BIOFILMS OF COLORECTAL CANCER

By

Christine Michelle Dejea

A dissertation submitted to Johns Hopkins University in
conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
January, 2015

ABSTRACT

The human colon hosts a diverse and metabolically complex community of microorganisms. While the colonic microbiome has long been suggested to contribute to the development of colorectal cancer (CRC), a definitive link has not been made. In this thesis, studies that define the bacterial associations of CRC in the genetically normal host (sporadic CRC) and the host predisposed to early colon carcinogenesis [familial adenomatous polyposis (FAP)] are presented. We demonstrate for the first time that bacterial biofilms are a common feature of many sporadic colorectal cancers, one of the leading malignancies in the United States and abroad. Colon biofilms, dense communities of bacteria encased in a mucus matrix that contact the colon epithelial cells, are a nearly universal feature of sporadic right colon tumors with broad extension over the right colon in the tumor host. Most remarkably, biofilm presence correlates with bacterial tissue invasion and a change in tissue biology including changes in oncogenic signaling pathways [E-cadherin, IL-6 and signal transducer and activator of transcription 3 (Stat3)], enhanced polyamine metabolites and increased cellular proliferation, changes capable of promoting oncogenic transformation. Deep sequencing revealed that biofilm communities on paired normal mucosa cluster with tumor microbiomes but lack distinct taxa differences. In striking contrast, biofilms were detected throughout the colons of individuals with the hereditary CRC condition (FAP) and lacked the diverse composition of biofilms associated with sporadic CRC. Namely, biofilms of FAP patients are comprised largely of pathogenic subtypes of *Escherichia coli* (*pks*+) and *Bacteroides fragilis* (*bft*+). This work introduces a new concept whereby the microbial community structural organization contributes to disease progression.

ACKNOWLEDGEMENTS

I would like to start by thanking my thesis advisor and mentor, Dr. Cynthia Sears, for taking me into her lab in my third year as a PhD student and offering me the opportunity to work on a project I am truly passionate about. As a mentor, Cindy has provided advice, assistance, support and respect throughout my graduate career. Since joining the lab I've had a lot of major life events: I got married, I lost one of the most important people in my life to cancer, and I had my first child. Cindy has been nothing but supportive and forgiving of the time they require (especially Chase), she is truly an example of work-life balance. I really appreciate everything she has done for me; she has been instrumental in my development as a scientist and a person. I would also like to thank Dr. Alan Scott for working with me so closely during my time as a graduate student. He has been an important secondary mentor for me and I appreciate and value his advice, guidance and support. Thank you to the remaining members of my committee, Dr. Drew Pardoll and Dr. Fengyi Wan. My project has always been a team effort and your discussions and guidance have been a huge help in its development. I owe a ton of thanks to the current and past members of the Sears lab. Thank you Shaoguang Wu, Sara White, Annemarie Boleij, Lizzy Hechenbleikner, Xinqun Wu, Keno Onwueme, LeaAnn Chen, June Chan, Payam Fathi, for helpful discussions, editing, and just making lab a fun and interesting place to come to every day. A special thank you to Liza Wick, for always being available for discussions (scientific and not). I would also like to thank the members of our extended (Pardoll) lab: Franck Housseau, Abby Geis, Nico Llosa and Erik Thiele Orberg for always being there for discussions and experimental help.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

General Introduction

1.1 Preface

Each year, approximately 1.2 million individuals are diagnosed with colon cancer worldwide [1]. As the second leading cancer affecting both men and women, colorectal cancer (CRC) claims the lives of over 600,000 individuals annually [1]. Greater than 90% of CRC cases are spontaneous, occurring in people with little or no family history of the disease. Once thought to be a cancer predominantly afflicting the Western world, incidence rates of CRC are rapidly increasing in countries that have historically been considered low risk, including South America, eastern Asia, and Eastern Europe [1]. This trend has been attributed to changes in dietary patterns, along with decreased physical activity, leading to a rise in obesity within these populations [2,3]. As a prominent public health threat, potential contributions to the development of CRC have been the focus of intense study. Colorectal carcinomas usually begin as benign tumors, called polyps or adenomas, which can develop anywhere along the colon from the epithelial cells lining the mucosa. Typically over a period of 10 or more years, some polyps become cancers. Importantly, however, colon cancer can be fully prevented by early detection and removal of polyps. This progression from normal epithelium to adenoma to adenocarcinoma has been well characterized by Fearon and Vogelstein to involve the cumulative accumulation of genetic mutations [4]. The proposed classes of optimal target genes include tumor suppressors and oncogenes, along with mismatch repair genes. Common examples include *APC* (a tumor-supressor gene), *KRAS* (an oncogene), and *MLH1 and MLH2* (mismatch repair genes) [4,5]. While there is general consensus about the stepwise transition to colorectal carcinoma, the initiating mechanism(s) remain unclear.

The notion that the endogenous enteric microbiome contributes to the etiopathogenesis of colon cancer has been proposed for decades. The human gastrointestinal (GI) tract is colonized by a vast and complex community of microorganisms totaling approximately 10^{13} bacteria composed of over 500 microbial species [6]. The commensal intestinal microbiota outnumbers human cells nine to one, and perhaps more impressively, their collective genes outnumber that of their human host 100 to one. The influences of this large microbial community are immense. The colon is colonized soon after birth, facilitating the essential roles played by the colon microbiota in host physiology, including mucosal immune development, regulation of cell proliferation and modulation of gene expression in host epithelial cells [7,8]. Other beneficial functions of the metabolically complex microbiome include providing usable forms of nutrients as a byproduct of metabolism and protection against exogenous pathogens. In the healthy colon, the microbiota interactions with the host are at homeostasis; however, intrinsic or extrinsic factors can cause perturbations, leading to abnormalities in microbiome composition or function that have been associated with several diseases, including inflammatory bowel disease (IBD) and colon cancer [9,10].

The entirety of the healthy human colon is covered by a mucus layer that consists of an inner gel-like layer and a loose outer layer, both primarily composed of a secreted network of highly glycosylated MUC2 mucins. Among the family of mucin genes expressed in the human colon, the gene product of *MUCB* has also been detected in minor quantities at the base of the crypt [11]. In addition, MUC5AC and MUC6 have been associated with colorectal adenomas and ulcerative colitis [12]. The outer mucus layer serves as a semipermeable network providing a habitat for commensal bacteria to reside in, while the inner gel-like mucus layer acts as a physical barrier excluding bacteria from direct contact with the epithelium [13]. It is likely that bacteria transiently penetrate the inner mucus barrier

in a healthy state; however, they are thought to be cleared quickly through host immune responses [13]. The inner mucus layer ranges in width from 30 to 170 μm in the human colon, increasing in depth from the ascending to the descending colon [14]. Bacteria mainly colonize two major niches within the human colon: the lumen and the outer mucus layer. Characterization of these distinct microbial communities has been the focus of a series of recent studies [6,15-17]. The distinction between these communities is important, as the microbial milieu in these two ecological niches may contribute differently to the etiology of disease. It is well accepted that microbial dysbiosis (an imbalance of the microbiota) with bacterial invasion and persistence in the inner mucus layer (biofilm formation), contributes to the development or progression of IBD [9,18,19]. Massive bacterial biofilms within the normally empty mucus layer, constituting invasions of greater than 10^9 bacteria/ml, were identified in 94% of Ulcerative Colitis patients, 98% of Crohn's Disease patients, and 78% of self-limiting colitis patients, compared to just 11% in controls [19]. The phylum-level 16S profiles were observed to involve a shift in major populations, most notably an increase in Proteobacteria and a decrease in Firmicutes and Bacteroidetes [18]. IBD is associated with an increased risk for the development of GI malignancies. While the development of CRC in the setting of IBD involves many of the same genetic mutations as the stepwise transition to sporadic CRC, the timing and frequency of these mutations often differ [20,21]. Furthermore, chronic colitis-associated CRC tends to be macroscopically heterogeneous compared with sporadic CRC and arises from flat dysplastic tissue rather than distinct polyps [21]. This stresses the importance of the major focus of this dissertation work, characterizing the microbial-epithelial interactions in various CRC disease states, analyses that have been less detailed to date [17,22,23].

1.2 Mechanisms of Bacterial Induced Oncogenesis

A significant amount of effort has been employed to determine the mechanisms of microbially induced oncogenesis. Proposed mechanisms include the inhibition, alteration or exacerbation of normal host responses such as apoptosis, inflammation and cellular proliferation (**CHAPTER 2**). Alternatively, bacteria may also promote cancer through production of secondary metabolites (**CHAPTER 3**), or direct effects on cell transformation through the production of oncogenic toxins (**CHAPTER 4**). An overview of the mechanisms of bacterial initiation or progression of oncogenesis is shown in Figure **1.1.**

Chronic Inflammation

The association between inflammation and tumorigenesis has been appreciated since 1863 when Rudolph Virchow hypothesized that cancer developed from sites of chronic inflammation, termed 'the chronic irritation hypothesis' [24]. Today, the connection between inflammation and cancer is well established; however, the mechanisms and pathways are not fully characterized. Infection triggers inflammation as a means to effectively combat an invading pathogen. Polymorphonuclear phagocytes are typically the first cells recruited to the site of infection, which serve as potent producers of proinflammatory cytokines and chemokines that amplify the response by recruiting more immune cells [25,26]. These cells produce an abundance of reactive oxygen species which can damage lipids, protein, and DNA leading to increased mutations in proliferating cells and ultimately alterations in cell turnover and death [25-28].

