorption ionization time-of-flight mass spectrometry by using machine learning approach. *PLoS ONE* **13**, e0194289 (2018).
12. Wolters, M. *et al.* MALDI-TOF MS fingerprinting allows for discrimination of major methicillin-resistant Staphylococcus aureus lineages. *International Journal of Medical Microbiology* **301**, 64–68 (2011).
13. Manukumar, H. M. & Umesha, S. MALDI-TOF-MS based identification and molecular characterization of food associated methicillin-resistant Staphylococcus aureus. *Sci Rep* **7**, 11414 (2017).
14. Nisa, S. *et al.* Combining MALDI-TOF and genomics in the study of methicillin resistant and multidrug resistant Staphylococcus pseudintermedius in New Zealand. *Sci Rep* **9**, 1271 (2019).
15. Christner, M. *et al.* Rapid MALDI-TOF Mass Spectrometry Strain Typing during a Large Outbreak of Shiga-Toxigenic Escherichia coli. *PLoS ONE* **9**, e101924 (2014).
16. Alanio, A. *et al.* Matrix-assisted laser desorption ionization time-of-flight mass spectrometry for fast and accurate identification of clinically relevant Aspergillus species. *Clinical Microbiology and Infection* **17**, 750–755 (2011).
17. Tenover, F. C., Arbeit, R. D. & Goering, R. V. How to select and interpret molecular strain typing methods for epidemiological studies of bacterial infections: a review for healthcare epidemiologists. Molecular Typing Working Group of the Society for Healthcare Epidemiology of America. *Infect Control Hosp Epidemiol* **18**, 426–439 (1997).
18. Parizad, E. G., Parizad, E. G. & Valizadeh, A. The Application of Pulsed Field Gel Electrophoresis in Clinical Studies. *J Clin Diagn Res* **10**, DE01–DE04 (2016).

19. Killgore, G. *et al.* Comparison of seven techniques for typing international epidemic strains of Clostridium difficile: restriction endonuclease analysis, pulsed-field gel electrophoresis, PCR-ribotyping, multilocus sequence typing, multilocus variable-number tandem-repeat analysis, amplified fragment length polymorphism, and surface layer protein A gene sequence typing. *J. Clin. Microbiol.* **46**, 431–437 (2008).
20. de Bruijn, F. J. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of Rhizobium meliloti isolates and other soil bacteria. *Appl. Environ. Microbiol.* **58**, 2180–2187 (1992).
21. Parolo, C. C. F. *et al.* Genetic diversity of Lactobacillus paracasei isolated from in situ human oral biofilms. *J. Appl. Microbiol.* **111**, 105–113 (2011).
22. Balloux, F. *et al.* From Theory to Practice: Translating Whole-Genome Sequencing (WGS) into the Clinic. *Trends Microbiol* **26**, 1035–1048 (2018).
23. Schürch, A. C., Arredondo-Alonso, S., Willems, R. J. L. & Goering, R. V. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches. *Clinical Microbiology and Infection* **24**, 350–354 (2018).
24. Murray, P. R. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: usefulness for taxonomy and epidemiology. *Clin. Microbiol. Infect.* **16**, 1626–1630 (2010).
25. Elinav, E., Garrett, W. S., Trinchieri, G. & Wargo, J. The cancer microbiome. *Nat Rev Cancer* **19**, 371–376 (2019).
26. Lasch, P. *et al.* Insufficient discriminatory power of MALDI-TOF mass spectrometry for typing of Enterococcus faecium and Staphylococcus aureus isolates. *Journal of Microbiological Methods* **100**, 58–69 (2014).

27. Wunschel, D. S. *et al.* Effects of varied pH, growth rate and temperature using controlled fermentation and batch culture on Matrix Assisted Laser Desorption/Ionization whole cell protein fingerprints. *Journal of Microbiological Methods* **62**, 259–271 (2005).
28. Valentine, N., Wunschel, S., Wunschel, D., Petersen, C. & Wahl, K. Effect of Culture Conditions on Microorganism Identification by Matrix-Assisted Laser Desorption Ionization Mass Spectrometry. *Applied and Environmental Microbiology* **71**, 58–64 (2005).
29. Rashidan, M. *et al.* Detection of B. fragilis group and diversity of bft enterotoxin and antibiotic resistance markers cepA , cfiA and nim among intestinal Bacteroides fragilis strains in patients with inflammatory bowel disease. *Anaerobe* **50**, 93–100 (2018).
30. Chart, H., Smith, H. R., La Ragione, R. M. & Woodward, M. J. An investigation into the pathogenic properties of Escherichia coli strains BLR, BL21, DH5alpha and EQ1. *J. Appl. Microbiol.* **89**, 1048–1058 (2000).
31. Ghyselinck, J., Van Hoorde, K., Hoste, B., Heylen, K. & De Vos, P. Evaluation of MALDI-TOF MS as a tool for high-throughput dereplication. *Journal of Microbiological Methods* **86**, 327–336 (2011).
32. Alatoom, A. A., Cunningham, S. A., Ihde, S. M., Mandrekar, J. & Patel, R. Comparison of Direct Colony Method versus Extraction Method for Identification of Gram-Positive Cocci by Use of Bruker Biotyper Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. *Journal of Clinical Microbiology* **49**, 2868–2873 (2011).
33. Wexler, H. M. Bacteroides: the Good, the Bad, and the Nitty-Gritty. *Clinical Microbiology Reviews* **20**, 593–621 (2007).
34. Zamani, S. *et al.* Detection of enterotoxigenic Bacteroides fragilis in patients with ulcerative colitis. *Gut Pathog* **9**, 53 (2017).

35. Purcell, R. V. *et al.* Colonization with enterotoxigenic Bacteroides fragilis is associated with early-stage colorectal neoplasia. *PLoS ONE* **12**, e0171602 (2017).

36. Goodwin, A. C. *et al.* Polyamine catabolism contributes to enterotoxigenic Bacteroides fragilis-induced colon tumorigenesis. *Proceedings of the National Academy of Sciences* **108**, 15354–15359 (2011).

37. Lee, L. A. *et al.* Yersinia Enterocolitica O:3: An Emerging Cause of Pediatric Gastroenteritis in the United States. *Journal of Infectious Diseases* **163**, 660–663 (1991).

38. Ray, S. M. *et al.* Population-Based Surveillance for *Yersinia enterocolitica* Infections in FoodNet Sites, 1996–1999: Higher Risk of Disease in Infants and Minority Populations. *CLIN INFECT DIS* **38**, S181–S189 (2004).

39. Clarke, S. C., Haigh, R. D., Freestone, P. P. E. & Williams, P. H. Virulence of Enteropathogenic Escherichia coli, a Global Pathogen. *Clinical Microbiology Reviews* **16**, 365–378 (2003).

40. Veenemans, J. *et al.* Comparison of MALDI-TOF MS and AFLP for strain typing of ESBL-producing Escherichia coli. *Eur J Clin Microbiol Infect Dis* **35**, 829–838 (2016).

41. Vithanage, N. R. *et al.* Species-Level Discrimination of Psychrotrophic Pathogenic and Spoilage Gram-Negative Raw Milk Isolates Using a Combined MALDI-TOF MS Proteomics–Bioinformatics-based Approach. *J. Proteome Res.* **16**, 2188–2203 (2017).

42. Zotta, T., Parente, E. & Ricciardi, A. Aerobic metabolism in the genus *Lactobacillus* : impact on stress response and potential applications in the food industry. *J Appl Microbiol* **122**, 857–869 (2017).

43. Straley, S. C. & Perry, R. D. Environmental modulation of gene expression and pathogenesis in Yersinia. *Trends Microbiol.* **3**, 310–317 (1995).

44. Tomoyasu, T. *et al.* Escherichia coli FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32. *EMBO J.* **14**, 2551–2560 (1995).

45. Bent, Z. W. *et al.* Transcriptomic Analysis of Yersinia enterocolitica Biovar 1B Infecting Murine Macrophages Reveals New Mechanisms of Extracellular and Intracellular Survival. *Infect. Immun.* **83**, 2672–2685 (2015).

46. Lau, J. T. *et al.* Capturing the diversity of the human gut microbiota through culture-enriched molecular profiling. *Genome Med* **8**, 72 (2016).

47. Wu, W.-K. *et al.* Optimization of fecal sample processing for microbiome study — The journey from bathroom to bench. *Journal of the Formosan Medical Association* **118**, 545–555 (2019).

48. Grenga, L., Pible, O. & Armengaud, J. Pathogen proteotyping: A rapidly developing application of mass spectrometry to address clinical concerns. *Clinical Mass Spectrometry* S2376999818300503 (2019) doi:10.1016/j.clinms.2019.04.004.

49. Karlsson, R. *et al.* Proteotyping: Proteomic characterization, classification and identification of microorganisms – A prospectus. *Systematic and Applied Microbiology* **38**, 246–257 (2015).

50. Arnold, R. J., Karty, J. A., Ellington, A. D. & Reilly, J. P. Monitoring the Growth of a Bacteria Culture by MALDI-MS of Whole Cells. *Anal. Chem.* **71**, 1990–1996 (1999).

51. Reich, M. Species Identification of Bacteria and Fungi from Solid and Liquid Culture Media by MALDI-TOF Mass Spectrometry. *J Bacteriol Parasitol* **01**, (2013).

52. Balážová, T. *et al.* The influence of culture conditions on the identification of *Mycobacterium* species by MALDI-TOF MS profiling. *FEMS Microbiol Lett* **353**, 77–84 (2014).

53. Šedo, O., Vávrová, A., Vad'urová, M., Tvrzová, L. & Zdráhal, Z. The influence of growth conditions on strain differentiation within the *Lactobacillus acidophilus* group using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry profiling: Growth conditions influence strain differentiation using MALDI-TOF MS. *Rapid Commun. Mass Spectrom.* **27**, 2729–2736 (2013).

54. Wieme, A. D. *et al.* Effects of Growth Medium on Matrix-Assisted Laser Desorption–Ionization Time of Flight Mass Spectra: a Case Study of Acetic Acid Bacteria. *Appl. Environ. Microbiol.* **80**, 1528–1538 (2014).

55. McTaggart, L. R., Chen, Y., Poopalarajah, R. & Kus, J. V. Incubation time and culture media impact success of identification of Nocardia spp. by MALDI-ToF mass spectrometry. *Diagnostic Microbiology and Infectious Disease* **92**, 270–274 (2018).

56. Usbeck, J. C., Kern, C. C., Vogel, R. F. & Behr, J. Optimization of experimental and modelling parameters for the differentiation of beverage spoiling yeasts by Matrix-Assisted-Laser-Desorption/Ionization–Time-of-Flight Mass Spectrometry (MALDI–TOF MS) in response to varying growth conditions. *Food Microbiology* **36**, 379–387 (2013).

57. Mandal, S. M., Pati, B. R., Ghosh, A. K. & Das, A. K. Letter: Influence of Experimental Parameters on Identification of Whole Cell *Rhizobium* by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Eur J Mass Spectrom (Chichester)* **13**, 165–171 (2007).

58. Shu, L.-J. & Yang, Y.-L. Bacillus Classification Based on Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry—Effects of Culture Conditions. *Sci Rep* **7**, 15546 (2017).

4. TLR6 Signaling Prevents Inflammation and Impacts Composition of the Microbiota During Inflammation-Induced Colorectal Cancer

Adapted from:
Jee-Hyun Kim, Melissa Kordahi, Denise Chac, and R. William DePaolo. TLR6 signaling prevents inflammation and impacts composition of the microbiota during inflammation-induced colorectal cancer. Cancer Prevention Research, 2019**.** doi: 10.1158/1940-6207.CAPR-19-0286

4.1. ABSTRACT

Tightly regulated immune responses must occur in the intestine in order to avoid unwanted inflammation, which may cause chronic sequela and leading to diseases such as colorectal cancer. Toll-like receptors play an important role in preventing aberrant immune responses in the intestine by sensing endogenous commensal microbiota and delivering important regulatory signals to the tissue. However, the role that specific innate receptors may play in the development of chronic inflammation and their impact on the composition of the colonic microbiota is not well understood. Using a model of inflammation induced colorectal cancer, we found that Lactobacillus species are lost more quickly in WT mice than TLR6-deficient mice resulting in overall differences in bacterial compositions. Despite the longer retention of Lactobacillus, the TLR6-deficient mice presented with more tumors and a worse overall outcome. Restoration of the lost Lactobacillus species suppressed inflammation, reduced tumor number and prevented change in the abundance of Proteobacteria only when given to WT mice, indicating the effect of these Lactobacillus are TLR6-dependent. We found that the TLR6-dependent effects of Lactobacillus could be dissociated from one another via the involvement of IL-10, which was necessary to dampen the inflammatory microenvironment, but had no effect on bacterial composition. Altogether, these data suggest that innate immune signals can shape the composition of the microbiota under chronic inflammatory conditions, bias the cytokine milieu of the tissue microenvironment, and influence the response to microbiota-associated therapies.

4.2. INTRODUCTION

Colorectal cancer (CRC) is the third most common form of cancer and the second leading cause of cancer-related deaths in the United States[1]. A majority of CRC cases are due to sporadic tumorigenesis, while only 10% are associated with a genetic predisposition[2,3]. The increased risk of developing CRC in people with metabolic syndrome, type 2 diabetes and inflammatory bowel disease[4] highlights a strong role for chronic inflammation in the etiology of inflammation-associated or colitis-associated CRC[5,6]. Other common features shared between these patient populations include changes within the composition of the intestinal microbiota and the development of immune reactivity against these intestinal microbes[7,8].

The intestine is home to a large microbial ecosystem that provides protective, structural and metabolic functions[9]. The importance of the microbiota in CRC has been illustrated in both animal models and in studies of patients with CRC[10–12]. In models of inducible and sporadic CRC, the composition of conventionally raised mice changes over the course of tumorigenesis[13]. Shifts in the composition of the microbiota have also been observed in human studies, which have not only found changes in the abundance of certain bacterial taxa but have identified a number of microbes thought to be drivers in the progression to malignancy. These include *Fusobacterium nucleatum*[14,15], *Bacteroides fragilis*[16,17] and *Escherichia coli*[18]. The importance of the microbiota in CRC has also been demonstrated using germ-free mice which develop less inflammation and fewer tumors than conventionally housed animals[19,20]. Despite these findings, the molecular mechanism contributing to the selection of certain bacterial families over another for an intestinal niche has not been established for CRC.