While bacterial infection was once thought of as an acute condition, it is clear that many bacteria are able to persist in the host and lead to chronic infections accompanied by inflammation. Study of the molecular mechanisms that link chronic infections to

inflammation and cancer is an area of intense investigation.   Persistent generation of microbially induced inflammation mediators such as TNF-α, IL-1, or even lipopolysaccharide on its own can lead to the induction of the NF- κB family of transcription factors, which have been shown to play a role in inflammation-driven carcinogenesis [27]. Some of the genes targeted are inflammatory cytokines including IL-6, IL-1β, and vascular endothelial growth factor (VEGF), which lead to a positive feedback loop of continuing inflammation [25-27]. Observations using human sera and tissue (including data presented in **CHAPTER 2**), as well as strong evidence through experimental mouse models have shown that IL-6 promotes survival, proliferation and progression to CRC in an inflammatory setting via the signal transducer pSTAT3 [29-31]. In addition, antiapoptotic genes, such as those of the *Bcl2* family, are upregulated by NF-κB, preventing routine cell turnover. Furthermore, expression of genes involved in cell-cycle regulation is altered (e.g., cyclins are upregulated and cell cycle inhibitors are downregulated). Ultimately, NF-κB plays a key role in inflammation-driven tumor development by generating an environment that promotes mutations and simultaneously prevents damaged cells from undergoing apoptosis, both key features of cancerous cells [25-28].

Reactive oxygen species (ROS) and nitric oxide (NO) are also generated by inflamed epithelial cells under the stress of bacterial toxin exposure or chronic bacterial infection [28,32]. These molecules play important roles in the initiation and progression of carcinogenesis by directly altering DNA, leading to mutations, deletions and chromosomal instability; if left unrepaired, these can lead to carcinogenesis [25,33]. In addition to direct effects on DNA, reactive species can influence cytoplasmic and nuclear signal transduction pathways. For example, reactive oxygen species (ROS) can direct cell proliferation and inhibit apoptosis through activation of the transcription factors MAPK (mitogen-activated

protein kinases), AP-1 (activator protein 1), and NF-κB [33]. Persistent asymptomatic bacterial infection of the colon in which the inner mucus layer is penetrated is proposed as capable of inducing chronic inflammation resulting in a cascade of diverse and complex events that compound to generate a procarcinogenic microenvironment. Bacterial breach of the mucus layer, and subsequent biofilm formation on the epithelium, is investigated in **CHAPTER 2** of this dissertation.

Oncogenic Bacterial Metabolites and Toxins

In addition to the indirect influences of chronic bacterial infection, there are also direct bacterial mechanisms of oncogenesis. Diets that are high in fat and protein provide substrates for microbial production of pro-carcinogenic products such as secondary bile acids, ammonia and polyamines. Strong evidence for the microbial production of polyamines on biofilm-covered tumor tissues is described in **CHAPTER 3**. Through their metabolically complex processes, bacteria also produce reactive species such as the derivatives of molecular oxygen including superoxide, hydrogen peroxide, and hydroxyl radicals [33,34]. These free radicals contribute to genomic instability by the mechanisms discussed above. Alternatively, several bacterial toxins have been identified that are predicted to be carcinogenic. These toxins have the capacity to modify host physiology leading either to direct DNA damage, augmentation of cellular proliferation and/or disruption of cellular differentiation and apoptosis [34]. One thoroughly studied example is CagA of *Helicobacter pylori*, which is considered to be the most important risk factor that links *H. pylori* infection to the development of gastric cancer [35,36]. CagA binds the cellular tyrosine phosphatase, SHP2, leading to modulation of cell structure [28]. It has also been shown to target multiple host proteins that regulate inflammation, and several studies suggest an ability to activate NF-κB and β-catenin signaling [37-40]. Recently, CagA was associated directly with a tumor

suppressor pathway when it was shown to usurp the tumor suppressor apoptosis-stimulating protein of p53 (ASPP2) and modify its activity thus promoting cell survival [41]. Strains of *H. pylori* expressing active VacA are associated with an increased risk of gastric cancer [42]. Another example of a bacterial toxin is the cytolethal distending toxin (CDT). This genotoxin is produced by several bacteria including selected strains of *Escherichia coli, Actinobacillus actinomycetemcomitans, Campylobacter jejuni, Shigella* spp., *Salmonella* spp., *Helicobacter hepaticus, Helicobacter cinaedi,* as well as other enterohepatic *Helicobacter spp.* [43]. CDT is composed of three subunits, one of which, CdtB, functions similar to mammalian DNase directly damaging host DNA [44]. Another CRC genotoxin family of interest, investigated in **CHAPTER 3**, can be found within the polyketide synthase (PKS) genotoxic island of *E. coli*. Recent studies have shown that this island encodes a hybrid peptide-polyketide, colibactin, capable of directly inducing DNA double strand breaks both in vitro and in vivo [45,46]. Further, deletion of the genotoxic PKS island from an *E. coli* strain diminished its oncogenic potential [47,48]. In addition to the toxins mentioned above, the microbial community contains a repertoire of toxins, listed in Table **1.1**, proposed to have oncogenic abilities. Of this group, toxins produced by *E. coli, Salmonella* spp., *Shigella* spp. or *Bacteroides fragilis* are potential contributors within the microbiota to CRC pathogenesis. One consistent feature of these bacterial carcinogenic mechanisms is that through either direct or indirect methods they interfere with key eukaryotic processes.

1.3 APC Min model of *Bacteroides fragilis*-induced CRC

One of the more compelling pieces of evidence displaying a direct link between an infectious bacterial agent in the induction of colorectal cancer lies with the murine models of enterotoxigenic *Bacteroides fragilis* (ETBF) infection. The genus *Bacteroides* is one of the most

numerically prominent members of the intestinal microbial flora. One species in particular, *Bacteroides fragilis*, is a Gram-negative obligate anaerobe and common symbiote colonizing nearly all humans. However, *B. fragilis* is also an important opportunistic pathogen, as it is the most common anaerobe isolated from clinical infections despite comprising only a small part (<1-2%) of the total microbiota [49,50]. Long recognized for roles in intestinal infections, more recently the molecular subtype, enterotoxigenic *Bacteroides fragilis* (ETBF), was revealed to induce colitis in wild-type C57Bl/6 mice and promote oncogenic transformation in multiple intestinal neoplasia (Min) mice (a murine intestinal cancer model) [51,52]. *B. fragilis* consists of two molecular subtypes termed nontoxigenic *B. fragilis* (NTBF) and ETBF. NTBF is proposed to be a probiotic organism, serving a crucial role in immune development and providing the host with usable forms of dietary products [7]. In contrast, ETBF has been identified as a cause of inflammatory diarrheal disease in animals and humans, and has also been suggested to be associated with active IBD and CRC [53-56]. Interestingly, a recent study by Zitomerskey et al found that ETBF carriage is potentially quite common in the US population, as they detected ETBF in 40% (6 out of 15) of healthy asymptomatic individuals between 31 and 66 years of age in Boston (MA,USA) [57].

To date, no ETBF strains have been fully sequenced. However, through identification and sequencing of a transposon-flanked pathogenicity island, ETBF was determined to encode a 20 kDa zinc-dependent metalloprotease termed *Bacteroides fragilis* toxin (BFT) [58,59]. BFT is the only known virulence factor of ETBF, and all strains harbor one of three highly related *bft* isoforms (*bft-1*, *bft-2*, or *bft-3*) present on the *B. fragilis* chromosome. All molecular isoforms are capable of exhibiting biological activity; however the relationship between isoform and disease severity is not yet known. BFT binds to a unknown colonic epithelial cell receptor, triggering rapid cleavage of the tumor suppressor protein E-cadherin, which, in

turn, frees its associated β-catenin, allowing its nuclear localization [60]. The subsequent

expression of the β-catenin/Wnt signaling pathway leads to an increase in colonic epithelial

cell proliferation. BFT also stimulates additional signaling pathways through NF-κB. While

the precise contribution of the plethora of colonic epithelial cell signaling triggered by BFT

to ETBF pathogenesis remains unknown, one clear biologic outcome is the recruitment and

activation of inflammatory cells, as well as epithelial cell secretion of proinflammatory

cytokines and ROS production [32,61,62]. A recent study by our group demonstrated that

ETBF induces a rapid-onset acute symptomatic colitis, followed by chronic subclinical

colonic inflammation and hyperplasia in specific pathogen-free (SPF) WT C57Bl/6 mice.

Unique to this model is the ability of ETBF to persistently colonize the mice for an extended

period of time after a single oral exposure to ETBF; in this study mice carried out to 16

months exhibited low-level colitis [52]. The acute ETBF murine colitis mimics the

inflammatory diarrhea detected in humans with ETBF infection, whereas the long-term

murine colonization is analogous to what is observed in ETBF colonization in the human

population, suggesting that ETBF carriers may be susceptible to asymptomatic ETBF-

induced colitis.

The pro-oncogenic cellular signaling induced by BFT in concert with the persistent chronic

inflammation induced by ETBF in WT mice, suggest that ETBF is an oncogenic bacterium.

This was recently tested by our group using the APC Min mouse strain, a well-established

cancer model in which loss of a single copy of the *Apc* gene predisposes mice to the

development of numerous tumors in the small intestine when the second allele

spontaneously mutates [63]. However, importantly for this animal model of bacterial-

induced carcinogenesis, adenomas are primarily observed in the small intestine and not in

the colon [64]. APC is a multidomain tumor suppressor protein that binds to and promotes

proteosomal degradation of β-catenin to regulate downstream Wnt signaling [64]. Loss or mutation of the *Apc* gene is the cause of the inherited disease familial adenomatous polyposis (FAP), and occurs in virtually all sporadic colon cancers [65]. By 4 months of age, sham APC Min mice developed an average of one to three tumors in the colon. By contrast, APC Min mice colonized with ETBF developed chronic asymptomatic colitis, with colon tumor foci detected as early as 1 week postinoculation. At 1 month of age, a marked increase in colon tumor formation (~12 tumors/mouse on average) occurs predominantly in the distal colon of ETBF-colonized mice. By contrast, Min mice colonized with NTBF did not exhibit colon tumors in excess of the sham mice. ETBF induces rapid activation of Stat3 both in the colonic epithelial cells, which are the targets of transformation in the colon, and in a subset of mucosal immune cells. Stat3 activation is required for Th17 cell development and, consistent with this, ETBF induces a rapid mucosal Th17 inflammatory response within 1 week of colonization. Colon tumors induced by ETBF also have a marked increase in Stat3 activation. Furthermore, excess tumor formation is significantly inhibited by administration of IL-17 blocking antibody, indicating that IL-17 is necessary for tumorigenesis in this model [51]. These studies suggest that persistent long-term colonization with ETBF may induce chronic colonic inflammation, with the potential for oncogenic transformation [51,52]. Furthermore, while Th17 inflammatory responses typically help the host control bacterial and fungal infection, the ETBF murine model demonstrates that endogenous Th17 responses can yield oncogenesis in the colon, a result supported by additional murine and human data [66-70].

1.4 Other animal models of bacterial influences on CRC

Animal models of bacterial driven oncogenesis have proven to be valuable tools in elucidating the link between microbes and CRC. Genetic knockout, germfree, and chemically-induced mouse models have been developed and extensively used in studies connecting bacteria and CRC. The APC Min model previously mentioned was the first mutant murine model for colon cancer and is an important tool, given the importance of inactivation of the *Apc* gene in the initiation of sporadic CRC. This initial APC Min mutant, carrying a truncation at codon 850 of the *Apc* gene, was identified among a colony of mice following random ethylnitrosourea mutagenesis [71]. Utilizing gene knockout technology, alternative *Apc* mutants have subsequently been constructed including mouse strains with truncations at codon 716 and 1638 that also develop polyps [64,72]. There are several additional genetically engineered models of intestinal neoplasia, extensively covered in a review by Taketo and Edelmann, including single knockouts of *Muc2*, *IL-10*, *Smad3*, and *Gαi2* ; and double knockouts of *APC* with *Smad4*, *TCRβ* with *p53*, *Gpx1* with *Gpx2*, *Tgfβ-1* with *Rag2* [73-81]. In the absence of intestinal microbiota, under germfree conditions, $Il10^{-/-}$, $Tcrb^{-/-}$; $Trp53^{-/-}$; $Gpx1^{-/-}$; $Gpx2^{-/-}$ and $Tgfb^{-/-}$; $Rag2^{-/-}$ mice all showed decreased or completely inhibited tumor formation [79-85]. Together, these studies indicate a role for the intestinal microbiota in development of inflammation and neoplasia. Mixed results have been reported in germfree APC Min mice. While Dove *et al.* noted a twofold decreased tumor load in the medial small intestine, they did not see a significant overall decrease in tumors [83]. By contrast, a recent study by Li *et al.* found a significant decrease of tumor load in both the small intestine and the colon. Furthermore, they identify two pathways triggered by microbiota, c-Jun/JNK and STAT3, which act to enhance tumor formation [84].

Mouse models of chemically induced colitis have also been used in studies to address the bacterial involvement in colitis and tumorigenesis. The most commonly used agents are azoxymethane (AOM) and dextran sulfate sodium (DSS). A recent study by Uronis *et al.* found that *IL10*-knockout mice that were colonized with complex microbiota and exposed to AOM developed tumors; however, germfree conditions abolished tumor formation [86]. It was further shown that conventional *IL-10/MyD88* double knockout mice showed no signs of tumor development upon treatment with AOM, suggesting that microbially-induced tumorigenesis in this system was dependent on the TLR/MyD88 pathway [86]. Johansson *et al.* showed that bacteria penetrate the inner mucus layer before inflammation is observed in the DSS colitis model, suggesting that invasion of the protective inner mucus layer and subsequent bacterial contact with the epithelium triggers the host immune response and inflammation [87]. Further studies have shown that DSS-induced colitis can be ameliorated under germ free conditions, indicating that the presence of microbiota facilitates DSS colitis [88]. A more recent study by Elinov *et al.* further supports the role of intestinal flora in the DSS-induced inflammation. Using mice deficient in the NLRP6 inflammasome, they showed that the resulting altered microbiota, characterized by increased levels of Bacteroidetes, led to an increased recruitment of inflammatory cells and worsened colitis upon DSS exposure when compared to wild-type mice [89]. This report, among others, emphasizes that the microbiota composition is shaped not only by diet, but by the host immune make-up, suggesting that human host polymorphisms modulating the inflammatory response may be important contributors to the influence of the microbiota on CRC pathogenesis [47,86,90-92]. Furthermore, several studies have shown that antibiotic treatment is capable of blocking colitis in the murine DSS colitis model [93-95]. Under germ-free conditions, DSS treatment

13

alone, however, is able to induce a slight inflammatory cell infiltration and edema, but without tumor induction [96].

While abundant data implicate the aggregate microbiome as a cofactor in colon tumor development, individual pathogens thought to promote colonic tumorigenesis have also been investigated using animal models. A study by Ellmerich and colleagues showed that *Streptococcus bovis*, long associated with colon cancer through epidemiological studies, is capable of markedly increasing the production of inflammatory cytokines and aberrant crypt foci in the colonic mucosa of rats through exposure to *S. bovis* cell wall antigens [97]. This study, however, lacked controls to demonstrate that the response was specific to *S. bovis* cell wall antigens. Furthermore, it is important to mention that the strain utilized in this study was later classified as *S. bovis* biotype II/1 (*Streptococcus Infantarius* subsp *Infantarius*) which shows a less convincing link to human CRC when compared with biotype I, a topic thoroughly covered in a recent review by Boleij and Tjalsma [98]. Another suspect, *Helicobacter hepaticus*, colonizes the liver and colon of several mouse strains and has been linked to hepatitis, chronic colitis and CRC, well-discussed in a recent review by Fox *et al.* [99] . It was recently shown that *H. hepaticus* triggers nitric oxide and TNF-α production, leading to inflammation and carcinogenesis in Rag2-deficient mice, implicating innate immune response induced by *H. hepaticus* as carcinogenic [100]. A subsequent study utilized transcriptional profiling of *H. hepaticus*-infected *Rag2*-knockout mice to reveal that colon and liver tissues exhibited different stress responses to infection. The colon was found to have a significant upregulation of genes involved in the generation of reactive species, while genes involved in DNA repair showed lower expression: this was directly contrasted with the liver, which showed upregulation of all major DNA repair pathways during infection [101]. These findings support the role of *H. hepaticus* in inflammation-induced carcinogenesis, and also

leads to interesting insights into the complexity of tissue specific microbial pathophysiology. Similarly the colon microbiota has been thought to play a role in the progression of certain diseases, such as HIV and HCV, both of which are conditions associated with an increased risk of cancer [102,103]. Other studies show that certain strains of *Enterococcus faecalis* produce extracellular superoxide and hydrogen peroxide that induces aneuploidy and tetraploidy in colonic epithelial cells [104,105]. *E. faecalis* also encodes a metalloprotease, GelE, that contributes to the development of colitis, dysplasia and adenocarcinoma in germ-free *IL10*-deficient mice [82,106]. However, to date, a link between *E. faecalis* and human CRC has not been identified [107].

In recent years, *Fusobacterium nucleatum* has become a putative candidate for contributing to CRC oncogenesis. Increased abundance of *Fusobacterium spp.* (most often *F. nucleatum*) is associated with CRC as determined through sequence analysis [23,108-110]. Recent experimental evidence, revealed the oncogenic potential of *F. nucleatum*. An IBD isolate of *F. nucleatum* induced an increase in small bowel tumors and colon adenomas in APC Min mice[109]. Similar to the findings in humans, *F. nucleatum* was enriched on tumor tissue compared to non-tumor tissue within an individual mouse. Furthermore, the authors identified a pro-inflammatory immune signature shared by *Fusobacterium*-associated APC Min tumors and human tumors with high Fusobacterium abundance, defined as >25% of sequence reads. These results were published in the same issue as the complementary data presented by Rubinstein *et al.* [111]. The authors suggest that a surface adhesion, FadA, confers virulence traits that could contribute to the carcinogenic properties of the organism. Specifically, FadA was found to facilitate the adherence to and invasion of colon carcinoma cell lines as well as tumors in xenograft mice resulting in increased cell proliferation. The authors further show that FadA binds to a specific extracellular domain of E-cadherin

triggering the β-catenin/Wnt signaling cascade and differentially promotes inflammation and oncogenesis. While the association between Fusobacterium and CRC tumors has been well established, *fad*A gene levels were also found to be increased in the colon tissues from individuals with CRC or adenomas compared to healthy subjects.

Another well-studied bacterial agent of interest is *Citrobacter rodentium*, which is known to induce self-limiting colitis, epithelial cell proliferation, and tumorigenesis in the murine colon [112]. *C. rodentium* is not a human pathogen, but is considered the mouse homolog of human attaching and effacing *E. coli* strains, which are yet another proposed procarcinogenic species. Early studies found that *C. rodentium* infection increases the carcinogenic effect of 1,2-dimethylhydrazine (DMH) treatment in NIH Swiss mice [113]. Later, a study revealed that *C. rodentium* infection leads to cytokinetic alterations and is sufficient to promote colon tumor development in APC Min mice [114]. Maddocks *et al.* reported that human attaching and effacing enteropathogenic *E. coli* (EPEC) downregulate DNA mismatch repair genes and provided preliminary data identifying these bacteria in human CRC [115]. A publication by Arthur *et al.* showed that tightly adherent *E. coli* strains harboring the PKS genotoxic island were able to induce tumor formation in AOM-treated *IL10*-deficient mice under germfree conditions [47]. The authors further showed that conventionally housed *IL10*-deficient mice developed an altered microbiome in association with colitis that occurred in 100% of the IL-10 deficient mice. Importantly, they showed that inflammation, not the carcinogen AOM, modifies the microbiota structure with emergence of potential procarcinogenic phyla. Further, a specific microbial virulence factor (the PKS island), not inflammation alone, was required for the microbially–induced carcinogenesis in this model. This study stresses the interplay between specific carcinogenic species, the microbial

community and the host. Consideration of these multifactorial influences is important when transitioning to studies concerning microbial involvement in human CRC.