Due to the proximity of the microbiota to the intestinal epithelium and underlying immune cells, tightly regulated communication must occur to prevent abnormal tissue responses that could lead to chronic inflammation and malignancy. Coordination of intestinal responses are initiated through the recognition of both microbial-derived and host-derived ligands by innate immune receptors, such as Nod-like receptors (NLR) and the Toll-like receptors (TLR)[21]. TLRs comprise a set of receptors that recognize conserved microbial motifs (e.g. LPS, flagellin), as well as endogenous danger signals (e.g. heat shock proteins)[22]. These receptors represent a line of defense against invading enteric pathogens[23], but also sense and respond to our own commensal bacteria in order to promote epithelial cell integrity[24], localized immune responses and even colonization of the host[25]. Disruption of these signals can lead to uncontrolled inflammation and changes within the

microbiota, which have significant roles in intestinal disease[26], tumorigenesis and tumor progression[10,27]. Toll-like Receptror-2 (TLR2) recognizes di- and tri-acylated bacterial lipoproteins when heterodimerized with TLR6[28] or TLR1[29], respectively. In a model of inflammation-associated CRC, mice deficient for TLR2 have shown an increase in size and number of colonic tumors, dysregulation of cytokine production and epithelial responses[30]. Since both TLR1 and TLR6 require TLR2 to signal[21], deletion of TLR2 also results in the disruption of their signaling, making it difficult to delineate the role that these receptors may play in sensing and responding to microbial ligands in CRC. Using a model of inflammation-associated CRC we sought to understand how TLR6 signaling would affect a disease associated with alterations of the microbiota in a highly inflamed environment. We found that TLR6 deficiency was associated with the development of more tumors and worse survival when housed with other TLR6KO mice or when cohoused with WT mice. Interestingly, cohousing provided protection against tumors in the WT mice due to the presence if Lactobacillus which induced IL-10 in a TLR6-dependent manner Overall these studies suggest that innate sensing of the microbiota during inflammation can influence the local cytokine milieu, the composition of the commensals and the response to microbiome based immune therapies.

4.3. MATERIALS AND METHODS

Animals

TLR6KO and littermate control mice were bred in-house and maintained in a specific pathogen free facility at the University of Southern California. All mice used in these experiments were generated from 5-7 different founder cages where all female mice were wild type for TLR6. At the time of weaning pups from different litters but of the same sex were mixed together for group housing to reduce founder and cage effects on the microbiome. Mice were fed a standard rodent diet AIN-78 (Lab diets, St. Louis, MO) and given non-chlorinated, non-acidified, distilled water ad libitum. All animal experiments were performed in accordance with institutional guidelines following experimental protocol review and approval by the Institutional Biosafety Committee and the Institutional Animal Care and Use Committee. Induction of inflammation-associated colorectal cancer 6- to 8-week-old female mice were injected intraperitoneally with 10 mg/kg azoxymethane (AOM; Sigma-Aldrich Chemical Co, St. Louis, MO). After 5 days, mice received 2.5% (w/v) dextran sulfate sodium (DSS; Affymetrix, Santa Clara, CA, molecular weight 35,000–50,000 kDa) in drinking water for 5 days followed by 16 days of regular drinking water. Mice

were subjected to two additional DSS cycles[50]. The clinical course of disease was monitored by measurement of body weight, observation of rectal bleeding, diarrhea, and bloody stool during DSS treatment. Mice were euthanized 66 days post-AOM injection using CO2 followed by cervical dislocation.

Pathology and Scoring

The gastrointestinal tract was removed and the colon separated from the cecum. Macroscopic assessment of colonic tumors was noted and images taken to determine tumor size using ImageJ software. The colon was placed into 10% neutral buffered formalin, prepared in a "Swiss roll" technique[51], routine-processed for paraffin embedding and stained with hematoxylin and eosin. Colons were scored by a veterinary pathologist blinded to experimental groups. Inflammation scores were based on the severity of mucosal loss, crypt inflammation, lamina propria mononuclear cells, neutrophils, epithelial hyperplasia mucosal epithelial hyperplasia, dysplasia presence and extent of pathology as modified from Suzuki et al[52] and Kennedy et al[53]. Proliferative lesions were graded according to Boivin et al[54] with invasive carcinoma defined as invasion into the lamina propria or into but not through the muscularis mucosa, sparing carcinoma.

Collection of luminal and mucosal associated microbiota samples

Colons were removed and placed onto a chilled glass plate on ice and opened longitudinally. Luminal contents were collected and placed in a 1.5 ml conical tube. The colon tissue was then rinsed with ice cold PBS and then washed by submerging tissues into clean sterile PBS and then vigorously shaken for 5 seconds. This was repeated 5 times and then the colon was placed back onto glass plate and glass slides were used to gently scrape the colon tissue removing the mucosa. This was then added to the luminal contents and used for sequencing.

454 Pyrosequencing, sequence curation and microbiome analysis

Microbiota samples were processed and sequenced at Research and Testing Laboratory (RTL; Lubbock, TX) based upon RTL protocols using a Roche FLX Titanium genome sequencer. Universal bacterial primers 28F 'GAGTTTGATCNTGGCTCAG' and 519R 'GTNTTACNGCGGCKGCTG' were used to amplify the variable regions V1-V3 of the 16S rRNA genes. 16s rRNA gene sequences were curated using mothur v.1.35.1[55]. Briefly, sequences were denoised using a flow gram denoising algorithm[56], aligned to Silva 16s rRNA sequence database[57], and pre-clustered to allow up to a 2-bp difference between sequences[58]. Chimeras were

detected using UCHIME[59] and were culled along with chloroplast and mitochondrial sequences. Sequences were then classified using the Ribosomal Database Project version 14 with a confidence score greater than 80%[60] and phylotyped to the family level. Prior to any data analysis the number of sequences were normalized per sample to at least 2000 sequences. Beta diversity was calculated using the Theta YC distance metric with the family-level data and visualized using principal coordinates analysis (PCoA).

Bacterial quantitative-PCR

DNA was extracted from proximal colonic contents using an ISOLATE Fecal DNA kit (Bioline, Taunton, MA). qPCR was performed on a CFX96 TouchTM Real-Time PCR Detection System (Bio-Rad Laboratories, Irvine, CA) using a SensiFAST SYBR® No-ROX Kit (Bioline) with 20 ng of bacterial DNA and the following primers: Helicobacteraceae primers forward (5'-CCGCAAATTGCAGCAATACTT-3') and reverse (5'-TCGTCC AAAATGCACAGGTG-3'); Lactobacillaceae 16S primers LabF362(5'- AGCAGTAGGGAATCTTCCA-3') and LabR677 (5'-CACCGCTACACATGGAG-3'); Porphyromonadaceae primers forward (5′-GGTGTCGGCTTAAGTGCCAT-3′) and reverse (5′-CGGA(C/T)GTAAGGGCCGTGC-3′). 10-fold serial dilutions of plasmid-based *H. hepaticus*, *L. reuteri*/*L. johnsonii* or *Tannerella forsythia* genomic DNA was used to generate a standard curve. Relative abundance was calculated by a ratio of the organism- specific DNA to total bacterial DNA used for the amplification.

Lactobacillus treatments

L. johnsonii, L. reuteri, and LGG (*L. rhamnosus GG*) were grown overnight at 37°C in *Lactobacilli* MRS broth (EMD Chemicals, Gibbstown, NJ). Cultures were diluted the next day in sterile PBS, and concentrated to 10^{10} CFU/ml. On day 2 of each DSS round, mice were administered 100 µl by intragastric gavage every other day for 10 days (total of 5 gavages).

ELISA

Colonic lamina propria was isolated, pelleted and protein determination was performed by Bradford assay. ELISA was performed according to product instructions from BD OptEIA kits (IFN-γ and IL-10) and R&D Systems kits (IL-13 and IL-17).

Quantitative reverse transcription PCR (qRT-PCR)

Colons were flushed with PBS, opened longitudinally and segmented into two sections. One third of the entire length of the colon measured up from the anus was denoted as distal, while the remaining two thirds of the colon was identified as proximal. RNA was extracted from mucosal scrapings using ISOLATE II RNA Mini Kit (Bioline, Taunton, MA) and reverse transcribed into cDNA with SensiFAST cDNA Synthesis Kit (Bioline). qRT-PCR was performed on a CFX96 TouchTM Real-Time PCR Detection System (Bio-Rad Laboratories, Irvine, CA) using a SensiFAST SYBR® No-ROX Kit (Bioline) with the following primers: Tbet forward 5'-AGCAAGGACGGCGAAGTT-3'; Tbet reverse 5'-GGGTGGACATATAAGCGGTTC-3'; Gata3 forward 5'-CAACCTCTACCCCACTGTG-3'; GATA3 reverse 5'-GATGTCCCTGCTCTCCTTG-3'; Rorc forward 5'-GCCTACAATGCCAACAACCACACA-3'; Rorc reverse 5'-ATTGATGAGAACCAGGGCCGTGTA-3'; Foxp3 forward 5'-AGAGAGGTATTGAGGGTGGG-3'; Foxp3 reverse 5'-GCTGAGATGTGACTGTCTTCC-3'; 18S forward 5'-GTAACCCGTTGAACCCCATT-3'; and 18S reverse 5'-CCATCCAATCGGTAGTAGCG-3'. Gene expression levels for each individual sample were normalized to that of 18S. Fold changes in gene expression were relative to unstimulated controls and calculated using the ΔΔCt method.

Dendritic cell isolation and stimulation

To purify dendritic cells, MLNs were digested with 400 units per ml of collagenase type IV (Sigma-Aldrich). After filtering, the cells were re-suspended in 22.5% Optiprep (Sigma-Aldrich), overlaid with Hank's Balanced Salt Solution Saline (HBSS, Sigma-Aldrich) and centrifuged at 670g for 30 minutes. Dendritic cells were harvested from the interface of the HBSS and Optiprep, incubated with anti-CD11c (Miltenyi), and positively selected for CD11c using Automacs (Miltenyi). After isolation, the cells were stimulated overnight with bacterial lysates at 37°C before performing the ELISAs.

rIL-10 and αIL-10 treatments

In some experiments, mice were treated with probiotic bacteria and/or, corresponding to the days of probiotic therapy, were administered 200 μg of anti-IL-10 or anti-IgG1 (Bio X Cell, West Lebanon, NH) intraperitoneally. In other experiments mice were administered 1 ng of rIL-10 (R&D) diluted in 500 μl sterile PBS given intraperitoneally 3 times during every DSS cycle: on

day 2, day 4 and the day after discontinuation of the DSS.

Immune fluorescence

Murine colons were rolled using the Swiss roll technique and fixed in 10% neutral buffer formalin (VWR International, Visalia, CA) overnight. Paraffin-embedded (without formalin) tissues were cut 5 μm thick. Hematoxylin and eosin, was performed by AML labs (Saint Augustine, FL) or in-house. Rabbit polyclonal IgG, (Abcam, ab27472), rabbit monoclonal IgG (Abcam, ab172730), anti-bcl2 (clone 3F11). BrdU staining was performed as per manufacturer's instructions (Abcam). For antigen retrieval, slides were placed into tubes containing sodium citrate buffer (pH 6.0) (Sigma-Aldrich, St. Louis, MO) buffer and heated at 99°C in water bath. The slides were washed and blocked before staining with primary antibody. The slides with primary antibody were incubated overnight at 4°C in a humid chamber and the following day, washed and incubated with secondary at 37°C for 1 hour. The slides were mounted with DAPI and confocal images were acquired using a Nikon Eclipse C1 laser-scanning microscope (Nikon, PA) fitted with a 60 Nikon objective (PL APO, 1.4NA) and Nikon image software.

Western Blot

Colonic scrapings of the whole colon were placed in radio-immunoprecipitation assay (RIPA) buffer (VWR) containing protease and phosphatase inhibitors (Thermofisher). Samples were normalized for protein content using Bradford reagent (Bio-Rad, 500-0205). Protein was separated by size by SDS-PAGE using Mini-Protein tris-glycine gels (4-15%) (Bio-Rad, 456-1083) and transferred to PVDF membrane for blotting. Membranes were blocked with 3% dry nonfat milk in Tris-Buffered Saline (TBS) containing 0.05% tween-20 and incubated with primary antibody overnight at 4°C ((anti-cleaved caspase 3 (clone 3F11, Abcam); anti-pAkt thr308 (clone 244F9, Abcam); anti-pStat3 Y705 (clone S727, Abcam)). Membranes were washed with TBS containing 0.05% tween-20 and incubated with goat anti-rabbit-HRP secondary antibody (Santa Cruz Biotechnology) for 2 hours at room temperature. Membranes were developed using SuperSignal West Femto Maximum Sensitivity kit (Thermo Scientific) according to manufacturer's instructions and imaged using a LiCOR Odyssey.

Statistics

Data are expressed either as the mean value + standard error of the mean (s.e.m.) or as individual values. Specific statistical tests used for each experiment are described in the figure legends and

performed in Graphpad PRISM (Graphpad Software LLC, San Diego CA). $p < 0.05$ was considered significantly different.

4.4. RESULTS

TLR6 deficiency exacerbates inflammation-associated CRC, while co-housing protects WT mice

TLR6-deficient (TLR6KO) and littermates (WT) were weaned from their mothers and separately housed by genotype (SH) or co-housed (CH) together, two weeks later the mice were treated with the mutagen azoxymethane (AOM) followed by three rounds of dextran sodium sulfate (DSS) over 66 days (Figure 4.1A). Despite little difference in weight loss (Figure 4.1B) TLR6-deficient mice had an overall worse survival outcome compared to the SH-WT mice regardless of whether they were co-housed or housed with only other TLR6KO's (Figure 4.1C). Both SH- and CH-TLR6KO mice showed an increase in the size and number of tumors (Figure 4.1D-E) compared to both the SH- and CH-WT mice. To our surprise, WT mice co-housed with TLR6KO mice exhibited smaller and less tumors than their CH-TLR6KO cage mates and compared to SH-WT mice (Figure 4.1F-G). Analysis of H&E sections of the colons from these mice indicated CH-WT mice had less pathological changes in the colon, less inflammation and a reduction in overall pathology score than both their CH-TLR6KO cage mates and SH-WT mice (Figure 4.1F-G).