1.5 Human studies

Despite a long quest, direct links between the bacterial microbiome and CRC in humans are not yet established. Culture-based, observational or case-control studies largely focusing on fecal analyses from patients with CRC and healthy control patients have suggested that *Bacteriodes*, *Streptococcus gallolyticus* subsp *gallolyticus (*previously known as *S. bovis* biotype I)*, E. coli* and *Enterococcus,* among others may be associated with development of CRC. Particularly notable over time has been the association of *Streptoccous gallolyticus* endocarditis and/or bacteremia with a high likelihood of having an underlying gastrointestinal tract malignancy, most commonly CRC [116-122]. *Clostridium septicum* aortitis and/or bacteremia have also been suggested as indicators of gastrointestinal malignancy [123]. Culture-based human studies combined with recent experimental mechanistic studies have provided the greatest support for potential roles for ETBF, *Streptococcus gallolyticus*, *Fusobacterium nucleatium,* enteroadherent *E. co*li and *E. coli* possessing the PKS island in human CRC [47,55,111,115,120].

Molecular approaches, in particular the advent of next generation sequencing techniques, have facilitated studies to examine more comprehensively the microbial associations of CRC (Table **1.2**). These approaches enhance culture-based methods because they allow detection of 'noncultivatible' microbes. Overall, available data suggest that the tumor-associated microbiome differs from that detected on matched normal tissue in the same patient.

Further the fecal microbiome of CRC patients appears to differ from that associated with their tumor and also from the fecal microbiome of healthy volunteers [124,125]. A wide range of bacteria have been reported as enriched in tumor tissue samples including *E. coli*, *Proteobacteria* (especially *Enterobacteriaceae*), *Bacteroides* spp., *Prevotella* spp., *Streptococcus* spp., *Peptostreptococcus* spp., *Enterococcus* spp. and *Fusobacterium* spp. [22,23,108,110,126]. However, the differences detected between sample groups varies among studies without clear patterns yet detected that might be useful, for example, to identify an individual at risk for or with CRC. The methodologic differences, varying sample types analyzed, varying populations studied as well as limited patient data provided makes differences among the studies difficult to interpret. Two studies from 2013, representing the largest set of CRC and matched normal tissue samples analyzed to date, identified a predominance of *Fusobacterium* spp. (*F. nucleatum* and other *Fusobacterium* spp.) associated with CRC as compared to adjacent normal tissue [23,108]. Of note, no healthy control populations were included in either study. Most studies have focused on patients with CRC; however, to begin to implicate bacteria in the pathogenesis of CRC, it is important to determine bacterial associations of colonic adenomas, precursors of CRC. Similar studies considering the unique pathogenic associations discussed earlier for IBD (see introduction) would also be helpful. In the one molecular study evaluating adenomas available to date, the bacterial population distributions also differed between adenoma and control patients when rectal biopsies of normal tissue were compared by 16S rRNA sequence analysis [17]. It is clear that additional studies are needed not only to delineate the microbial populations associated with CRC compared to diverse control populations, but also to understand how the microbial populations may relate to disease outcome and may contribute to the pathogenesis of CRC.

1.6 Thesis Aims

The overall goal of this thesis was to systematically characterize the bacterial community adherent to and directly adjacent to the colon epithelium of various pathological disease states: sporadic CRC (adenoma and carcinoma), familial CRC (adenomas from patients with familial adenomatous polyposis), as well as healthy subjects. This was accomplished through the following specific aims.

Specific Aim 1: Identify the composition and spatial arrangement of the bacterial community directly adherent to the colonic epithelium of sporadic colorectal cancer patients and healthy control subjects.

Specific Aim 2: Determine the metabolomic contribution of the bacterial biofilm community identified on select sporadic colorectal cancers.

Specific Aim 3: Characterize the composition and spatial arrangement of the bacterial community directly adherent to the colonic epithelium of familial adenomatous polyposis patients.

Figure 1.1 Overview of tissue- and cell-level mechanisms of bacterial oncogenesis. (A) In the healthy human colon, the inner mucus layer serves as a physical barrier separating the mucosal epithelium from luminal contents. The mucus layer is further protected through epithelial cell secretion of antimicrobial peptides and plasma cell secretion of IgA. This spatial segregation largely maintains the host–microbe homeostasis; nevertheless, bacterial invasion of the inner mucus layer does occur. (B) It is this perturbation that facilitates direct interactions between microbes and host cells, resulting in pathology. The precise mechanisms by which the bacterial community may induce oncogenesis when invading the inner mucus layer are, as yet, uncertain. (C & D) By contrast, for select bacteria for which preliminary epidemiologic data suggest an association with some human colorectal cancer, linkages between the mechanism of action of secreted toxins and colorectal cancer are shown. (C) Genotoxin colibactin secreted by *Escherichia coli* harboring the PKS island damages DNA. DNA damage by colibactin can be direct and/or through as yet unidentified colonic epithelial and/or other mechanisms. (D) Steps supported by experimental data regarding the action of BFT secreted by ETBF. See text for details.

DC: Dendritic cell; ETBF: Enterotoxigenic *Bacteroides fragilis*; Mf: Macrophage; MDSC: Myeloid-derived suppressor cell; NOS: Nitric oxide synthase; PMN: Polymorphonuclear cell; ROS: Reactive oxygen species.
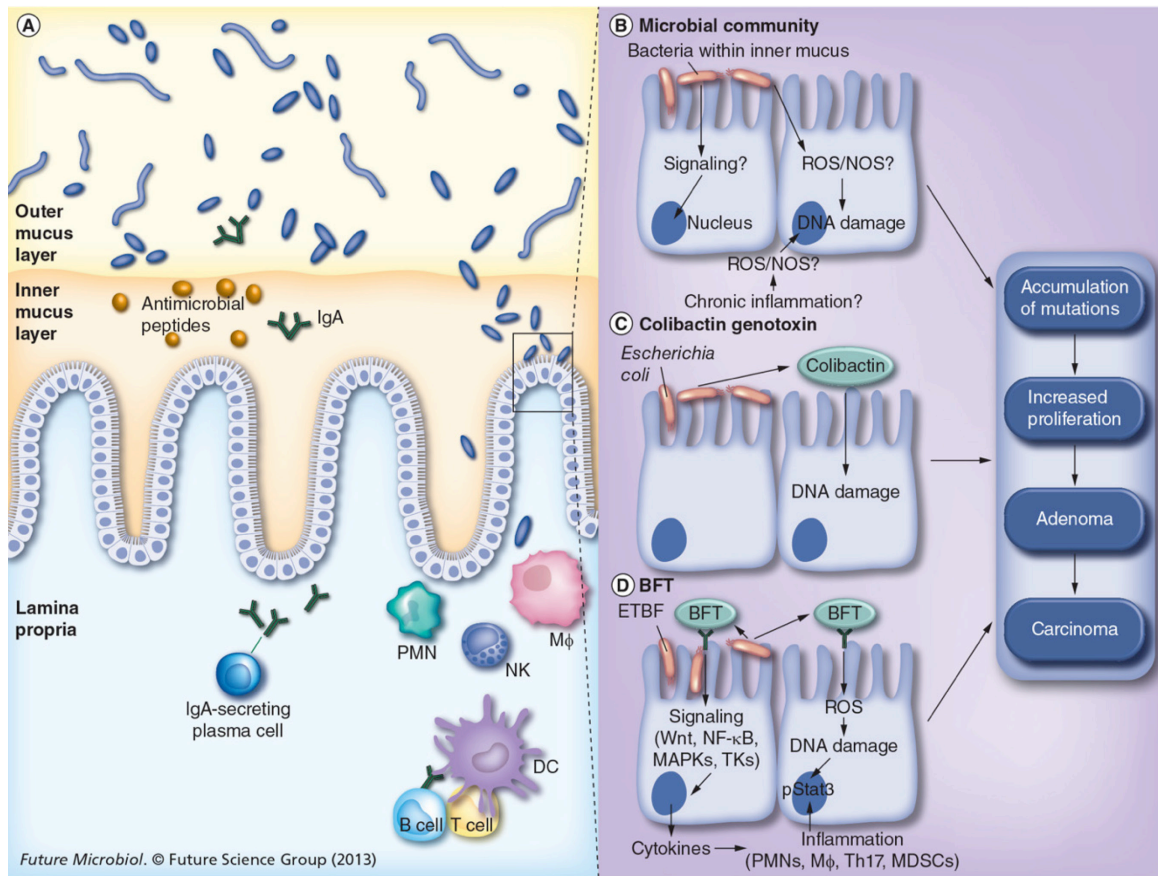
Figure 1.1