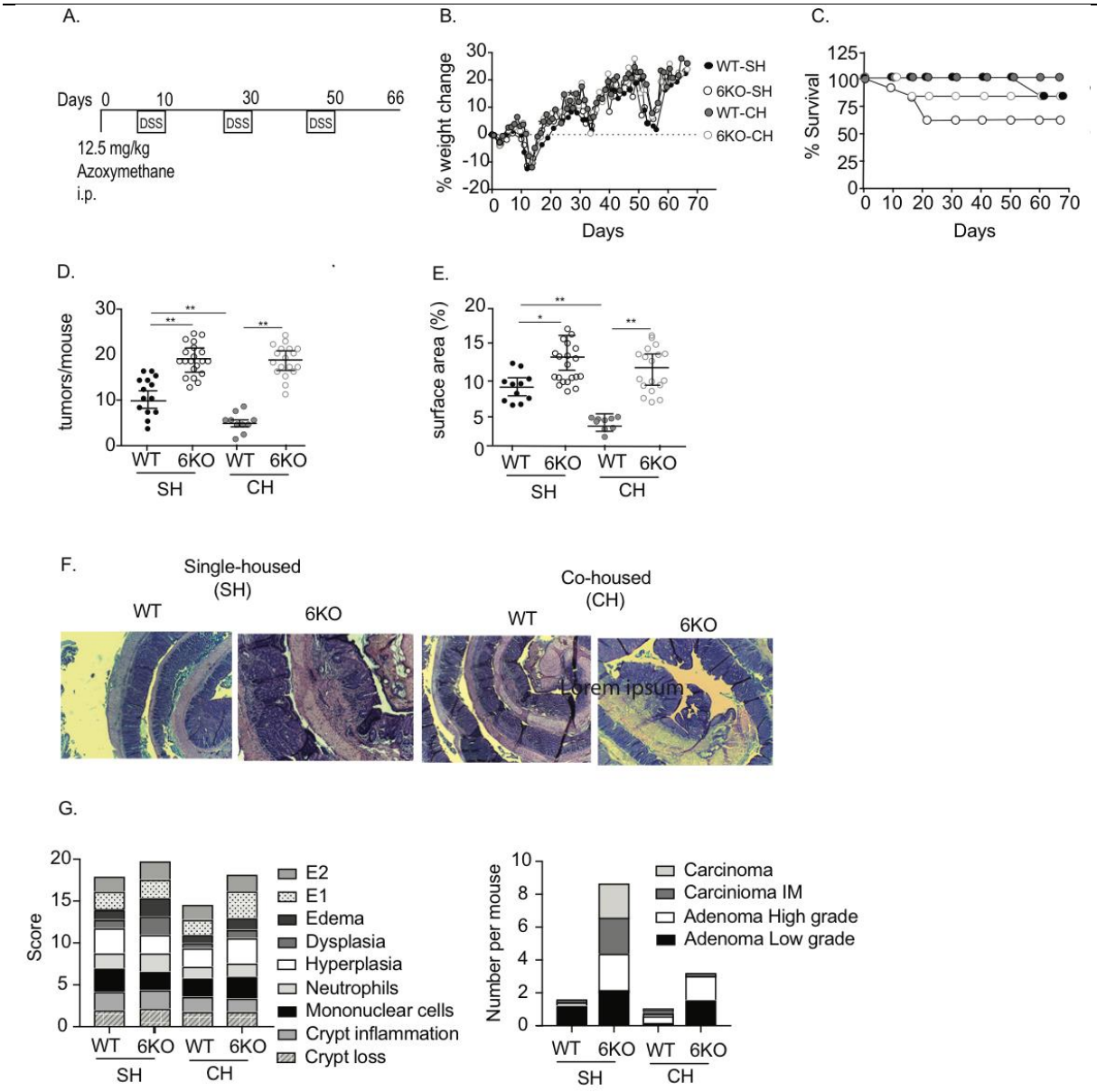


Figure 4.1: TLR6 signaling reduces the severity of inflammation-associated colorectal cancer

(A) Induction procedure for AOM/DSS model. (B) Percent change in weight of TLR6-deficient (TLR6KO) and littermates (WT) when co-housed (CH) or separately housed (SH) by genotype. (C) Percent survival over the course of AOM/DSS treatment, (D) Number of tumors and (E) the total surface area of tumors quantified using image j software. (F) Representative images of Hematoxylin and Eosin stained colons of mice collected on day 66. (G) Pathology scores of colonic tissues assessing inflammation and grade of lesion (see Methods for detailed criteria used to make these assessments). B-E, Data is pooled from 11-20 mice over 4 independent

experiments and expressed as the mean ± s.e.m (B-C) or as individual points (D- E). F-G, data is pooled from 4-6 mice per group. *, $p < 0.05$. **, $p < 0.01$ (B, C) Wilcoxon Log-rank test, (D, E) Two-way ANOVA with Bonferroni post-hoc tests, (G) Unpaired Mann-Whitney t-test comparing; WT CH vs WT SH; KO-SH vs WT-SH; KO-CH vs WT-CH; KO-SH vs KO-CH).

Consistent with their increased tumor number and mortality, the TLR6KO mice had significantly worse dysplasia (p=0.02) and epithelial hyperplasia (p=0.03) compared to SH-WT mice (Figure 4.1G). Despite no significant differences in inflammatory parameters, hyperplasia, dysplasia or adenomas, the SH-TLR6KO mice had more crypt loss, high-grade adenocarcinomas and carcinomas than SH-WT mice (Figure 4.1F-G). Compared to SH-WT mice, co-housing significantly reduced epithelial hyperplasia (p=0.050) and dysplasia (p = 0.02) in WT mice. In contrast the CH-TLR6KO mice had significantly more hyperplasia, dysplasia, a greater extent of disease and more adenomas than the CH-WT mice (Figure 4.1F-G).

Our group and others have demonstrated a role for TLR6 and its binding partner, TLR2, in modulating inflammation via the induction of IL-10. Analysis of the lamina propria of the colon of AOM/DSS treated mice showed that TLR6KO mice, regardless of whether they were co- or single housed, had significantly higher levels of IL-17, IFN-γ and corresponding transcription factors *Rorc* and *Tbet*, compared to WT mice housed similarly (Figure 4.2A-B). Interestingly, only the CH-WT mice, which develop less tumors, had significantly elevated levels of IL-10 and higher expression of the transcription factors Gata3 and Foxp3 compared to SH-WT mice, and their co-housed TLR6KO counterparts (Figure 4.2A-B). Taken together, these data indicate that loss of TLR6 signaling causes an increase in tumor number and worsens survival, while co-housing protects WT mice from inflammation and developing tumors.

The exacerbated inflammatory response following AOM/DSS observed in the TLR6KO mice could be a result of deficient immune-modulatory mechanisms or mediated by a hyper-inflammatory microbiota associated with the TLR6KO mice. To begin to address both of these questions we measured IL-10 and IL-12p40 production from naive WT and TLR6KO bone-marrow derived dendritic cells (BMDCs) after stimulation with colonic contents harvested from

cohoused and separately housed mice at day 66 of AOM/DSS treatment, or naïve age-matched controls. IL-10 production was significantly higher in WT BMDCs cultured with colonic contents from either naïve or AOM/DSS-treated cohoused mice compared to TLR6KO BMDCs stimulated with the same contents (Figure 4.2C). In contrast, separately housed WT and TLR6KO contents elicited much more IL-12p40 and very little IL-10 (Figure 4.2C). When the BMDCs were derived from TLR6KO mice, we observed a poor IL-10 response and elevated IL-12p40 levels from all samples except naïve WT and TLR6KO colonic contents (Figure 4.2C). Taken together these data indicate that the colonic contents of co-housed TLR6KO and WT mice seemed to induce an IL-10 dominant response and dampen the levels of IL-12p40, while the colonic contents of both separately housed WT and TLR6KO promoted a more inflammatory response with high IL-12p40 and low IL-10. These data also suggest that in the context of TLR6 deficiency, colonic contents normally associated with high IL-10 instead induce a robust IL-12p40 response.

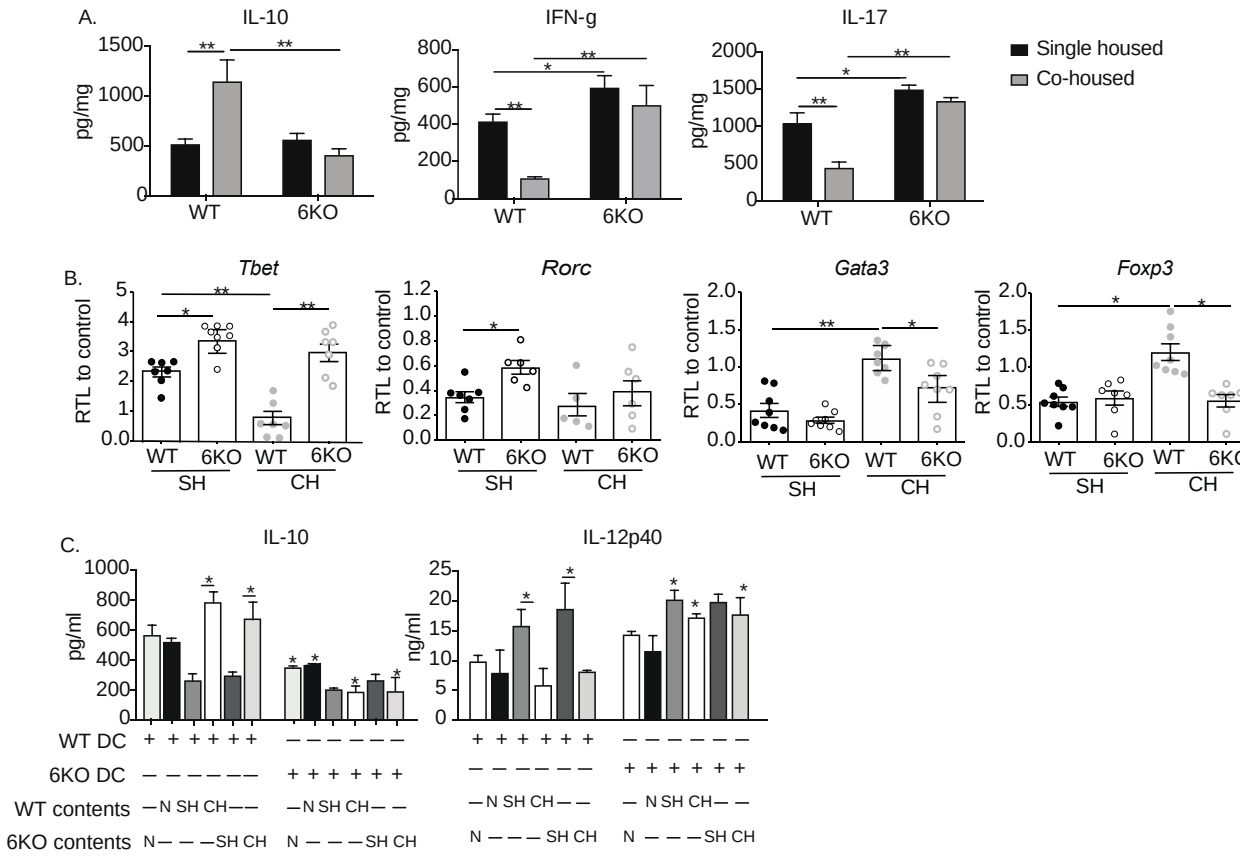


Figure 4.2: Colonic lysates from co-housed mice are less inflammatory and induce IL-10 in a TLR6-dependent manner.

(A) Levels of indicated cytokines from colonic lysates measured by ELISA. (B) Relative mRNA transcript levels of indicated transcription factors from colonic lysates were assessed by quantitative RT-PCR using naïve WT and TLR6KO mice as controls. (C) IL-10 (left) and IL-12p40 (right) protein production from BMDC generated from indicated naïve mouse (WT or TLR6KO) and stimulated with colonic lysates from indicated donors for 24 hours (N, naïve; SH, single-genotype housed; CH, cohoused; +, present; -, absent). A-C, data is expressed as mean + s.e.m of 6-8 mice from 2 independent experiments. *, $p < 0.05$. **, $p < 0.01$ Two-way ANOVA with Bonferroni post-hoc test.

Both genotype and housing influence dysbiosis associated with AOM/DSS

We hypothesized that both genotype and housing would contribute to alterations in the composition of the microbiota after AOM/DSS treatment. 16S rRNA sequencing was performed by scraping the colon with glass slides to capture both luminal and mucosal associated microbiota in naïve age-matched mice and day 66 AOM/DSS-treated mice. Analysis of the OTU's at the Phylum level revealed that co-housing significantly affected the composition of WT and TLR6KO mice, both naively and after AOM/DSS treatment (Figure 4.3A). Prior to AOM/DSS treatment, the colonic contents of SH-WT and SH-TLR6KO mice were dominated by taxa within Firmicutes, while co-housing WT and TLR6KO mice together lead to a dominance of Bacteroidetes (Figure 4.3A). Following AOM/DSS treatment there were major shifts in the Phylum level, regardless of genotype or housing (Figure 4.3A). Despite the absence of TLR6, co-housed mice shared a similar microbiome, dominated by the families Lactobacillaceae, Clostridiaceae and Erysipelotrichaceae (Figure 4.3B). In contrast, genotype seemed to impact the composition of the separately housed mice after AOM/DSS treatment in which we observed a significant increase in Helicobacteraceae and a small increase in Porphyromonadaceae in SH-WT mice (Figure 4.3B). This compositional phenotype was reversed in SH-TLR6KO mice who had much a higher abundance of Porphyromonadaceae and only a slight increase in Helicobacteraceae (Figure 4.3B).