Table 1. Oncogenic bacterial toxins and linkage to human colorectal cancer				
Toxin	**Bacterial Species**	**Potential Oncogenic Mechanisms**	**Human CRC Studies†**	**Ref.**
CagA	*Helicobacter pylori*	Binds cellular SHP2, leading to cytoskeletal rearrangements; dysregulates NFkB and β-catenin signalling; hijacks the tumor suppressor ASPP2 of p53 to inhibit cell cycle progression	Meta-analysis of 16 case-control studies revealed that infection with *H. pylori* strains containing cagA increased the risk for gastric cancer 1.64 fold over *H. pylori* strains lacking cagA	
VacA	*H. pylori*	Induces the formation of vacuolar compartments, and increases proliferation through stimulation of VEGF; suppresses immune responses and inhibits apoptosis; has also been reported to induce apoptosis, an anti-tumor mechanism	Several epidemiological studies have reported that *H. pylori* strains that express active toxin have a higher association with gastric cancer than strains expressing inactive toxin (as measured *in vitro*)	
PMT	*Pasteurella multocida*	Activates Rho GTPases influencing cell signaling and proliferation	Not available	
MAP	*Citrobacter rodentium*	Targets host mitochondria, causes epithelial barrier disruption and cytoskeletal rearrangement	Not relevant. *C rodentium* is a murine pathogen	
Colibactin	*Escherichia coli*	Directly induces double DNA strand breaks *in vivo* and *in vitro*	67% and 21% of CRC (N=21) and control (N=24)patients, respectively, were positive for PKS+ *E. coli*	
CDT	*Escherichia coli, Actinobacillus actinomycetemcomitans, Helicobacter spp., Shigella spp., Haemophilus spp., Campylobacter spp,*	A subunit of CDT called Cdtb (for binding subunit) acts as a DNase directly damaging host DNA	16 and 0% of CRC (n=38) and control (n=31) patients, respectively, were positive for E. coli encoding *cdt*	
CIF	*Escherichia coli, Burkholderia pseudomallei, Yersinia pseudotuberculosis, Photorhabdus luminescens and Photorhabdus asymbiotica*	Inhibits apoptosis	8 and 0% of CRC (n=38) and control (n=31) patients, respectively, were positive for E. coli encoding *cif*.	
BFT	*Bacteroides fragilis*	Cleaves E-cadherin leading to Wnt/beta catenin nuclear signaling and increased colon epithelial cell proliferation; direct DNA damage in colonic epithelial cells, induces oncogenic Stat3/Th17 mucosal inflammation and cell signaling; induces spermine oxidase and NF-kB cell signaling	38% and 12% of CRC (N=73) and control (N+59) patients, respectively, were positive for ETBF	
FadA	*Fusobacterium nucleatum*	Cleaves E-cadherin leading to beta catenin nuclear signaling resulting in increased colon epithelial cell proliferation	*fadA* gene expression was increased (> 2 log) on adenocarcinoma tissue (n=19) compared with tissue from control patients (n=14)	

† except for H. pylori where association of select virulence determinants to gastric cancer are provided.
CRC: Colorectal cancer; ETBF: Enterotoxigenic Bacteroides fragilis

Table 1.2. Molecular analyses of microbiota in human colorectal cancer					
Population Studied	**Method**	**Sites analyzed**	**Key Findings**	**Comments**	**Ref.**
Tubular-villous adenoma (29 patients/83 biopsies[†]), Screening colonoscopy (31 patients/97 biopsies[‡]), Colonoscopy for symptoms (34 patients/113 biopsies[§]); Germany	16S rDNA PCR (5′, 600-bp product) on colon biopsies with 40-100 cloned 16S sequences analyzed/biopsied; gentamicin protection assay	not stated	Number of 16S PCR-positive patients (number with *Escherichia coli*): 1 (1) asymptomatic control, 10 (4) symptomatic controls, 27 (18) adenoma, 28 (24) carcinoma Gentamicin protection assay detected intracellular *E. coli* in biopsies of 81% of adenoma and CRC patients (n=16) and 0% of controls (n=25)	No antibiotics in prior 8 weeks; sites of biopsies and pathology of biopsies (i.e., tumor or normal mucosa) not stated; bowel preparation not stated	
Adenoma patients (21), nonadenoma (control) patients (23); participants from DHS study; USA	16S rRNA PCR analyzed by T-RFLP; 16S rRNA clone libraries (≥30 clones/patient) for four control patients; 16S rRNA FISH	Colonoscopy-obtained biopsy samples, 10-12 cm from anal verge	T-RFLP data: bacterial community distribution differed between adenoma and control patients Clone data: adenoma patients had >Proteobacter, <Bacteroidetes (phyla), >Faecalibacterium, >Dorea, <Bacteroides and Coprococcus (genera) FISH: bacteria present in mucus layer in adenoma and controls without difference noted	No antibiotics in prior 8 weeks; only rectal biopsies analyzed (normal tissue); excluded patients with prior CRC or adenoma; all received polyethelyne glycol bowel preparation	
Fecal population: CRC patients (60) and normal colonoscopy patients (119) Mucosal population: normal colon (22) and CRC (22; 16 colon, 6 rectal); France	Replicate 16S rRNA (V3-V4) pyrosequencing of 6 controls and 6 CRC fecal samples (n = 24 samples); qPCR for select genera and species[¶] on all stools and 44 mucosal DNA samples	Feces collected between 3 days and 2 weeks prior to colonoscopy; mucosal samples from surgery	PCA indicated the phylogenetic core of CRC and control stools differ qPCR revealed higher Bacteroides/Prevotella in CRC versus control patients	No prior history of cancer or colitis; specific exposure to antibiotics not stated; fecal and mucosal results not discussed separately; site of mucosal tissue harvesting in colon not stated	
6 CRC patients; The Netherlands	16S rRNA (V1-V3) pyrosequencing	Tumor and adjacent normal mucosa 5-10 cm from tumor in surgical resection specimens	On- and off-tumor bacterial populations differed significantly, although consistent patient-to-patient changes were not shown Overall, tumor tissue associated with Bacteroidetes and putative commensals, whereas Proteobacteria, especially Enterobacteriaceae, favored off-tumor tissue	No patient details provided; site of tissue acquisition in colon not stated; no healthy control population	

[†]*Nineteen with new adenoma and ten with colonoscopy within 1 year of polypectomy (six with recurrent adenoma).*

[‡]*Ten new diagnoses of CRC: 15 with colonoscopy within 1 year of partial colectomy (seven adenoma, five adenoma and CRC); three with colonoscopy 5–10 years after CRC resection.*

[§]*Normal colonoscopy.*

[¶]*Clostridium/Leptum group, Clostridium/Coccoides group, Bacteroides/Prevotella group, Escherichia coli, Bifidobacterium genus, Lactobacillus/Leuconostoc/ Pediococcus group and Faecalibacterium prausnitzii species.*

[#]*Eleven genera were reported as differing between tumor and matched normal tissues by unweighted UniFrac PCA.*

[**]*Sixteen and ten genera were reported as differing between rectal swab and stool, respectively, of CRC versus healthy volunteers by unweighted UniFrac PCA.*

[‡‡]*Distance of normal tissue from tumor not defined. CRC: Colorectal cancer; OTU: Operational taxonomic unit; PCA: Principal component analysis; qPCR: Real-time PCR; T-RFLP: Terminal restriction fragment length polymorphism.*

Population Studied	Method	Sites analyzed	Key Findings	Comments	Ref.
CRC patients (46), healthy volunteers (56); China	16S rRNA (V3) pyrosequencing; PCR to detect Bacteroides species and butyryl- CoA transferase genes	Fecal samples collected before surgery for CRC patients and in association with routine clinic visits for controls	PCA differentiated CRC from control patients. Bacteroides species, except Bacteroides fragilis and butyrate-producing bacteria, were enriched in control samples, whereas the family Enterobacteriaceae and genera Streptococcus, Peptostreptococcus, Enterococcus and Fusobacterium, among others, were enriched in CRC samples	Antibiotic exposure not defined for CRC group; no antibiotic exposure for 3 months in control group; other patient data not stated	
CRC patients (46) including stool (21), rectal swab (32), tumor tissue with matched normal tissue (2–5 and 10–20 cm from tumor) (27); healthy volunteers (56) including stool (22) and rectal swab (34); China	16S rDNA (V1–V3) bacterial pyrosequencing	Feces and rectal swab prior to bowel cleansing; tissue samples from surgical resection	Microbiota of CRC and lumen differ by PCA analysis, tumor and matched normal tissues were similar; however, in additional analyses, genus-level[#], but not phylum-level, differences were detected in tumor versus matched normal tissues. Both mucosally adherent bacteria (rectal swab comparison) and luminal microbial composition (stool) differed between CRC and healthy volunteers[††]	No antibiotics within 1 month of sampling; no patient data provided; 'within patient' versus 'between patient' analyses not clearly defined	
11 CRC and matched normal tissue samples[‡‡] for metagenomic sequencing; 88 CRC and matched normal tissue samples for qPCR; Canada	Paired-end, PCR-amplified RNA libraries sequenced on Illumina® GAIIx platform; qPCR for Fusobacterium nucleatum rRNA	CRC and matched normal colon mucosa from tumor tissue repository of surgical specimens	Only F. nucleatum subsp. nucleatum was markedly disproportionate between CRC and matched control tissues (nine of 11 subjects ≥twofold higher for F. nucleatum Illumina reads in CRC tissue and ~70% of 88 patients evaluated by qPCR)	Only paper to date to evaluate transcriptionally active bacteria; colon sites of tissue samples not defined; no patient data provided; no healthy control population	
Nine CRC and adjacent normal tissue samples for whole-genome sequencing; 95 CRC and adjacent normal tissue samples for PCR and 16S rRNA pyrosequencing; Spain; USA; Vietnam	Whole-genome Illumina sequencing; 16S rDNA (V3–V5) pyrosequencing; FISH using all bacteria and Fusobacterium-targeted probes	CRC and adjacent normal colon mucosa from surgical specimens	Overall Fusobacterium species enriched in CRC (range: <1 to >20% relative abundance in nine Illumina-sequenced samples; 16S rDNA sequencing, qPCR and FISH suggested increased Fusobacterium in CRC compared with normal colon tissues)	Colon sites of tissue samples not defined; no patient data provided; OTU with greatest similarity to F. nucleatum was the most dominant phylotype within CRC, but some tumors contained multiple Fusobacterium species; no healthy control population	

Table 1.2. Molecular analyses of microbiota in human colorectal cancer

[†]Nineteen with new adenoma and ten with colonoscopy within 1 year of polypectomy (six with recurrent adenoma).