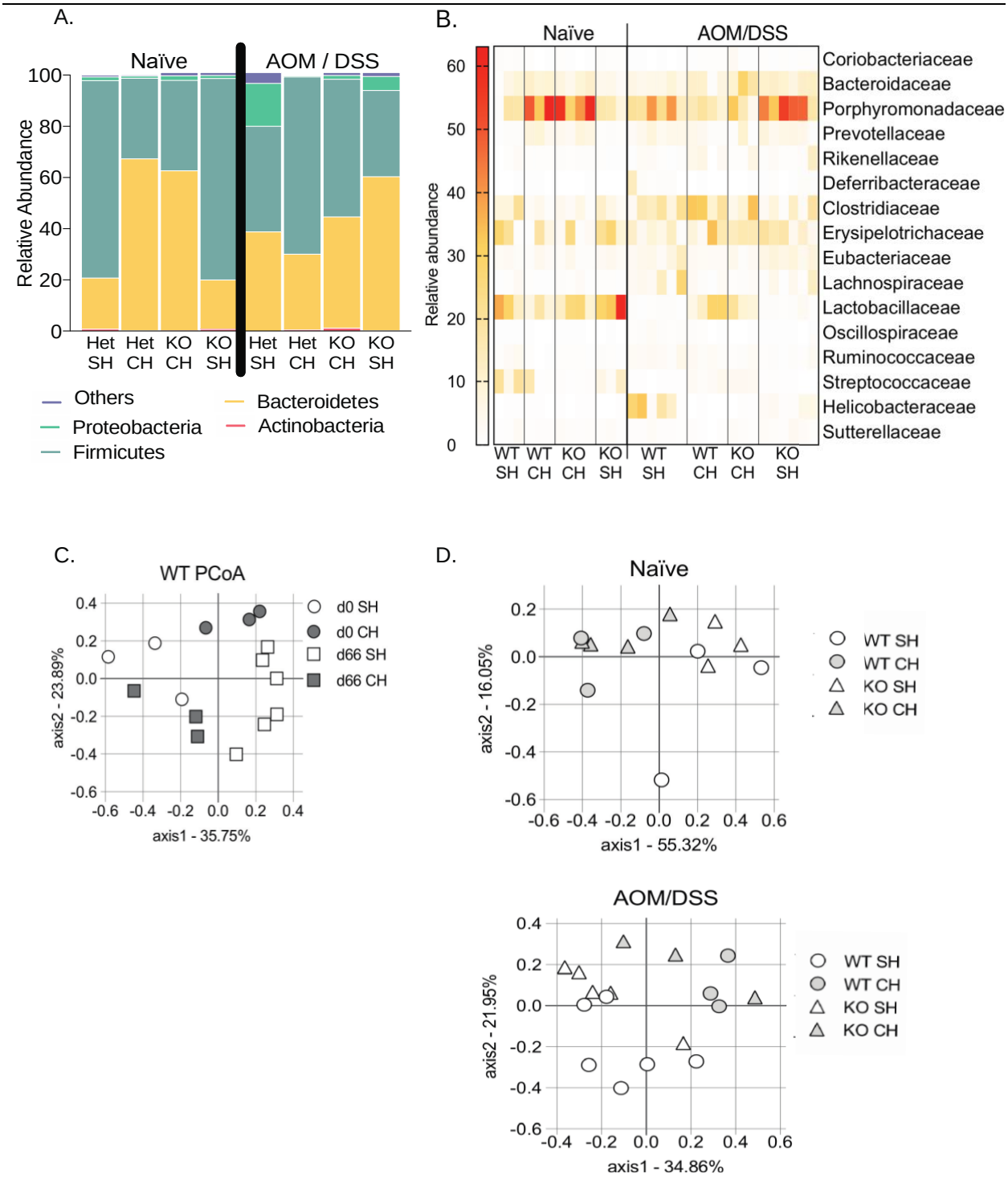


Figure 4.3: Compositional changes within the microbiota are influenced by TLR6 expression and housing status.

(A) Relative abundance of Phylum level microbiota from colons of naïve mice and mice treated with AOM/DSS under different housing conditions assessed by 16S rRNA sequencing. (B) Heatmap of family level abundance of colonic microbiota from colons of naïve and AOM/DSS treated mice under different housing conditions assessed by 16s rRNA sequencing. Bacterial families with less than 1% were not depicted. (C) Principal coordinates analysis (PCoA) of colonic microbiota samples of naïve (d0) and AOM/DSS treated (d66) WT mice. (D) Principal coordinates analysis (PCoA) of colonic microbiota samples of naïve and AOM/DSS treated WT and TLR6KO mice under different housing conditions. A-D. Data is pooled from two independent experiments with a total of 3-6 mice per group. A-D data is presented as mean + s.e.m. B-D individual points for each mouse are shown. C-D, $p < 0.001$, AMOVA.

Principal coordinate of analysis (PCoA) of naïve and SH-WT mice revealed that the colonic contents cluster distinctly depending on whether they were taken pre- or post-AOM/DSS treatment, whereas CH-WT mice have a microbiota more similar to naïve mice regardless of AOM/DSS treatment (Figure 4.3C). PCoA revealed a higher degree of likeness pre-AOM/DSS treatment in co-housed WT and TLR6KO mice than to their respective single house counterparts, suggesting that housing exerted a strong influence over composition in a steady state (Figure 4.3D). This influence was maintained under inflammatory conditions as the co-housed WT and TLR6KO mice clustered separately from both SH-WT and SH-TLR6KO following AOM/DSS treatment (p=0.006, AOM/DSS SH-WT vs CH-WT) (Figure 4.3D). However, the compositional changes that occurred in the SH-WT and SH-TLR6KO during AOM/DSS treatment were highly influenced by genotype as they showed distinct clusters in the PCoA (Figure 4.3D). Genotype and housing influence the kinetics and composition of the gut microbiota during AOM/DSS. The genus-level bacteria that were significantly affected by the interaction of housing and genotype by 3-way ANOVA, included *Porphyromonas* (p<0.0001), *Tannerella* (p<0.0001), *Prevotella* (p=0.0003), *Lactobacillus* (p<0.0001) and *Clostridium* (p=0.0122) (Supplementary Figure 4.1A). In order to more fully understand the changes in the composition, we collected colonic samples over the course of AOM/DSS treatment and performed PCR using primers that bind within conserved areas of 16S rDNA to identify bacterial families[31]. We found that the housing status had a significant impact on levels of bacterial DNA in the colon (Supplementary Figure 4.1B). In naïve co-housed mice roughly 50% of the total 16S rDNA was identified as Porphorymonadaceae, compared to only 15% in naïve single-housed mice of both genotypes (Supplementary Figure

4.1B). During AOM/DSS treatment, the levels of Porphorymonadaceae decreased in the co-housed mice regardless of genotype, increased in the SH-TLR6KO and stayed relatively the same in SH-WT mice (Supplementary Figure 4.1B). In all untreated mice 30-40% of the total 16S DNA belonged to Lactobacillaceae. Upon AOM/DSS treatment we found that the SH-TLR6KO mice were able to maintain levels of Lactobacillaceae DNA until the second round of DSS, while these levels dropped significantly after the first round of DSS in SH-WT mice (Supplementary Figure 4.1B). In the co-housed mice we found that housing status had a positive effect on the amount of Lactobacillaceae DNA, as neither CH-WT nor CH-TLR6KO mice exhibited a reduction after the first or second rounds of DSS, and both genotypes maintained relatively consistent levels of Lactobacillaceae DNA throughout the experiment (Supplementary Figure 4.1B). After the first round of DSS, when the levels of Lactobacillaceae DNA dropped in the SH-WT mice, we found that Helicobacteraceae DNA increased from 1% to 15% (Supplementary Figure 4.1B). Unlike the SH-WT, we saw no increase in Helicobacteraceae in the co-housed mice or the SH-TLR6KO mice (Supplementary Figure 4.1B). Taken together, these data demonstrate the absence of TLR6 signaling affects the compositional changes during AOM/DSS and that co-housing allows a maintenance of probiotic Lactobacillaceae, suggesting that both genotype and housing influence the kinetics and composition of the gut microbiota during AOM/DSS.

Restoring commensals ameliorates disease and reduces inflammatory cytokines in WT mice via a TLR6-dependent mechanism.

Many Lactobacillus exist within our GI tract and confer health benefits, thus the overall reduction in the abundance of Lactobacillus observed in SH-WT and SH-TLR6KO during inflammation-associated CRC led us to hypothesize that a re-colonization strategy using Lactobacillus could be efficacious in treating disease. To accomplish this, mice were administered a mixture of two taxa that were the most reduced as determined by OTUs in our tumor model (*L. johnsonii* and *L. reuteri*) or a species that was not found either naïvely or post-AOM/DSS (*L. rhamnosus GG*). These OTUs were confirmed via 16S sequencing and through the bacterial isolation and culture of colonic contents from mice pre- or post-AOM/DSS (data not shown). To ensure that any effect on tumor development was not due to the metabolism of AOM by the *Lactobacillus*, the re-colonization was performed 5 days post-AOM injection, beginning simultaneously with DSS administration, and given every other day for a total of 4 feedings. This regimen was repeated for each round of DSS (Figure 4.4A).

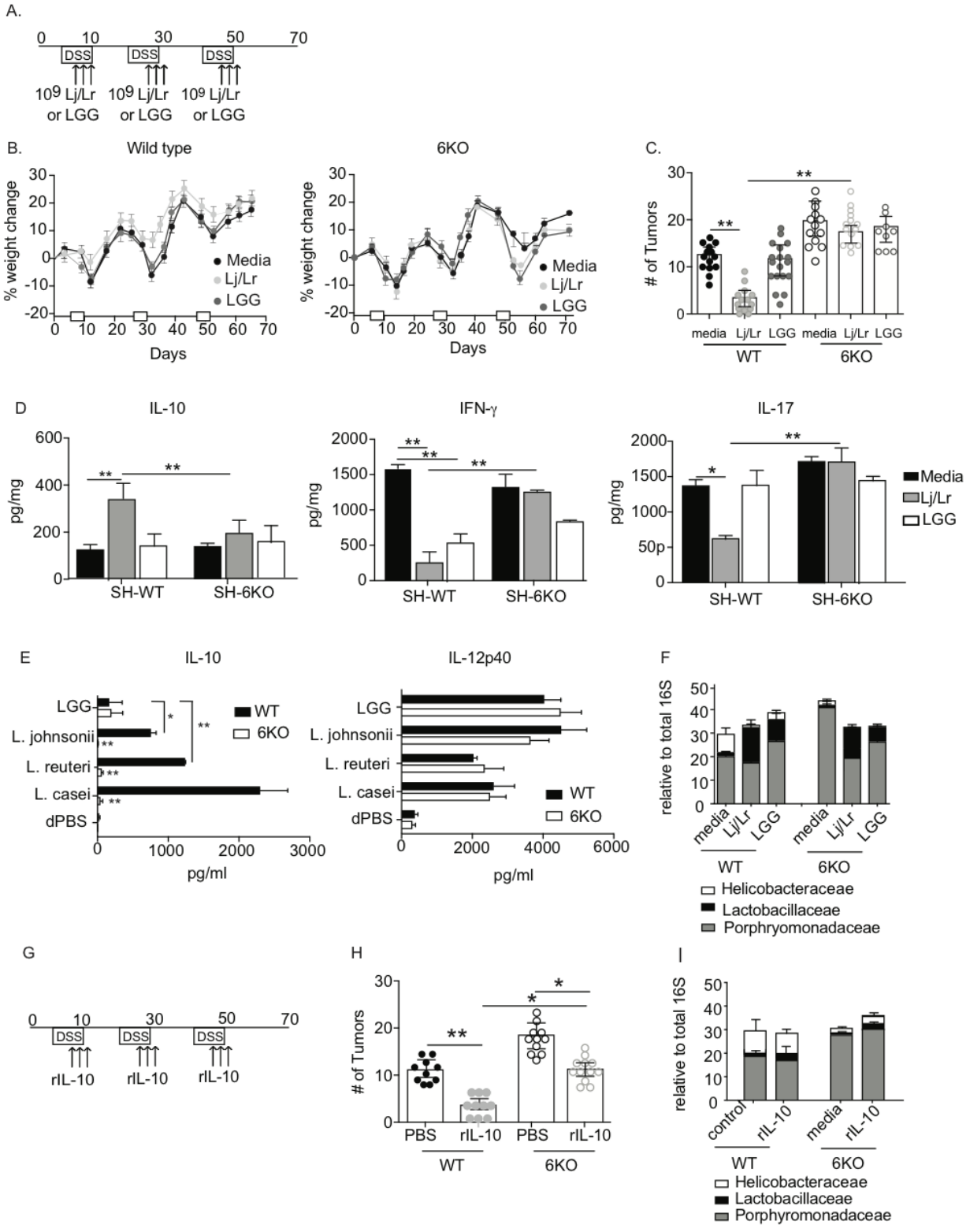

A.
0 10 30 50 70
DSS DSS DSS
10^9 Lj/Lr or LGG
10^9 Lj/Lr or LGG
10^9 Lj/Lr or LGG
B.
Wild type
6KO
% weight change
Days
Media
Lj/Lr
LGG
C.
# of Tumors
media Lj/Lr LGG media Lj/Lr LGG
WT
6KO
D
IL-10
IFN-γ
IL-17
pg/mg
SH-WT
SH-6KO
Media
Lj/Lr
LGG
E
IL-10
IL-12p40
LGG
L. johnsonii
L. reuteri
L. casei
dPBS
WT
6KO
pg/ml
F
relative to total 16S
media Lj/Lr LGG media Lj/Lr LGG
WT
6KO
Helicobacteraceae
Lactobacillaceae
Porphryomonadaceae
G
0 10 30 50 70
DSS DSS DSS
rIL-10 rIL-10 rIL-10
H
# of Tumors
PBS rIL-10 PBS rIL-10
WT
6KO
I
relative to total 16S
control rIL-10 media rIL-10
WT
6KO
Helicobacteraceae
Lactobacillaceae
Porphyromonadaceae

Figure 4.4: Restoration of Lactobacillus reduces tumor burden and suppresses inflammation in a TLR-6 dependent and independent manner.

(A) Methods for Figure 5 showing mice received an oral inoculation of sham (media) (n=9), L. johnsonii and L. reuteri (Lj/Lr) (n=11) or LGG (n=6) at the time-points indicated in the text and Materials and Methods. (B) % weight change and (C) number of macroscopic tumors from mice. (D) Levels of indicated cytokines from colonic lysates as measured by ELISA. (E) Concentration of IL-10 and IL-12p40 produced from WT or TLR6KO BMDC stimulated for 24 hours with 106 cfu indicated Lactobacillus species. (F) Relative DNA levels of indicated bacterial families to total 16S rDNA from colon taken at day 66. (G) Same method as in (A, above) except WT mice were given an i.p. injection of rIL-10 instead of Lactobacillus and (H) the number of tumors and (I) the relative DNA levels of indicated bacterial families to total 16S rDNA from colon at day 66. A-I, Data is the mean + s.e.m of 7-12 mice pooled from 3 independent experiments. *, $p< 0.05$; **, $p<0.01$, C-F, Student's t-test; H-I Two-way ANOVA with Bonferroni post-hoc tests.

Physical observations after administration of *L. johnsonni/L. reuteri* (Lj/Lr) revealed no effect on weight (Figure 4.4B), however the SH-WT mice treated with Lj/Lr had significantly fewer tumors than the media control treated SH-WT mice (Figure 4.4C). Further, the reduction in tumor number seemed to be dependent upon TLR6 expression as Lj/Lr treatment failed to reduce tumor number or size in the SH-TLR6KO mice (Figure 4.4C). In contrast, treating SH-WT mice with *L. rhamnosus GG* (LGG) failed to suppress both tumor number and size (Figure 4.4C).

Analysis of the mucosal cytokines associated with Lactobacillus treatment revealed distinct cytokine patterns. SH-WT mice treated with Lj/Lr had a significant increase in anti-inflammatory IL-10 and subsequent suppression of inflammatory cytokines, IFN-γ and IL-17, compared to control media treated WT mice (Figure 4.4D). The observed cytokine effects were also mediated by TLR6 as Lj/Lr treated SH-TLR6KO mice were unable to induce IL-10 or suppress IFN-γ or IL-17 (Figure 4.4D). In contrast, treatment with LGG had no effect on cytokine production and seemed to be independent of TLR6 signaling as WT and TLR6KO mice had similar cytokine patterns (Figure 4.4D).

In addition to their ability to secrete anti-microbial peptides[32] and shift the pH of the colonic environment to kill pathogenic bacteria[33], *Lactobacillus* are also well known for their anti-inflammatory properties[34]. Our group and others have shown that TLR6 signaling in antigen presenting cells promotes IL-10 production[35] and polarize anti-inflammatory T cells[36]. As we observed IL-10 production in SH-WT mice treated with Lj/Lr, but not by LGG, we wanted to assess the ability of a panel of lactic acid-producing bacteria to induce IL-10 and IL-12p40 and determine their dependency on TLR6 by using bone-marrow derived dendritic cells (BMDC) from naive WT and TLR6KO mice. *L. johnsonii*, *L. reuteri* and *L. casei* induced IL-10 and IL-12p40 from WT BMDCs, but only IL-10 was dependent upon the expression of TLR6 (Figure 4.4E). LGG stimulated BMDC were able to produce IL-10 regardless of TLR6 expression and had a robust TLR6-independent IL-12p40 response (Figure 4.4E). These data suggest that LGG induces low amounts of IL-10 through a TLR6-independent mechanism and could explain its inability to protect WT mice.

We next wanted to determine if the Lj/Lr treatment had any effect on the microbiota by screening the colonic contents of mice at the end of the AOM/DSS treatment by qPCR using the same primers used in Figure 4. Confirming our 16S and qPCR data, we detected high amounts of Helicobacteraceae DNA in colons from SH-WT mice, while SH-TLR6KO mice had higher amounts of Porphyromonadaceae DNA and only a small amount of DNA from Helicobacteraceae (Figure 4.4F). Treatment with Lj/Lr was able to reduce the levels of Helicobacteraceae and Porphyromonadaceae in both the SH-WT and SH-TLR6KO mice, respectively (Figure 4.4F). Despite the lack of any anti-tumor activity in our model, we found that LGG treatment was also able to reduce the amount of Helicobacteraceae and Porphyromonadaceae DNA recovered from the colons of both SH-WT and SH-TLR6KO mice. Interestingly, treatment by either Lj/Lr or LGG resulted in an increase in Lactobacillaceae DNA in both WT and TLR6KO mice (Figure 4.4F).

IL-10 is responsible for the partial reduction in tumor number but is dispensable for changes in the microbiota

The molecular pathobiology of CRC has implicated inflammation in the promotion of tumor progression, invasion, and metastasis[37]. A clear example of this is the finding that patients with inflammatory bowel disease are at higher risk of CRC[38]. One mechanism that could account for

the reduction in tumor number observed in SH-WT mice given Lj/Lr may be the ability to limit inflammation via the production of IL-10. To evaluate whether IL-10 alone was sufficient to suppress tumor number and limit colonization by Helicobacteraceae, we substituted Lj/Lr treatments with direct intraperitoneal (i.p.) administration of recombinant IL-10 (rIL-10) (Figure 4.4G). This protocol allowed us to determine if IL-10 could mediate any of the phenotype observed in Lj/Lr treated WT mice. Further, it would allow us to determine the requirement for TLR6, as administration of rIL-10 to TLR6KO mice would bypass the need for TLR6 signaling, and these mice should behave more similarly to the WT mice treated with AOM/DSS. Similar to treatment with Lj/Lr, rIL-10 provided a partial reduction in tumor number in SH-WT mice (Figure 4.4H), but it failed to reduce the amount of Helicobacteraceae DNA (Figure 4.4H-I). As expected, the TLR6KO mice treated with rIL-10 treatment also had a significant reduction in the number and size of tumors compared to control treated TLR6KO mice given AOM/DSS (Figure 4.4H). However, the number of tumors was still significantly higher in the TLR6KO mice treated with rIL-10 than WT mice given rIL-10 (Figure 4.4H) suggesting that there may be other molecular pathways contributing to the development of tumors in the TLR6KO mice. Similar to WT mice, we found no effect on the level of Helicobacteraceae or Porphyromonadaceae DNA in the colons of the rIL-10-treated TLR6KO mice.

The requirement for *Lactobacillus*-induced IL-10 was also further determined by co-administering a neutralizing antibody against IL-10 beginning one day prior to treatment with Lj/Lr and continuing for two days post-treatment (Figure 4.5A). Co-administration of an IL-10-neutralizing antibody with Lj/Lr had little effect on the weight of SH-WT or SH- TLR6KO mice (Figure 4.5B). Neutralizing IL-10 in the absence of Lj/Lr had no effect on the number of tumors observed in WT or TLR6KO mice (Figure 4.5C). Co-administration of anti-IL-10 with Lj/Lr only partially reversed the reduction in tumor number and had no effect on the levels Helicobacteraceae DNA in WT mice (Figure 4.5C-D). As expected, neutralizing IL-10 during Lj/Lr treatment in TLR6KO mice had little effect, likely due to the inability of TLR6KO mice to produce IL-10 during treatment (Figure 4.5B-D). Analysis of the mucosal cytokines demonstrated that neutralizing IL-10 was able to reverse the reduction in the levels of IFN-γ and IL-17 observed in Lj/Lr treated SH-WT mice (Figure 4.5E). Taken together, these results demonstrate that IL-10 production by Lj/Lr can partially explain the reduction in tumor size and number. While IL-10 did play a partial role in the reduction of tumors, it had no impact on the amount of Helicobacteraceae, Porphyromonadaceae

or Lactobacillaceae DNA. These data thereby dissociate the anti-inflammatory effects of IL-10 from the regulation of the microbiota in inflammation-associated CRC.

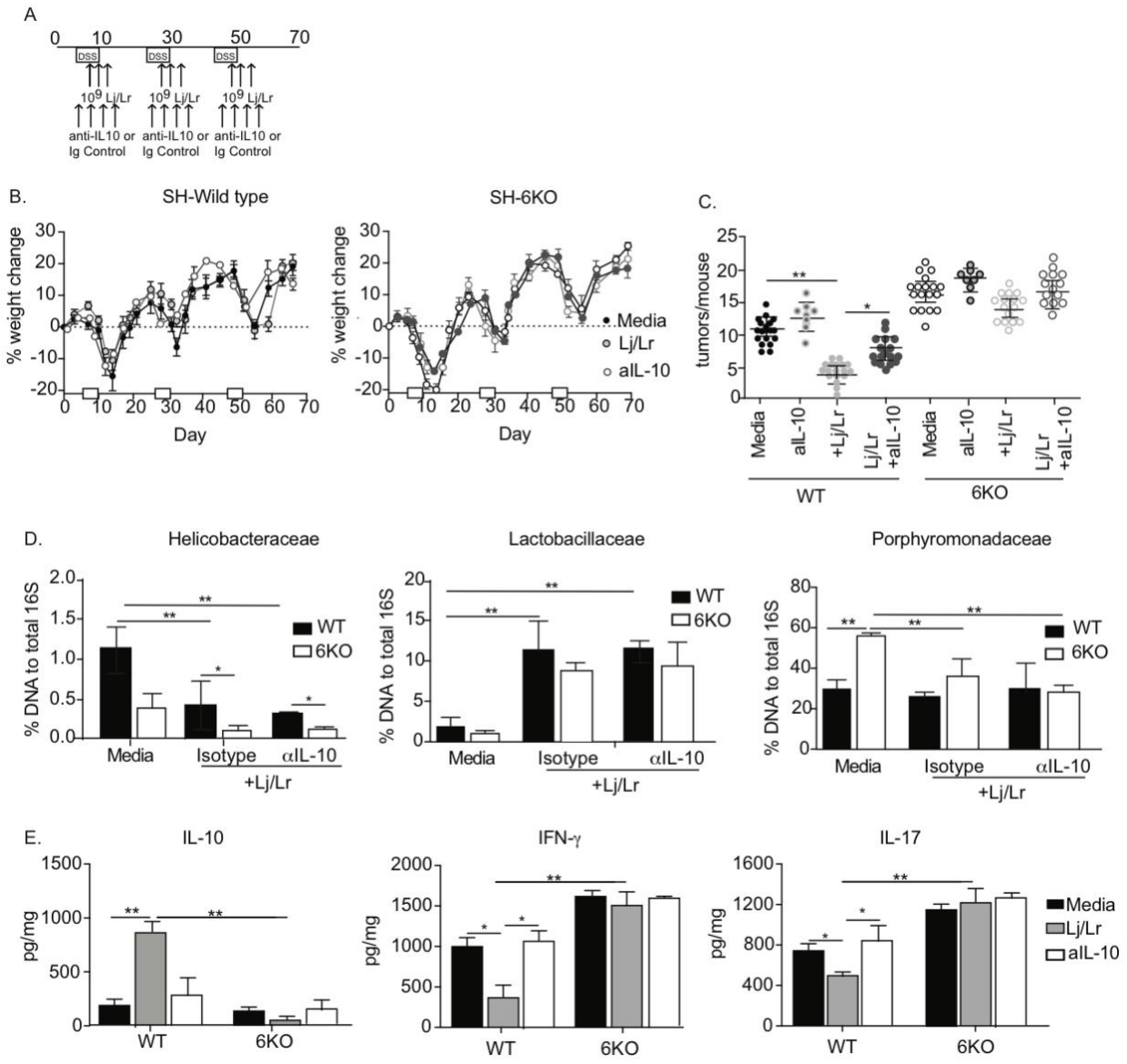


Figure 4.3: Requirement for IL-10 dissociates Lactobacillus effects on inflammation and composition.

(A) Methods. Mice were treated as described in Figure 5A and Materials and Methods but instead a monoclonal antibody against IL-10 was administered i.p. at indicated time points

during Lj/Lr treatments. (B) % weight change and (C) number of tumors from mice receiving Lactobacillus plus anti-IL-10 or control immunoglobulin (isotype). (D) Relative DNA levels of indicated bacterial families to total 16S rDNA from colon taken at day 66. (E) Levels of cytokines detected from colonic lysates of mice at day 66 by ELISA. B-E, Data is represented as the mean ± s.e.m of 8-12 mice pooled from 3 independent experiments. C-E, *, $p < 0.05$. **, $p < 0.01$. Two-way ANOVA with Bonferroni post-hoc test.

Impaired apoptosis during AOM/DSS in TLR6-deficient mice

While the data above provide evidence as to how co-housing would lead to reduced tumors in the WT mice, it does not address why both single and co-housed TLR6KO mice have more tumors than the SH-WT mice. We hypothesized that in the absence of TLR6 an important cellular regulatory pathway may be altered leading to an environment more sensitive to mutagenesis. We also hypothesized that this effect would be independent of a specific microbiota signature, as the CH-TLR6KO and SH-TLR6KO have distinct colonic microbial compositions.

One of the early steps in CRC progression is the increase in epithelial cell proliferation and/or an inhibition of apoptosis. To assess the proliferation of enterocytes, we performed immunohistochemical staining on colon sections of AOM/DSS treated WT and TLR6KO mice for Ki67 and bromodeoxyuridine (5-bromo-2'-deoxyuridine, BrdU). Ki67 is a commonly used marker that identifies cells that have recently undergone proliferation by labelling S, G1, and G2 phases of the cell cycle, while BrDU incorporates into newly synthesized DNA in actively replicating cells during S phase. There was an increase in the number of Ki67+ cells found in both the co-housed and single genotype housed TLR6KO mice compared to WT mice (Figure 4.6A). Two hours following injection of BrdU into AOM/DSS treated mice the colons were removed and stained for actively proliferating cells. In contrast to the Ki67 staining, we identified on average 3-4 BrdU+ cells per crypt, regardless of housing or genotype (Figure 4.6A). These data indicate that the disruption of TLR6 signaling or the housing status was not impacting the turnover of intestinal stem cells. In order to assess whether the TLR6KO mice may have defects in apoptosis[39], we performed immunohistochemistry on AOM/DSS colons looking for the anti-apoptotic factor Bcl-2. Analysis of the stained tissue revealed many Bcl-2+ cells within the epithelium, as well as in the lamina propria of SH-TLR6KO and CH-TLR6KO mice compared to WT mice (Figure 4.6B).

TLR2 signaling has been shown to signal through the PI3K/Akt pathway[40]. Pi3K/Akt has been shown to phosphorylate Stat3 which has been shown to influence apoptosis and proliferation signals in enterocytes[39,40]. We performed Western blot analysis on the colonic enterocytes at day 28 during AOM/DSS to assess if signaling pathways upstream and downstream of Bcl-2 were disturbed in the TLR6KO mice, thereby driving less apoptosis and making these mice more sensitive to AOM/DSS. Both SH- and CH-TLR6KO mice had lower levels of cleaved caspase-3 compared to WT mice (Figure 4.6C). Despite the reduction in cleaved caspase-3 and the increase in Bcl-2, the TLR6KO mice had and similar levels of phosphorylated Akt and phosphorylated Stat3 (Figure 4.6C). Taken together these data suggest that the deficiency in TLR6 may lead to an anti-apoptotic response that is Stat3- and Akt-independent but involves Bcl-2 and cleaved caspase-3.