[‡]Ten new diagnoses of CRC: 15 with colonoscopy within 1 year of partial colectomy (seven adenoma, five adenoma and CRC); three with colonoscopy 5–10 years after CRC resection

[§]Normal colonoscopy.

[¶]Clostridium/Leptum group, Clostridium/Coccoides group, Bacteroides/Prevotella group, Escherichia coli, Bifidobacterium genus, Lactobacillus/Leuconostoc/Pediococcus group and Faecalibacterium prausnitzii species.

[#]Eleven genera were reported as differing between tumor and matched normal tissues by unweighted UniFrac PCA.

[††]Sixteen and ten genera were reported as differing between rectal swab and stool, respectively, of CRC versus healthy volunteers by unweighted UniFrac PCA.

[‡‡]Distance of normal tissue from tumor not defined.

CRC: Colorectal cancer; OTU: Operational taxonomic unit; PCA: Principal component analysis; qPCR: Real-time PCR; T-RFLP: Terminal restriction fragment length polymorphism.

Chapter 2

Microbiota organization rather than composition is an underlying feature of colorectal cancers

2.1 Abstract

Sporadic colorectal cancer (CRC) results from accumulated DNA mutations in colonic epithelial cells. Environmental factors clearly affect CRC incidence but the mechanisms through which these factors function are unknown. One prime candidate is an altered colonic microbiota. Here we show that the mucosal microbiota organization is a critical factor associated with oncogenic progression in a subset of CRC. We identified invasive polymicrobial bacterial biofilms, structures previously associated with nonmalignant intestinal pathology, nearly universally (89%) on right-sided tumors (13/15 CRCs, 4/4 adenomas) but on only 12% of left-sided tumors (2/15 CRCs, 0/2 adenomas). Surprisingly, patients with biofilm-positive tumors, whether cancers or adenomas, all had biofilms on their tumor-free mucosa far distant from their tumors. Bacterial biofilms were associated with diminished colonic epithelial cell E-cadherin and enhanced epithelial cell IL-6 and Stat3 activation as well as increased crypt epithelial cell proliferation in normal colon mucosa. High throughput sequencing revealed no consistent bacterial genus associated with tumors, regardless of biofilm status. However, Unifrac distance analysis revealed that biofilm communities on paired normal mucosa, distant from the tumor itself, cluster with tumor microbiomes as opposed to biofilm negative normal mucosa bacterial communities. Colon mucosal biofilm detection may predict increased risk for development of sporadic CRC.

2.2 Introduction

When healthy, the colon is covered by a mucus layer that segregates the microbiota from direct contact with the host colonic epithelium [13]. Breaches of this protective mucus layer with resulting increased contact between mucosal microbiota and the colonic epithelial cells have been proposed as a critical first step in inciting changes in tissue biology and/or inflammation that yield inflammatory bowel disease [127,128]. Concomitant with increased

access to the mucosal epithelium, microbial community communication (such as quorum sensing) is predicted to change, thereby modifying microbial structure and function and often resulting in biofilm formation [129]. Biofilms are defined as aggregations of microbial communities encased in a polymeric matrix that adhere to either biological or non-biological surfaces. Biofilms that invade the colonic mucus layer and come into direct contact with mucosal epithelial cells indicate a pathologic state [130,131]. Biofilms characterize numerous chronic mucosal disease states in and outside of the colon (including inflammatory bowel diseases, cystic fibrosis, pharyngo-tonsillitis, otitis media, rhinosinusitis, urethritis and vaginitis) where direct bacterial contact with epithelial cells results in perturbed epithelial function and chronic inflammation [132]. However, no association of biofilms with CRC pathologic states has been reported.

2.3 Results

We systematically studied the microbial communities associated with surgically-resected colorectal tumors (CRC and adenomas) compared with paired pathologically tumor-free mucosa (here-in referred to as "normal") obtained from patients at Johns Hopkins Hospital, Baltimore, MD (JHH; Table S2.1) and the University of Malaya Medical Centre, Kuala Lumpur, Malaysia (UMMC; Table S2.2). The normal colon tissues were obtained from the margins of the resected specimens furthest from the site of the tumor (Figure S2.1). In addition, we studied colon biopsies obtained from healthy individuals without colorectal tumors and without a diagnosis of inflammatory colonic disease undergoing routine screening colonoscopy at JHH (Table S2.3).

Using samples from JHH, we first compared the spatial relationship of the microbiota with the host mucus layer and colonic epithelium using fluorescent *in situ* hybridization (FISH).

Carnoy's solution-fixed, paraffin-embedded tissues, known to preserve the mucus layer, were used whenever possible. To detect all bacterial populations, we hybridized the tissues with a probe (Eub338) targeting the conserved 16S ribosomal RNA bacterial domain [16]. Bacterial biofilms were defined as massive bacterial invasions ($>10^9$ bacteria/ml) of the mucus layer spanning at least a linear distance of 200 μm of the epithelial surface. Bacterial biofilms were identified by FISH analysis on 50% (15/30) and 67% (4/6) of all evaluated CRCs and adenomas, respectively (Figure 2.1A, Figure S2.2). Bacterial biofilm presence on tumors was ordered by geographical location along the colonic axis. Unexpectedly, when tumors were ordered by geographic location along the colonic axis, tumors in the ascending colon and hepatic flexure were biofilm-positive in 87% (13/15) and 100% (4/4) of CRCs and adenomas, respectively, whereas tumors located in the transverse and descending colon displayed biofilms in 13% (2/15) and 0% (0/2) of CRCs and adenomas, respectively ($P=0.0001$ for carcinomas and $P=0.067$ for adenomas, Fisher's exact test, Figure 2.1C). Surgical resection samples from the UMMC confirmed this geographical ordering of bacterial biofilm presence on tumors. Namely, in this Malaysian population, all four tumors harvested from the ascending colon and hepatic flexure were biofilm positive whereas 22% of tumors (4/17 CRCs and 0/1 adenoma) from the transverse and descending colon were biofilm positive (Table S2.2). Biofilm presence was not associated with age, gender, race, CRC stage, tumor size, bowel preparation or histopathologic classification.

All biofilm-covered CRCs and adenomas exhibited the remarkable feature of bacterial invasion into the tumor mass (Figure 2.2A, Figure S2.3A,B, white arrows) not detected in biofilm-negative tumors. Scanning electron microscopy (SEM) of a tumor sample subset was consistent with the FISH results, revealing both direct bacterial:epithelial surface contact and

a dense biofilm comprised of mixed bacterial morphologies on all ascending colon tumors with few mucosal bacteria detected on tumors distal to the hepatic flexure (Figure 2.1D). These data further confirm that a breach of the colonic protective mucus layer is strikingly dictated by colon geographic location.

To determine if biofilm formation was specific for the tumor microenvironment, we next used FISH to examine the paired normal colon tissues obtained from the surgical resection margin furthest from the tumor mass (Figure S2.1). No biofilms were detected on the paired normal surgically-resected colon tissues from patients with biofilm-negative tumors (adenomas or CRCs). In striking contrast, all but one normal colon tissue sample from patients with biofilm-covered tumors were biofilm positive. This was true for patients with both adenomas and carcinomas, regardless of their location within the colon. Of note, the single surgically-resected normal tissue on which we failed to detect a biofilm was fixed in formalin rather than Carnoy's and thus not optimized for mucus preservation [133]. While biofilm bacterial density did not differ between tumors (CRCs or adenomas) and their paired normal colon tissues, biofilm depth was significantly increased on tumor samples when compared to their paired normal colon tissues (P=0.001 for CRCs, P=0.028 for adenomas; Figure 2.1B). These findings demonstrate that biofilm formation represents a broad regional alteration in host epithelial:microbiota association not restricted to tumor tissue.

Screening colonoscopy biopsies from healthy individuals were typically covered with a mucus layer devoid of bacteria (Figure 2.1A, Figure S2.2). A subset of colonoscopy biopsies (15/120, 13%) revealed thin bacterial biofilms with an average density of 10^8 bacteria/ml (Figure 2.1B). Biofilm formation on colonoscopy biopsy tissues did not differ by colon location (8/60, right colon vs. 7/60, left colon) (Figure S2.1B). Thus, the right colon does

not have a greater likelihood of bacterial biofilm development in a cancer-free host. These findings in the healthy host are consistent with past reports detecting biofilms on ~15% of biopsies from asymptomatic individuals[134].