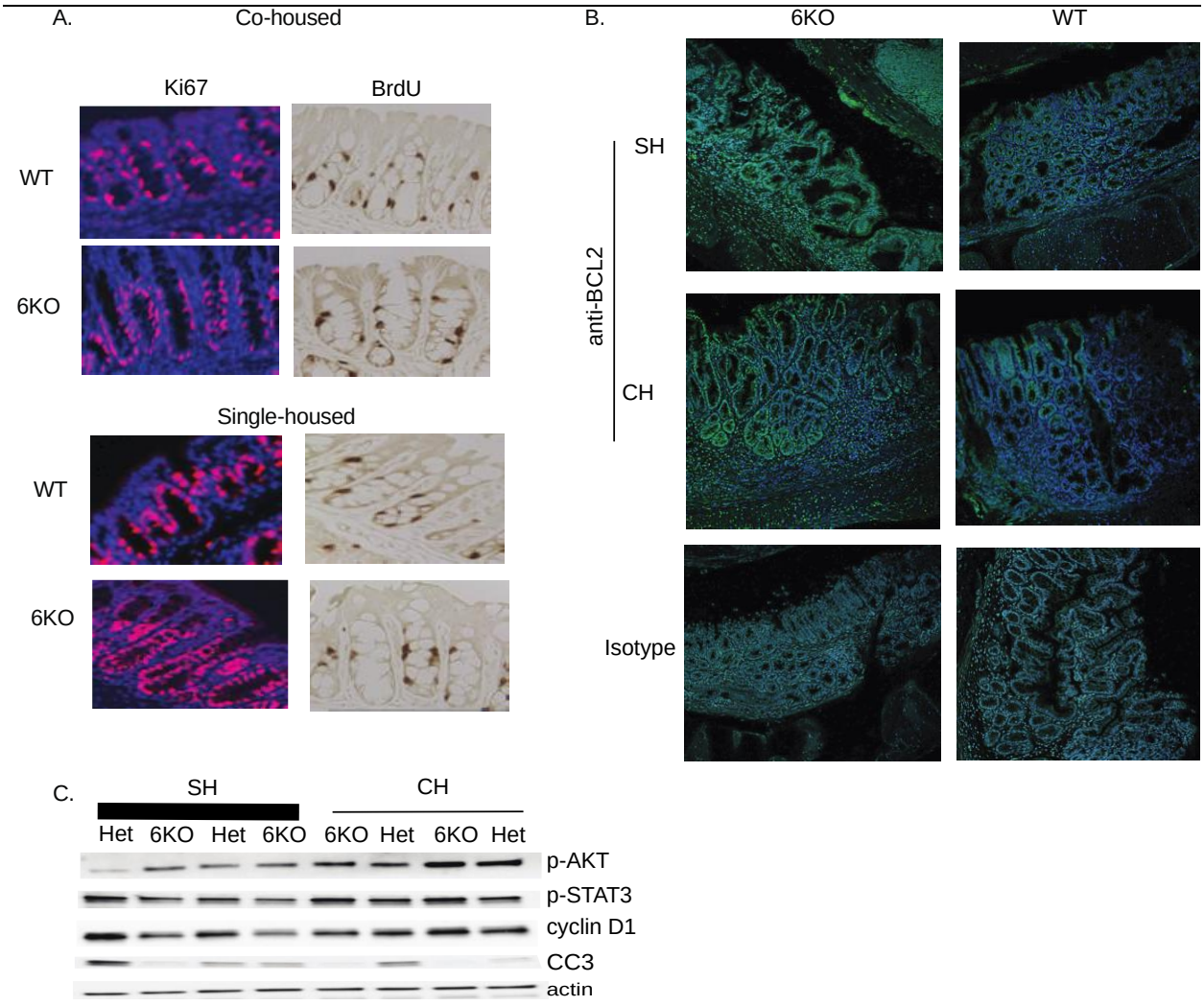


Figure 4.4: Proliferation and apoptosis of colonocytes are altered in TLR6KO mice.

(A) Colons from mice at day 28 post-AOM/DSS stained with Ki67 or from mice injected with BrdU 2 hours before the end of the experiment. Representative image from one of 3-5 mice per group (B) Histological sections of the colon from mice at day 28 stained with anti-Bcl-2 antibody or isotype control. Representative image from one of 3-6 mice per group (C) Western blot of whole colon from individual mice at day 28 post-AOM/DSS blotted for phosphorylated Akt, phosphorylated Stat3, and Cleaved caspase-3.

4.5. DISCUSSION

In the present study, we established a potential treatment for CRC using a microbiome-associated therapy that restores commensal species lost during inflammatory disease. Compositional data generated by 16S rRNA sequencing of colons from AOM/DSS treated WT mice revealed a distinct microbiota from their naïve counterparts. Most noticeably, was an increase in the abundance of cancer promoting Helicobacteraceae, and a loss of beneficial Lactobacillaceae. Restoring *Lactobacillus* dampened inflammation and reduced tumor number. These data underscore the need for a more personalized approach in using probiotics to treat or ameliorate disease. The rationale "not all probiotics are created equal" has been used to explain the failure of probiotics in clinical trials, focusing on differences between bacterial species, strain or dose. However, data from this study and others[41,42] suggest that the personal relationship established between the innate immune system, intestinal cells and the resident commensals determine whether a given probiotic will be effective. Our study also demonstrates that the effects on inflammation-associated CRC by *L. johnsonii* and *L. reuteri* (Lj/Lr) are mediated by multiple mechanisms that can be dissociated from each other by the involvement of IL-10. While Lactobacillus has been shown to treat disease via IL-10-dependent[43,44] and IL-10-independent[45–48] mechanisms, our study demonstrates that the most efficient anti-tumor effects were associated with inhibiting inflammation and restoring microbial ecology by reducing the abundance of Proteobacteria.

Genetics play an important part in disease susceptibility and response to treatment, yet their impact on the microbiota is still not understood. Here, genetic ablation of Tlr6 in mice was associated with an increased incidence of tumors, and an insensitivity to Lj/Lr therapy. Our data suggests that expression of TLR6 is critical for the production of IL-10 during treatment with Lj/Lr, confirming our previous studies linking TLR6 signaling with IL-10[49] and further implicating the role of TLR6 in the sensing of endogenous microbial signals. TLR6 signaling did have an effect on the kinetics associated with the loss of *Lactobacillus*, with a more gradual decrease observed in the TLR6KO mice. The prolonged presence of the *Lactobacillus* had no beneficial effect on tumor number in TLR6KO mice, likely due to their inability to sense immune-modulatory ligands through TLR6 and induce IL-10. However, the prolonged presence of *Lactobacillus* in AOM/DSS-treated TLR6KO mice had a protective effect on WT mice when co-housed together. It may be that the WT mice were able to maintain Lactobacillus levels due to ingestion/exposure of *Lactobacillus* in

TLR6KO stool through coprophagy and this helped prevent tumors in these mice. Indeed, our data shows that restoring *Lactobacillus* species during DSS exposures was able to reduce the tumor burden in SH-WT mice in part through production of IL-10.

Compositional analysis of the colonic microbiota by 16S sequencing and qPCR revealed that SH-WT mice had an increase in Helicobacteraceae, while the TLR6KO mice had an increase in Porphyromonadaceae. Each of these increases occurred when Lactobacillaceae was lowest after the first or second round of DSS in the WT and TLR6KO mice, respectively. It is plausible that Helicobacteraceae has only a small window of opportunity when a comfortable niche is created by the absence of Lactobacillus. If this window is missed than other commensals, such as Porphyromonadaceae, may fill that niche. Our data would suggest that innate immune sensing of the endogenous microbiome impacts dysbiosis through altering the kinetics in which bacterial species are reduced and by contributing to the inflammatory milieu of the local tissue microenvironment. These data imply that the genetics of the host are linked to dysbiosis through expression of such receptors and could possibly explain why many of the polymorphisms associated with IBD involve innate immune sensing.

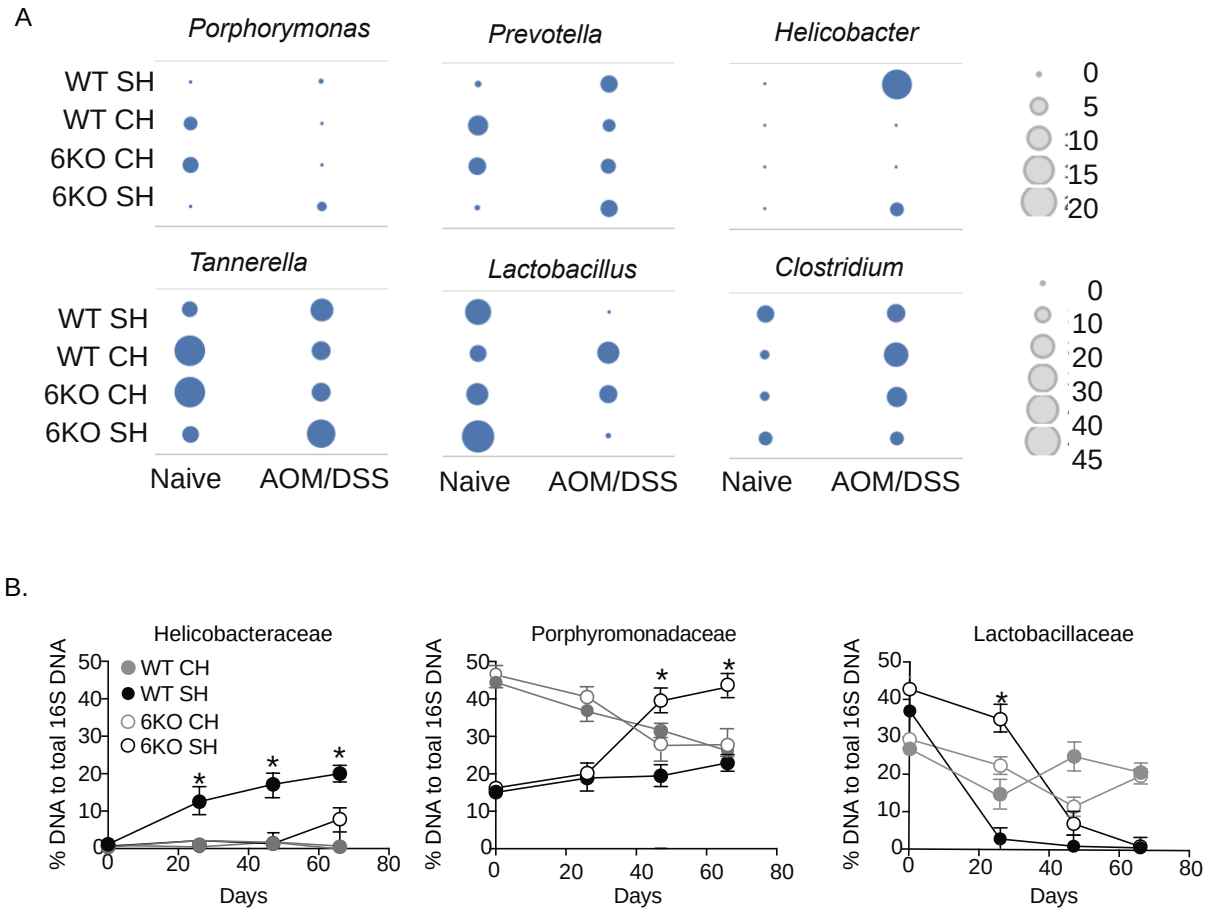
The contribution of innate signaling to the cytokine milieu of the local tissue microenvironment must also not be overlooked. Here, IL-10 production by the *Lactobacillus* was almost completely dependent on TLR6 as BMDCs deficient in TLR6 failed to produce IL-10 after stimulation by these bacteria, but also by the finding that the beneficial effects of *Lactobacillus* are lost in TLR6KO mice. Importantly, we observed that after treatment with *Lactobacillus*, decreases in Helicobacteraceae and Porphyromonadaceae were both IL-10 and TLR6-independent. While this study did not identify the mechanism by which Lj/Lr modulates the microbiota, treatment with rIL-10 or neutralization of IL-10 during Lj/Lr administration had little effect on the levels of Helicobacteraceae and Porphyromonadaceae, thus ruling out a role of IL-10 in dysbiosis and indicating an IL-10-independent role for Lactobacillus in maintaining microbial ecology.

Regardless of housing, TLR6KO mice had higher tumor numbers and higher mortality compared to the WT mice. This increased tumor number is likely independent of the microbiota as single and cohoused TLR6KO mice had very different microbiomes at both the induction of the AOM/DSS and after the 66 days. We found that TLR6KO mice have higher levels of Ki67+ cells

in the intestinal epithelium during AOM/DSS. Further molecular analysis revealed an increase in the anti-apoptotic molecule Bcl-2 in the TLR6KO mice regardless of housing status. In line with the increase in Bcl-2, we also observed diminished cleaved caspase-3. TLR2 signaling has been shown to activate PI3K/Akt[39] and can regulate Bcl-2 via phosphorylation of Stat3[41–43]. However, we found that Akt, and Stat3 were not phosphorylated in the TLR6KO mice at day 28 of AOM/DSS were unchanged and may not be involved in the TLR6 signaling cascade leading to Bcl-2 activation.

Altogether, our findings highlight several important concepts and introduce novel paradigms in the relationship between chronic inflammation and its effects on the microbiota. First, microbiota-associated therapies can protect against tumorigenesis if they are specific and personal. Here, this effect was mediated through a TLR6-dependent induction of anti-inflammatory IL-10 and by altering the abundance of cancer-promoting commensals. Second, the genotype of an individual shapes the composition of the microbiota under chronic inflammatory conditions, biases the cytokine milieu of the tissue microenvironment and influences the response to microbiota-associated therapies. Lastly, our study establishes TLR6 as an important molecule involved in modulating the microbiota and immune response during inflammation-associated CRC. We originally examined TLR6 because of its profound requirement for IL-10 induction by *Lactobacillus* but discovered that TLR6 signaling influences the kinetics of compositional changes within the microbiota during inflammation. The strong influence of genetics and environment on the microbiota and the response to microbiota-associated therapy suggest that a personalized approach to understanding the microbiota during disease will be critical to the development of new microbiota-associated therapies.

4.6. SUPPLEMENTAL MATERIAL



Supplemental Figure 4.1: TLR6 signaling impacts dysbiosis associated with AOM/DSS.

(**A**) Proportion of genus-level bacteria affected by the interaction of housing and genotype obtained by 16S rDNA sequencing. (**B**) Colonic contents were scraped from mice at the indicated time points. DNA was amplified using bacterial family specific primers for Helicobacteraceae, Porphyromonadaceae and Lactobacillaceae. Data is the mean ± s.e.m. of 5-7 mice pooled over two experiments. *, $p < 0.05$, **, $p < 0.01$, Two-way ANOVA with individual post-hoc tests (**A**) and Student's t-test (**B**).