We next evaluated the composition and spatial organization of specific bacteria within the biofilms using fluorescence spectral imaging employing eleven group- and species-specific FISH probes to target the majority (85%) of the major groups of bacteria identified by sequencing. Combinations of nine probes were selected for simultaneous hybridization to the tissues (Table S2.4). Sixteen biofilm-covered tumors (three adenomas from two patients, 13 CRCs) and their paired normal mucosa were available for analysis along with normal mucosa from five right and four left colonoscopy biopsies (in three cases, biopsies were available from both the right and left colon). All biofilms, whether associated with normal colon mucosa or tumor tissue, were polymicrobial (Figure 2.2A-C). Predominant bacterial phyla associated with adenomas and CRCs were *Bacteroidetes* and *Firmicutes* (family *Lachnospiraceae* including *Clostridium*, *Ruminococcus*, and *Butyrivibrio*). A subset of tumors harbored predominant populations of *Fusobacteria* (4/16) or *Gammaproteobacteria* (1/16) (determined to be the *Enterobacteriaceae* family) (Figure S2.3). Using this multi-probe method, the tumor-invading bacterial groups present in all CRC and adenoma tissues were also identified in the biofilm bacterial composition on the tumor surface consistent with tissue invasion by a subset of these bacterial groups (Figure 2.2A, Figure S2.3A,B). Biofilms identified on surgically-resected, normal tissues were also consistently diverse, comprised of *Bacteroidetes*, *Lachnospiraceae*, and *Gammaproteobacteria* (Figure 2.2B). Biofilms detected on normal mucosa obtained at colonoscopy from patients without CRC were similarly composed of *Bacteroidetes* and *Lachnospiraceae* (Figure 2.2C). A subset of biofilm-positive

surgically-resected normal mucosa from patients with tumors (4/16, 3 tissues from CRC patients and one from an adenoma patient) also revealed bacterial invasion into the colonic epithelial cells or submucosa (Figure 2.2D and Figure S2.4). No mucosal biopsies with or without biofilms from healthy individuals revealed invasive bacteria.

To further evaluate the colonic microbiota associated with these samples, high-throughput 454 pyrosequencing targeting the hypervariable V3-V5 region of the 16S ribosomal RNA gene was performed on DNA extracted from the mucosa of 23 CRCs, two adenomas and their paired surgically-resected normal tissues and 22 biopsies obtained on colonoscopy of healthy control patients (11 right and left-matched pairs, none biofilm positive).

Sequence analysis revealed substantial overlap between tumors (adenomas or CRCs) and paired normal tissue bacterial membership at the genus level; tumor bacterial membership was a complete subset of their normal pair in 52% of tumor:normal sets. Among the 25 tumor:normal paired tissue samples, eight CRCs (32%, 4 right and 4 left), but not their paired surgically-resected normal tissues, were *Fusobacteria* dominant ($\geq$25% of total sequences) (Figure 2.2E and Figure S2.5), a finding compatible with recent reports [23,108,109,135]. No biopsies from healthy controls displayed dominant membership of *Fusobacteria* (Figure 2.2E and Figure S2.5). Collectively, our data are consistent with the concept of 'on tumor:off tumor' communities as previously reported[136].

We detected nine differentially abundant genus level groups in colonoscopy biopsies compared to paired normal biofilm-positive tissue samples from tumor patients (FDR<3%) including significant enrichment of *Lactococcus*, *Leuconostoc*, and *Comamonas* and other *Burkholderiales* members in colonoscopy biopsies, and, conversely, a 10-fold relative increase

of a candidate *Ruminococcaceae* member in surgically-resected paired normal biofilm-positive tissue samples. In contrast, we detected significant depletion of *Bacilli* and some *Bacteroidetes* members in surgically-resected normal biofilm-positive tissue samples with on average 28- and 7-fold lower relative abundance than surgically-resected normal biofilm-negative samples, respectively (FDR<5%) (data not shown).

The differences between tissues with and without a biofilm from the tumor host were highlighted by unweighted Unifrac distance analysis and principal coordinate analysis (PCoA), which revealed a striking progression of bacterial dysbiosis in biofilm positive relative to biofilm negative mucosa, despite the minimal differentially abundant taxa between the two groups (Figure 2.2E,F; Figure S2.6). These analyses revealed discrete clustering of normal colonoscopy biopsies relative to tumor-associated communities (Figure 2.2F, Figure S2.6); microbial populations from these two sample types were the most structurally divergent (P<8e-7). Communities from CRC-associated normal mucosa without biofilm were on average significantly closer in overall structure to healthy colonoscopy biopsy populations than to CRC-associated communities (P=0.001). In striking contrast, biofilm-positive normal tissue communities were significantly closer in structure to CRC-associated populations than to those of biopsies from healthy individuals (P=1e-8). This distinction supports the notion that biofilm presence correlates with the dysbiosis detected within the tumor-associated microbiota. Our findings suggest that stepwise colon mucosal microbial community dysbiosis, largely with depletion of common microbiota community members, parallels the transition from normal colon mucosa to CRC.

Can biofilms modify epithelial biology before initiation of transformation? To evaluate this conjecture, we conducted analyses to evaluate colonic epithelial cell biologic changes relevant

to carcinogenesis including barrier permeability (using E-cadherin detection as a marker) [137,138], IL-6 levels, Stat3 activation [139,140], proliferation and apoptosis in normal tissues from CRC patients as well as from healthy individuals. Loss of E-cadherin activates Wnt signaling in colon cancer and IL-6-driven Stat3 activation in colonic epithelial cells is critical to colon carcinogenesis in multiple murine models. These analyses showed marked differences between biofilm positive and biofilm negative normal colon tissues from the CRC host. Namely, biofilm positive normal tissues in the CRC host displayed significantly reduced crypt cell E-cadherin (Figure 2.3A,C) with significantly increased epithelial cell IL-6 (Figure 2.3A,D,E and Figure S2.7). Consistent with the IL-6 colonic epithelial cell localization, Stat3 activation (measured by pStat3 immunohistochemistry) was significantly increased in the epithelial cells (Figure 2.3A,F). Consistent with the action of Stat3 to promote epithelial cell proliferation and survival[139], we further detected a significant increase in crypt epithelial cell proliferation, as measured by Ki67 staining, in normal tissues covered with biofilms from CRC patients compared to normal tissues without a biofilm, also from CRC patients (P≤0.0001, Figure 2.4A). In contrast, biofilm negative normal colon tissues from the CRC host displayed intact E-cadherin (Figure 2.3A,C and Figure S2.9). Lamina propria IL-6 and Stat3 activation did not differ between biofilm positive and negative normal colon tissues, both from the CRC host (Figure 2.3A, Figures S2.8, S2.10A). Notably, although epithelial cell E-cadherin was quantitatively unchanged (Figure 2.3B, Figure S2.11A,B) in biofilm positive vs. biofilm negative colonoscopy biopsies from healthy subjects without CRC, E-cadherin localized to the basal pole of the epithelial cells in biofilm positive colonoscopy biopsies from healthy subjects without CRC (Figure 2.3B, insets). IL-6 detection was also significantly increased in biofilm positive colonoscopy biopsies from healthy subjects without CRC (Figure 2.3B, Figure S2.11C) whereas Stat3 activation was

similar in biofilm positive and negative colonoscopy biopsies from healthy subjects (Figure 2.3B, Figure S2.10B). Nonetheless, epithelial cell proliferation was significantly increased in biofilm positive colonoscopy biopsies from healthy individuals (P≤0.01, Figure 2.4B). In parallel, TUNEL staining was performed to determine if increased proliferation was simply a byproduct of increased cell turnover (Figure S2.12). Importantly, epithelial cell apoptosis was not increased in biofilm positive tissues suggesting that the increased epithelial proliferation measured is a pro-oncogenic state. This contrasts with the biofilm negative normal tissues from the CRC host where lower proliferation (Figure 2.4A) was associated with significantly increased apoptosis (Figure S2.12). Spearman correlations between Ki67 staining and genus-level relative abundances were sought in samples stratified by biofilm positive and negative status. No bacterial genera significantly correlated with the Ki67 counts in any subgroup.

2.4 Discussion

While it has been long suspected that bacteria contribute to chronic inflammation leading to CRC, this is the first time that bacterial biofilms, a known driver of tissue inflammation [128], have been identified in CRC. Further our data show biofilm formation in both the colon cancer host and healthy subjects is associated with reduced or redistributed colonic epithelial cell E-cadherin consistent with increased epithelial permeability. Our detection of enhanced IL-6 associated with biofilm formation even in healthy subjects without CRC suggests that early biofilm formation can initiate pro-carcinogenic tissue inflammation; in the cancer host with biofilm formation, IL-6 is notably localized in colonic epithelial cells with Stat3 activation. The IL-6 family of proinflammatory cytokines and their downstream effector Stat3 have been shown to promote CRC through increased epithelial proliferation, diminished apoptosis and/or angiogenesis [139,140]. Thus, our data support a model

whereby biofilm formation enhances epithelial permeability that increases direct access of bacterial antigens/mutagens to an unshielded epithelial surface and promotes pro-carcinogenic tissue inflammation. Collectively these events are predicted to induce epithelial cell mutations with consequent increased proliferation of colonic epithelial cells. In this regard, a key observation linking biofilm formation to tumor biology is our identification of the tight association between mucosal biofilm formation and the pro-cancerous state of increased epithelial proliferation. Individual genetic polymorphisms likely govern the composition of the mucosal immune response to the mucosal biofilms with Th17-dominant mucosal immune responses increasingly associated with oncogenesis and poor outcomes in CRC [70,141].

It has been hypothesized that differences in diet have significant effects on the gut microbiome and, thus, whether diet relates to biofilm status is an interesting question. Definitive diet-microbiome correlations require large epidemiologic studies that are difficult to control. We do note that the primary associations between biofilms on tumor and associated normal mucosa predominantly in right-sided colon cancers are faithfully reproduced in a population of CRC patients from Malaysia, who, in general, have a very different diet and environmental exposures than patients from the US. That biofilm patterns in colon cancer are so similar in such different populations, despite likely differences in microbiota, emphasizes the importance of microbiota structure.