4.7. REFERENCES

1. Siegel RL, Miller KD, Fedewa SA, et al. Colorectal cancer statistics, 2017. *CA Cancer J Clin*. 2017;67(3):177-193. doi:10.3322/caac.21395

2. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and Heritable Factors in the Causation of Cancer — Analyses of Cohorts of Twins from Sweden, Denmark, and Finland. *N Engl J Med*. 2000;343(2):78-85. doi:10.1056/NEJM200007133430201

3. Grady WM. Genetic testing for high-risk colon cancer patients. *Gastroenterology*. 2003;124(6):1574-1594. doi:10.1016/s0016-5085(03)00376-7

4. Elinav E, Nowarski R, Thaiss CA, Hu B, Jin C, Flavell RA. Inflammation-induced cancer: crosstalk between tumours, immune cells and microorganisms. *Nat Rev Cancer*. 2013;13(11):759-771. doi:10.1038/nrc3611

5. Gupta RB, Harpaz N, Itzkowitz S, et al. Histologic inflammation is a risk factor for progression to colorectal neoplasia in ulcerative colitis: a cohort study. *Gastroenterology*. 2007;133(4):1099-1105; quiz 1340-1341. doi:10.1053/j.gastro.2007.08.001

6. Shalapour S, Karin M. Immunity, inflammation, and cancer: an eternal fight between good and evil. *J Clin Invest*. 2015;125(9):3347-3355. doi:10.1172/JCI80007

7. Manichanh C, Borruel N, Casellas F, Guarner F. The gut microbiota in IBD. *Nat Rev Gastroenterol Hepatol*. 2012;9(10):599-608. doi:10.1038/nrgastro.2012.152

8. Maynard CL, Elson CO, Hatton RD, Weaver CT. Reciprocal interactions of the intestinal microbiota and immune system. *Nature*. 2012;489(7415):231-241. doi:10.1038/nature11551

9. Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. Host-bacterial mutualism in the human intestine. *Science*. 2005;307(5717):1915-1920. doi:10.1126/science.1104816

10. Arthur JC, Perez-Chanona E, Mühlbauer M, et al. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science*. 2012;338(6103):120-123. doi:10.1126/science.1224820

11. Shen XJ, Rawls JF, Randall T, et al. Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. *Gut Microbes*. 2010;1(3):138-147. doi:10.4161/gmic.1.3.12360

12. Chen W, Liu F, Ling Z, Tong X, Xiang C. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PLoS ONE*. 2012;7(6):e39743. doi:10.1371/journal.pone.0039743

13. Zackular JP, Baxter NT, Iverson KD, et al. The gut microbiome modulates colon tumorigenesis. *MBio*. 2013;4(6):e00692-00613. doi:10.1128/mBio.00692-13

14. Flanagan L, Schmid J, Ebert M, et al. Fusobacterium nucleatum associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. *Eur J Clin Microbiol Infect Dis*. 2014;33(8):1381-1390. doi:10.1007/s10096-014-2081-3

15. Castellarin M, Warren RL, Freeman JD, et al. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res*. 2012;22(2):299-306. doi:10.1101/gr.126516.111

16. Sears CL, Geis AL, Housseau F. Bacteroides fragilis subverts mucosal biology: from symbiont to colon carcinogenesis. *J Clin Invest*. 2014;124(10):4166-4172. doi:10.1172/JCI72334

17. Wu S, Rhee K-J, Albesiano E, et al. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat Med*. 2009;15(9):1016-1022. doi:10.1038/nm.2015

18. Dalmasso G, Cougnoux A, Delmas J, Darfeuille-Michaud A, Bonnet R. The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment. *Gut Microbes*. 2014;5(5):675-680. doi:10.4161/19490976.2014.969989

19. Uronis JM, Mühlbauer M, Herfarth HH, Rubinas TC, Jones GS, Jobin C. Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. *PLoS ONE*. 2009;4(6):e6026. doi:10.1371/journal.pone.0006026

20. Vannucci L, Stepankova R, Kozakova H, Fiserova A, Rossmann P, Tlaskalova-Hogenova H. Colorectal carcinogenesis in germ-free and conventionally reared rats: different intestinal environments affect the systemic immunity. *Int J Oncol*. 2008;32(3):609-617.

21. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol*. 2010;11(5):373-384. doi:10.1038/ni.1863

22. Kordahi MC, William DePaolo R. The Influence of the Microbiota on the Etiology of Colorectal Cancer. In: Sun J, Dudeja PK, eds. *Mechanisms Underlying Host-Microbiome Interactions in Pathophysiology of Human Diseases*. Boston, MA: Springer US; 2018:167-193. doi:10.1007/978-1-4939-7534-1_8

23. Rasmussen SB, Reinert LS, Paludan SR. Innate recognition of intracellular pathogens: detection and activation of the first line of defense. *APMIS*. 2009;117(5-6):323-337. doi:10.1111/j.1600-0463.2009.02456.x

24. Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, Edberg S, Medzhitov R. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell*. 2004;118(2):229-241. doi:10.1016/j.cell.2004.07.002

25. Round JL, Lee SM, Li J, et al. The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. *Science*. 2011;332(6032):974-977. doi:10.1126/science.1206095

26. Chow J, Tang H, Mazmanian SK. Pathobionts of the gastrointestinal microbiota and inflammatory disease. *Curr Opin Immunol*. 2011;23(4):473-480. doi:10.1016/j.coi.2011.07.010

27. Trinchieri G. Cancer and Inflammation: An Old Intuition with Rapidly Evolving New Concepts. *Annu Rev Immunol*. 2012;30(1):677-706. doi:10.1146/annurev-immunol-020711-075008

28. Takeuchi O, Kawai T, Mühlradt PF, et al. Discrimination of bacterial lipoproteins by Toll-like receptor 6. *Int Immunol*. 2001;13(7):933-940. doi:10.1093/intimm/13.7.933

29. Takeuchi O, Sato S, Horiuchi T, et al. Cutting edge: role of Toll-like receptor 1 in mediating immune response to microbial lipoproteins. *J Immunol*. 2002;169(1):10-14. doi:10.4049/jimmunol.169.1.10

30. Lowe EL, Crother TR, Rabizadeh S, et al. Toll-like receptor 2 signaling protects mice from tumor development in a mouse model of colitis-induced cancer. *PLoS ONE*. 2010;5(9):e13027. doi:10.1371/journal.pone.0013027

31. Kamdar K, Khakpour S, Chen J, et al. Genetic and Metabolic Signals during Acute Enteric Bacterial Infection Alter the Microbiota and Drive Progression to Chronic Inflammatory Disease. *Cell Host Microbe*. 2016;19(1):21-31. doi:10.1016/j.chom.2015.12.006

32. Janakiram NB, Rao CV. The role of inflammation in colon cancer. *Adv Exp Med Biol*. 2014;816:25-52. doi:10.1007/978-3-0348-0837-8_2

33. Bansal P, Sonnenberg A. Risk factors of colorectal cancer in inflammatory bowel disease. *Am J Gastroenterol*. 1996;91(1):44-48.

34. Bedi A, Pasricha PJ, Akhtar AJ, et al. Inhibition of apoptosis during development of colorectal cancer. *Cancer Res*. 1995;55(9):1811-1816.

35. You X, Liu L, Zeng Y, et al. Macrophage-activating lipopeptide-2 requires Mal and PI3K for efficient induction of heme oxygenase-1. *PLoS ONE*. 2014;9(7):e103433. doi:10.1371/journal.pone.0103433

36. Alessi DR, Kozlowski MT, Weng QP, Morrice N, Avruch J. 3-Phosphoinositide-dependent protein kinase 1 (PDK1) phosphorylates and activates the p70 S6 kinase in vivo and in vitro. *Curr Biol*. 1998;8(2):69-81. doi:10.1016/s0960-9822(98)70037-5

37. Ohbayashi N, Ikeda O, Taira N, et al. LIF- and IL-6-induced acetylation of STAT3 at Lys-685 through PI3K/Akt activation. *Biol Pharm Bull*. 2007;30(10):1860-1864. doi:10.1248/bpb.30.1860

38. Zushi S, Shinomura Y, Kiyohara T, et al. STAT3 mediates the survival signal in oncogenic ras-transfected intestinal epithelial cells. *Int J Cancer*. 1998;78(3):326-330. doi:10.1002/(SICI)1097-0215(19981029)78:3<326::AID-IJC12>3.0.CO;2-4

39. Leslie K, Lang C, Devgan G, et al. Cyclin D1 is transcriptionally regulated by and required for transformation by activated signal transducer and activator of transcription 3. *Cancer Res*. 2006;66(5):2544-2552. doi:10.1158/0008-5472.CAN-05-2203

40. Barton BE, Karras JG, Murphy TF, Barton A, Huang HF-S. Signal transducer and activator of transcription 3 (STAT3) activation in prostate cancer: Direct STAT3 inhibition induces apoptosis in prostate cancer lines. *Mol Cancer Ther*. 2004;3(1):11-20.

41. Faith JJ, Guruge JL, Charbonneau M, et al. The long-term stability of the human gut microbiota. *Science*. 2013;341(6141):1237439. doi:10.1126/science.1237439

42. Lee SM, Donaldson GP, Mikulski Z, Boyajian S, Ley K, Mazmanian SK. Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature*. 2013;501(7467):426-429. doi:10.1038/nature12447

43. Fiorentino DF, Zlotnik A, Mosmann TR, Howard M, O'Garra A. IL-10 inhibits cytokine production by activated macrophages. *J Immunol*. 1991;147(11):3815-3822.

44. Mocellin S, Panelli MC, Wang E, Nagorsen D, Marincola FM. The dual role of IL-10. *Trends Immunol*. 2003;24(1):36-43.

45. Ballal SA, Veiga P, Fenn K, et al. Host lysozyme-mediated lysis of Lactococcus lactis facilitates delivery of colitis-attenuating superoxide dismutase to inflamed colons. *Proc Natl Acad Sci USA*. 2015;112(25):7803-7808. doi:10.1073/pnas.1501897112

46. Peña JA, Rogers AB, Ge Z, et al. Probiotic Lactobacillus spp. diminish Helicobacter hepaticus-induced inflammatory bowel disease in interleukin-10-deficient mice. *Infect Immun*. 2005;73(2):912-920. doi:10.1128/IAI.73.2.912-920.2005

47. Thirabunyanon M, Boonprasom P, Niamsup P. Probiotic potential of lactic acid bacteria isolated from fermented dairy milks on antiproliferation of colon cancer cells. *Biotechnol Lett*. 2009;31(4):571-576. doi:10.1007/s10529-008-9902-3

48. Chong ESL. A potential role of probiotics in colorectal cancer prevention: review of possible mechanisms of action. *World J Microbiol Biotechnol*. 2014;30(2):351-374. doi:10.1007/s11274-013-1499-6

49. DePaolo RW, Kamdar K, Khakpour S, Sugiura Y, Wang W, Jabri B. A specific role for TLR1 in protective TH17 immunity during mucosal infection. *Journal of Experimental Medicine*. 2012;209(8):1437-1444. doi:10.1084/jem.20112339

50. Volzing K, Borrero J, Sadowsky MJ, Kaznessis YN. Antimicrobial peptides targeting Gram-negative pathogens, produced and delivered by lactic acid bacteria. *ACS Synth Biol*. 2013;2(11):643-650. doi:10.1021/sb4000367

51. Moolenbeek C, Ruitenberg EJ. The "Swiss roll": a simple technique for histological studies of the rodent intestine. *Lab Anim*. 1981;15(1):57-59.

52. Suzuki R, Kohno H, Sugie S, Nakagama H, Tanaka T. Strain differences in the susceptibility to azoxymethane and dextran sodium sulfate-induced colon carcinogenesis in mice. *Carcinogenesis*. 2006;27(1):162-169. doi:10.1093/carcin/bgi205

53. Kennedy RJ, Hoper M, Deodhar K, Erwin PJ, Kirk SJ, Gardiner KR. Interleukin 10-deficient colitis: new similarities to human inflammatory bowel disease. *Br J Surg*. 2000;87(10):1346-1351. doi:10.1046/j.1365-2168.2000.01615.x

54. Boivin GP, Washington K, Yang K, et al. Pathology of mouse models of intestinal cancer: consensus report and recommendations. *Gastroenterology*. 2003;124(3):762-777. doi:10.1053/gast.2003.50094

55. Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*. 2009;75(23):7537-7541.

56. Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ. Removing noise from pyrosequenced amplicons. *BMC Bioinformatics*. 2011;12:38.

57. Pruesse E, Quast C, Knittel K, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res*. 2007;35(21):7188-7196.

58. Schloss PD, Gevers D, Westcott SL. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS One*. 2011;6(12):e27310.

59. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*. 2011;27(16):2194-2200.

60. Cole JR, Wang Q, Cardenas E, et al. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res*. 2009;37(Database issue):D141-

5. Conclusions and Perspectives

In this Chapter, I will be summarizing key concepts about the relationship of our gut microbes with our bodies in general and the role that they play in the etiology of colorectal cancer in particular. As previously mentioned, the human "microbiome" refers to the microorganisms colonizing our bodies as well as their genomes and contributes significantly to our health and diseases[1]. One of the human body sites inhabited by a particularly large microbial community is the digestive tract, which contains at least one hundred billion bacteria with the largest number of gut microbes residing on the surface of the mucous membrane of the colon. The intestinal commensal microflora represents an interface of the human host with environmental factors as well as a translating agent of these factors for the body, making it of great importance to the health of the host[2]. It can aid digestion through the metabolism of carbohydrates and amino acids, the elimination of xenobiotics, the production of vitamins, and local intestinal peristalsis[3]. Our intestinal commensals also play a significant role in the growth and differentiation of intestinal epithelial cells, in the resistance to colonization and elimination of pathogens, as well as in the maturation of the immune system through ensuring tolerance to food and bacterial antigens[4]. Shifts in the gut microbiome composition, gene expression, and metabolism due to environmental factors as well as host factors have been linked to various disease states such as obesity, diabetes, metabolic syndrome, inflammatory bowel diseases, and colorectal cancer[5].

CRC alone remains a major cause of morbidity and mortality worldwide. Mortality from CRC is generally associated with advanced metastatic disease, when treatment is relatively ineffective[6]. The disease occurs in a stepwise fashion beginning with abnormal cell proliferation leading to the development of adenomatous polyps, which are CRC precursors[7]. Despite the availability of modern therapies, the mortality rate from CRC remains high, prompting an urgent need for better biomarkers and more effective treatments. In that regards, understanding the involvement of the microbiota in CRC development holds the promise of unraveling microbiome-based diagnostics and therapeutic targets.

In 1974, Reddy et al. suggested for the very first time a relationship between commensal bacterial and intestinal carcinogenesis in a study involving germ-free rats[8]. Ever since, mounting evidence from metagenomics analyses and animal models have continued to support the role of the intestinal microbiome in CRC[9]. Moreover, many theories have been elaborated about the involvement of certain specific bacterial species in CRC tumorigenesis since McCoy and Mason

first suggested a link between Enterococcus and carcinoma of the sigmoid in the middle of the 20th century[10]. In 2011, Sears and Pardoll proposed the 'alpha-bug' hypothesis, in which species such as *Bacteroides fragilis* exert a fundamental pro-oncogenic, enterotoxigenic role, thereby contributing to the start of CRC[11]. Then, Tjalsma et al. proposed the driver–passenger model in 2012, in which a driver bacterium (e.g., B. fragilis) can lead to a multistep progress of carcinogenesis through inflammation, increased cellular proliferation, and/or the production of genotoxins in the colon[12]. Based on this theory, Hajishengallis et al. proposed the 'Keystone hypothesis', in which 'key pathogens', even at low abundance, facilitate colonization by 'accessory pathogens'[13]. This invasion is followed by the destabilization of host responses, resulting in 'dysbiosis' of the microbiota where certain other microorganisms can contribute to more inflammation. Lastly, another theory suggests a 'hit-and-run' action of 'tumor-initiating' bacteria. The best example of this mechanism is probably *Helicobacter pylori* infections in gastric cancers, whereby the bacterial toxin CagA exerts a pro-oncogenic action leading to epigenetic alterations in the gastric mucosa[14]. This model suggests that potentially passing pathogenic bacteria are required for the initiation, but not for the maintenance, of a cancerous phenotype.

Despite the fact that further efforts are still required to resolve the entire underlying pro-oncogenic mechanisms linking bacteria to CRC, we do know that certain bacteria may affect CRC development through directly or indirectly interacting with host cells. The ways these interactions occur involve invasion of the relatively sterile mucus layer lining the intestinal epithelium and adhesion to the intestinal mucosa with biofilm formation, translocation of bacteria and bacterial factors across the epithelial barrier of the gut, host-defense modulation through bacteria-immune cells interactions, and the production of bacterial metabolites such as extracellular superoxide, genotoxins, or short chain fatty acids. As dietary substrates and other environmental modifiers mediate changes in the microbiota, the inflammation of the intestinal mucosa that ensues allows invasion of the mucus layer by various bacterial invaders[15]. Subacute chronic inflammation of the intestinal epithelium will induce damage to the epithelial barrier overtime, compromising its integrity. These changes in the permeability of the epithelial lining of the colon will allow certain commensal bacteria and bacterial factors to come in contact with the mucosal innate immune cells. The innate immune system uses pattern recognition receptors, which have the ability to identify microbial commensals by molecular structure. These receptors include TLRs, a class of proteins expressed on antigen-presenting cells, macrophages, and dendritic cells (DCs). In the situation of intestinal barrier damage by microorganisms, antigen presenting cells recognize these

microorganisms through TLR receptors, causing the secretion of cytokines activating the immune response. For example, it has been proven and further shown in this body of work that TLR2 binding plays an important role in the carcinogenesis of colorectal cancer, through inducing inflammation and hyperproliferation of the epithelium[16]. Bacterial lipopolysaccharide (LPS) in the gut is a bacterial molecular structure that has been shown to strongly stimulate innate immune signaling, thereby further compromising gut homeostasis and normal host physiology[4]. In a cross-sectional study looking at patients with adenoma, it was shown that blood levels of LPS, likely due to leakage of microbial products, were higher than in control individuals[17]. A polymorphism in the LPS-binding protein (*LBP*) gene (rs2232596) was also associated with higher risk of CRC[18]. A large prospective study of mostly white individuals also found that higher levels of LPS were associated with increased risk of CRC in men[19] whereas a prospective study in a racially diverse cohort observed that, compared with individuals in the first quartile of plasma LBP level, those in the third, but not fourth, quartile had an increased risk of CRC[20].

All in all, researchers have generated exciting preliminary data for the role of the microbiota in CRC development, prevention, and treatment to this day. So, where do we really stand regarding the actual use of microbiome-biomarkers and therapeutics in the clinic? First, we do know that CRC is a disease that is mostly sporadic, stemming from the host's lifestyle factors. Obesity and physical inactivity affect the intestinal micro metabolome and its role in colorectal carcinogenesis[21]. Red and processed meat may also increase CRC risk by increased bacterial production of secondary bile acids, hydrogen sulfide (H_2S), and trimethylamine oxide (TMAO)[22]. On the other hand, dietary fibers were shown to be beneficial as they promote and enrichment of SCFA-producing bacteria and increased production of SCFAs that modulate the intestinal immune and metabolic response and inhibit CRC development[23]. So, first and foremost, ensuring a healthy diet and lifestyle is a protective behavioral approach that significantly decreases the risk of developing CRC.

Fecal microbiome analysis offers the promise of a quick and non-invasive way of diagnosing CRC, when compared to the more invasive process of undergoing a colonoscopy, that is the gold-standard for CRC diagnosis today. Fecal microbiome analysis can identify patients with adenomas with reasonable levels of accuracy (with area under the receiver operating curves ranging from 0.55 to 0.67 in validation studies), although this is a lower value than for detection of CRC[24]. The fecal microbiome data with scores from risk factor–based models or results of screening tests (such as fecal occult blood test and fecal immunochemical test) increases the accuracy of detection for

advanced adenomas[25] (For example, addition of fecal *Fusobacterium nucleatum* quantitation to fecal immunochemical testing). However, before fecal microbiomes can be reliably used in CRC screening, it is important to realize the incredibly large patients' heterogeneity complicating studies using microbial features to construct analysis models. Moreover, the variations in the technical aspects of these studies (such as stool collection methods, timing and type of bowel preparation for colonoscopy, and sequencing and analysis methods) constitute further challenges that need to be considered. Moreover, as studies are increasingly focusing on tissue-adherent microbiome for its physiological relevance, the fecal microbiome might ultimately become less relevant. An alternative option would be to focus on sequencing colonic biopsies harvested during colonoscopies instead and assessing whether the data gathered, and the patterns identified can help improve the sensitivity of detection of colon neoplasia. The quantification of systemic LPS in the bloodstream might also hold promising potential. So does the detection of IL-8 and IL-12p40, as our study in Chapter one of this Thesis clearly confirmed elevated levels of both cytokines in patients with early colon neoplasia.

In addition to affecting CRC development, the gut microbiota has been shown to modulate the response to cancer therapy and susceptibility to toxic adverse effects[26]. For example, Iida et al. reported in 2013 that intestinal dysbiosis can affect the efficacy of an immunotherapy (CpG-oligonucleotide-tide) and oxaliplatin, a platinum compound used in chemotherapy for CRC and other cancers[27]. Both therapies had reduced efficacy in germ-free mice. *F. nucleatum* has also been shown to promote resistance of CRC cells to chemotherapy[28]. On the other hand, sphingolipid containing *Bacteroides thetaiotaomicron* and certain *Bacteroides fragilis* been associated with less inflammation and promotion of intestinal homeostasis[29]. Certain Bacteroides species have also been associated with higher efficacy of cytotoxic T-lymphocyte associated protein 4 blockers (CTLA4), immunotherapy, possibly by affecting interleukin 12–dependent-helper 1 cell–mediated immune responses[30]. *Bifidobacterium, A. muciniphilia*, and *Faecalibacterium* were all shown to increase the efficacy of programmed cell death 1 (PDCD1) blocking agents, possibly by increasing recruitment of T cells to tumors and their functions there[31,32,33]. TLR6 signaling by lactic acid bacteria was also shown to reduce tumors in our study published in cancer research prevention earlier this year[34]. Moreover, specific bacteria have been associated with resistance to the development of immune checkpoint inhibitor–associated colitis, and fecal microbiota transplants might be used to treat this form of colitis[35]. Our study in Chapter one showed that mucosal

Bacteroides fragilis can trigger high levels of chemokine IL-8 in the colon. IL-8 has been shown to promote CRC cell proliferation, invasion, migration, and angiogenesis[41]. Thus, blocking the IL-8 receptors CXCR1 and CXCR2 could constitute an interesting anti-tumor therapeutic strategy for CRC prevention and management. Lastly, certain dietary interventions can be also considered to modulate the gut microbiota in patients receiving cancer therapy. Higher intake of nutrients such as fiber, marine omega-3 fatty acid, vitamin D, or calcium; coffee; plant-based low-carbohydrate diet all have immunomodulatory and microbiota-modifying effects and have been associated with increased survival times of patients with CRC, as they might increase the efficacy and reduce the adverse effects of immunotherapies or other therapeutic agents.

In order to further understand the intestinal microbiota and its manipulation in CRC screening, prevention, and treatment, more integrated prospective studies and mechanistic studies are needed. Prospective studies with detailed diet and lifestyle data collected long before participants develop CRC are for instance needed in order to better characterize the long-term influence of environmental exposures on the gut microbiome and their effects on CRC prevention. Furthermore, we need to uncover the specific mechanisms by which these diet and lifestyle factors influence the gut microbiome and risk of CRC. We happen to have at our disposal an arsenal of tools ranging from *in vitro* to *in vivo* models such as gnotobiotic mouse models, gut-on-chip models, and 3D intestinal organoids. We also have multi-Omic approaches from genomics, to transcriptomics, proteomics, and metabolomics. Paired with functional data stemming from mechanistic models, these data can be integrated in computational modeling in order to study the 'genotype' to 'phenotype' relations as well as the microbe-microbe and host-microbe interactions. Thus, this field of research holds the promise that the microbiome may one day be integrated as a biomarker into everyday diagnostics in order to better the current tools at our disposal for diagnosing CRC and treating early. One can also anticipate a future where microbiome data will be fully integrated to diagnose CRC and predict personalized treatment strategies in order to increase patient quality of life and survival.

5.1. REFERENCES:

1. Wang, B., Yao, M., Lv, L., Ling, Z. & Li, L. The Human Microbiota in Health and Disease. *Engineering* **3**, 71–82 (2017).

2. Cho, I. & Blaser, M. J. The Human Microbiome: at the interface of health and disease. *Nat Rev Genet* **13**, 260–270 (2012).
3. Rowland, I. *et al.* Gut microbiota functions: metabolism of nutrients and other food components. *Eur J Nutr* **57**, 1–24 (2018).
4. Belkaid, Y. & Hand, T. Role of the Microbiota in Immunity and inflammation. *Cell* **157**, 121–141 (2014).
5. Hills, R. D. *et al.* Gut Microbiome: Profound Implications for Diet and Disease. *Nutrients* **11**, (2019).
6. Mariotto, A. B., Robin Yabroff, K., Shao, Y., Feuer, E. J. & Brown, M. L. Projections of the Cost of Cancer Care in the United States: 2010–2020. *J Natl Cancer Inst* **103**, 117–128 (2011).
7. Dulal, S. & Keku, T. O. Gut Microbiome and Colorectal adenomas. *Cancer J* **20**, 225–231 (2014).
8. Reddy, B. S., Weisburger, J. H., Narisawa, T. & Wynder, E. L. Colon Carcinogenesis in Germ-free Rats with 1,2-Dimethylhydrazine and N-Methyl-N′-nitro-N-nitrosoguanidine. *Cancer Res* **34**, 2368–2372 (1974).
9. Wong, S. H. & Yu, J. Gut microbiota in colorectal cancer: mechanisms of action and clinical applications. *Nat Rev Gastroenterol Hepatol* **16**, 690–704 (2019).
10. McCOY, W. C. & Mason, J. M. Enterococcal endocarditis associated with carcinoma of the sigmoid; report of a case. *J Med Assoc State Ala* **21**, 162–166 (1951).
11. Sears, C. L. & Pardoll, D. M. Perspective: Alpha-Bugs, Their Microbial Partners, and the Link to Colon Cancer. *J Infect Dis* **203**, 306–311 (2011).
12. Tjalsma, H., Boleij, A., Marchesi, J. R. & Dutilh, B. E. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nat. Rev. Microbiol.* **10**, 575–582 (2012).
13. Hajishengallis, G., Darveau, R. P. & Curtis, M. A. The Keystone Pathogen Hypothesis. *Nat Rev Microbiol* **10**, 717–725 (2012).
14. Ding, S.-Z., Goldberg, J. B. & Hatakeyama, M. Helicobacter pylori infection, oncogenic pathways and epigenetic mechanisms in gastric carcinogenesis. *Future Oncol* **6**, 851–862 (2010).
15. Chassaing, B. & Gewirtz, A. T. Identification of Inner Mucus-Associated Bacteria by Laser Capture Microdissection. *Cell Mol Gastroenterol Hepatol* **7**, 157–160 (2018).

16. Li, T.-T., Ogino, S. & Qian, Z. R. Toll-like receptor signaling in colorectal cancer: Carcinogenesis to cancer therapy. *World J Gastroenterol* **20**, 17699–17708 (2014).
17. Kang, M. *et al.* Association of plasma endotoxin, inflammatory cytokines and risk of colorectal adenomas. *BMC Cancer* **13**, 91 (2013).
18. Chen, R., Luo, F.-K., Wang, Y.-L., Tang, J.-L. & Liu, Y.-S. LBP and CD14 polymorphisms correlate with increased colorectal carcinoma risk in Han Chinese. *World J. Gastroenterol.* **17**, 2326–2331 (2011).
19. Kong, S. Y. *et al.* Serum Endotoxins and Flagellin and Risk of Colorectal Cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC) Cohort. *Cancer Epidemiol. Biomarkers Prev.* **25**, 291–301 (2016).
20. Citronberg, J. S. *et al.* Plasma lipopolysaccharide-binding protein and colorectal cancer risk: a nested case-control study in the Multiethnic Cohort. *Cancer Causes Control* **29**, 115–123 (2018).
21. Russell, W. R., Duncan, S. H. & Flint, H. J. The gut microbial metabolome: modulation of cancer risk in obese individuals. *Proc. Nutr. Soc.* **72**, 178–188 (2013).
22. Oostindjer, M. *et al.* The role of red and processed meat in colorectal cancer development: a perspective. *Meat Science* **97**, 583–596 (2014).
23. den Besten, G. *et al.* The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* **54**, 2325–2340 (2013).
24. Song, M., Chan, A. T. & Sun, J. Influence of the Gut Microbiome, Diet, and Environment on Risk of Colorectal Cancer. *Gastroenterology* **158**, 322–340 (2020).
25. Zackular, J. P., Rogers, M. A. M., Ruffin, M. T. & Schloss, P. D. The human gut microbiome as a screening tool for colorectal cancer. *Cancer Prev Res (Phila)* **7**, 1112–1121 (2014).
26. Vivarelli, S. *et al.* Gut Microbiota and Cancer: From Pathogenesis to Therapy. *Cancers (Basel)* **11**, (2019).
27. Iida, N. *et al.* Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* **342**, 967–970 (2013).
28. Mima, K. *et al.* Fusobacterium nucleatum and T-cells in Colorectal Carcinoma. *JAMA Oncol* **1**, 653–661 (2015).
29. Brown, E. M. *et al.* Bacteroides-Derived Sphingolipids Are Critical for Maintaining Intestinal Homeostasis and Symbiosis. *Cell Host & Microbe* **25**, 668-680.e7 (2019).

30. Shui, L. *et al.* Gut Microbiome as a Potential Factor for Modulating Resistance to Cancer Immunotherapy. *Front Immunol* **10**, (2020).
31. Sivan, A. *et al.* Commensal Bifidobacterium promotes antitumor immunity and facilitates anti–PD-L1 efficacy. *Science* **350**, 1084–1089 (2015).
32. Routy, B. *et al.* Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97 (2018).
33. Gopalakrishnan, V. *et al.* Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97–103 (2018).
34. Kim, J.-H., Kordahi, M. C., Chac, D. & DePaolo, R. W. Toll-like Receptor-6 Signaling Prevents Inflammation and Impacts Composition of the Microbiota During Inflammation-Induced Colorectal Cancer. *Cancer Prev Res* (2019) doi:10.1158/1940-6207.CAPR-19-0286.
35. Som, A. *et al.* Immune checkpoint inhibitor-induced colitis: A comprehensive review. *World J Clin Cases* **7**, 405–418 (2019).