The primary findings of this study are that the vast majority of right-sided CRCs are associated with a dense bacterial biofilm and that the normal colonic mucosa from patients whose tumors are covered with biofilms (whether right- or left-sided) are biofilm positive. None of the normal mucosa from patients with biofilm-negative CRCs possessed a biofilm. These findings introduce the concept that the organization, as opposed to the species

composition *per se*, of the mucosa-associated microbial community is an important factor in CRC pathogenesis, particularly in the proximal colon. Microsatellite instability, hypermethylation, hypermutation (not all correlating with the presence of microsatellite instability) and the BRAF(*V600E*) mutation have also been associated with right colon cancer[65].

Based on the numbers of patients and normal individuals with colonic biofilms, we speculate that colorectal cancers develop in two different settings: individuals with biofilms and individuals without them. Based on the data described here, the risk of developing CRC is more than 5-fold higher in the patients with biofilms compared to those without biofilms. This risk is considerably higher than that reported for other environmental associations with CRCs [142,143]. One can envision minimally invasive assays to evaluate the presence of these biofilms as well as probiotic treatments that could eliminate them. Based on these findings, prospective epidemiologic studies to directly test these hypotheses are currently being designed.

2.5 Materials and Methods

Ethics Statement

This study was approved by the Johns Hopkins Institutional Review Board and the Medical Institutional Review Board and UMMC Medical Ethics Committee at the University of Malaya. All samples were obtained in accordance with the Health Insurance Portability and Accountability Act (HIPAA).

Patient Selection and Sample Acquisition

Colon tumors (adenomas and cancers) and paired normal tissues were collected from patients undergoing surgery at Johns Hopkins Hospital or the University of Malaya Medical Centre in Kuala Lumpur, Malaysia. All tissue not needed for pathologic diagnosis was rapidly preserved in formalin, Carnoy's solution and/or RNAlater (Qiagen Inc., Germantown, MD) for analysis. Patients who received pre-operative radiation and/or chemotherapy or with a personal history of CRC were excluded. For patients in this study, two mechanical bowel preparations were routinely used and recorded (polyethylene glycol [Miralax™ or Fortrans™] or Fleet Phospho-soda™ enema [PE]). The proportion of individuals who received polyethylene glycol (PEG) vs. PE use prior to surgery was identical in the biofilm positive and negative groups. Pre-operative intravenous antibiotics were administered in all cases (cefotetan, clindamycin/gentamicin or cefoperazone/metronidazole). No patient received pre-operative oral antibiotics. Dietary information was not available.

Healthy control patients undergoing screening colonoscopy or colonoscopy for diagnostic work-up (eg, anemia) at Johns Hopkins Hospital were enrolled. All patients underwent a standard mechanical bowel preparation. Mucosal biopsies from grossly normal colon were taken from the right (cecum or ascending) and left (descending or sigmoid) colon during the colonoscopy. All tissue was rapidly preserved in formalin, Carnoy's solution and/or RNAlater for analysis. Patients who had a personal history of CRC, inflammatory bowel disease or were treated with antibiotics within the past three months were excluded.

Fluorescent in situ hybridization (FISH)

Fixed, paraffin-embedded tissues were sectioned to 5 µm thickness and de-waxed following standard procedures. Sections were stained with Periodic acid Schiff (PAS) to confirm mucus presence and preservation and successive sections were hybridized with the Eub338

universal bacterial probe and with a nonsense probe to test for nonspecific binding of probes. Slides were imaged using a Nikon E800 microscope with NIS elements software or Zeiss LSM 510 META laser scanning microscope with LSM imaging software (for confocal imaging). Paired images are presented at identical exposure intensities.

Oligonucleotide probes were synthesized and conjugated at the 5' end to the fluorophores listed in Table S2.4 (Life Technologies). Probes were applied to slides at a concentration of 2 pmol/μl of each probe in prewarmed hybridization buffer (900 mM NaCl, 20 mM Tris pH 7.5, 0.01% SDS, 20% formamide). Slides were incubated at 46°C in a humid chamber for 2 hours, and washed at 48°C for 15 minutes in wash buffer (215 mM NaCl, 20 mM Tris pH 7.5, 5 mM EDTA). Slides were dipped in water, then in 100% ethanol, air-dried, and coverslips were mounted using ProLongGold antifade reagent (Life Technologies).

Biofilm Bacterial Quantification

Biofilm bacterial density and depth were measured using slides hybridized with the universal bacterial probe, Eub338, and imaged at 1000x magnification with a Nikon E800 microscope and Nikon NIS elements viewing software.

Measures of bacterial density were based on the following model. A 10x10 μm square placed over a region of a 5 μm thick tissue section (500 μm^3) constitutes a volume of 5×10^{-10} ml. One bacterium in this volume is equivalent to 2×10^9 bacteria/ml. The visual distinction of a single bacterium is lost but spaces can still be seen between the bacteria when 250 bacteria occupy a 10x10 μm space; these cases were assigned a concentration of 10^{11} bacteria/ml. A solid mat of bacteria with no discernible spaces between the bacteria constitutes an increase to 2500 bacteria in a 10x10 μm space; these cases were assigned a concentration of 10^{12}

bacteria/ml [144]. The mean of five (10x10 μm) fields was used to determine bacterial density.

The biofilm depth was measured using ImageJ software calibrated with an image of a stage micrometer from the same microscope and magnification used in the images being quantified. Biofilm depth was calculated as the mean of five measurements taken along a 200 μm span of the biofilm.

Scanning Electron Microscopy (SEM)

Tissue samples were fixed in 2% glutaraldehyde, 2% paraformaldehyde in 0.1 M sodium cacodylate (NaCaco), 3 mM CaCl, 1% sucrose pH 7.4 overnight with gentle rocking. Samples were rinsed three times in washing buffer (0.1 M NaCaco, 3 mM CaCl, 3% sucrose), and placed in 1% osmium tetroxide in 1 M NaCaco for 1 hour in the dark. Samples were rinsed twice in distilled water followed by dehydration in an ethanol series. Samples were next placed in a 1:1 mixture of 100% ethanol to hexamethyldisilazane (HMDS) for two washes of 10 minutes each. This was followed by three washes with 100% HMDS for five minutes each. Samples were then removed and placed in a vacuum desiccant overnight followed by gold palladium coating before viewing under a Leo Zeiss Field emission SEM. Samples were scored by two independent observers for biofilm presence and morphologies.

Fluorescence Spectral Imaging and Unmixing

Samples that were determined to have a bacterial presence by universal probe were next analyzed by fluorescence spectral imaging as described above (see Fluorescent in situ hybridization Methods) using 9 probes simultaneously, targeting broad phylogenetic groups and subgroups (Table S2.4) [145-155] .

Spectral images were acquired with a Zeiss LSM 780 laser scanning confocal microscope with a 32-channel GaAsP detector and Zeiss ZEN software. All images were acquired with a Zeiss Plan-Apochromat 40x/1.4 NA(420762-9900) objective; 2x line averaging, 2048x2048 pixel frame size, 1.58 μs pixel dwell time; and 8.7 nm spectral resolution. Five fields of view were selected per sample. Spectral images of each field of view were acquired sequentially with six different lasers proceeding from long to short excitation wavelength: HeNe633 (633nm), HeNe594 (594nm), DPSS561-10 (561nm), Ar514 (514nm), Ar488 (488nm), and Diode 405-30 (405nm).

FISH probe reference spectra were measured from spectral images of pure populations of cultured bacterial cells singly labeled with the appropriate taxon-specific FISH probe. Tissue autofluorescence reference spectra were measured from spectral images of tissue subjected to the FISH procedure but without probe, and imaged under experimental imaging conditions.

Linear unmixing was performed with a custom Mathematica script using a least squares method. Each spectral image was unmixed independently using the appropriate reference spectra for the excitation wavelength. For each field of view, unmixed channels for each FISH probe were extracted from the unmixing results corresponding to the appropriate excitation wavelength. Extracted unmixed channels were compiled and colorized in ImageJ using the Image5D plugin.

Sample Preparation for Sequencing

Mucosal samples from surgically-removed tumors, paired surgical normal tissues and colonoscopy biopsies were collected in the pathology or endoscopy suites at Johns Hopkins Hospital and immediately placed in RNAlater (Qiagen Inc. Germantown, MD) and stored at

-80°C. Tissue samples (100-500 mg) were placed in a 15 ml conical tube with 2.5 ml Qiagen buffer ASL. Samples were incubated at 95°C for 15 minutes with frequent vortexing to remove bacteria from the epithelial surface. Following the dislodging of mucosal associated bacteria, 1.4 ml of supernatant was removed and cells were thoroughly lysed using a Barocycler NEP2320 (Pressure Biosciences, Inc. South Easton, MA), by cycling between atmospheric pressure, 0 psi to 25,000 psi while maintaining a temperature of 60°C. Following pressure lysis, DNA was extracted using the QIAamp DNA Stool Kit (Qiagen). Recovered genomic DNAs were quantitated using a Nanodrop spectrophotometer (Bio-Rad Life Science Research, Hercules, CA). The V3-V5 region of bacterial 16S rDNA was amplified and sequenced following the procedures described by the Human Microbiome Project standard protocol (http://www.hmpdacc.org/doc/16S_Sequencing_SOP_4.2.2.pdf). Briefly, the V3-V5 region of 16S rDNA was amplified with PCR primers (357F 5' CCTACGGGAGGCAGCAG 3' and 926R 5' CCGTCAATTCMTTTRAGT 3') that were appended with Roche 454 Titanium FLX library adapter sequences. All B-adapter primers were identical, while A-adapter primers also contained a unique barcode of 5-10 nucleotides to allow indexing of individual samples. Each sample was PCR amplified for 30 cycles with Phusion HF DNA polymerase (New England Biolabs Inc. Ipswich, MA). PCR products were purified by gel electrophoresis. Bands of the appropriate size were excised from the gel and purified using the QIAquick Gel Extraction Kit (Qiagen Inc. Germantown, MD). Purified DNAs were quantified using the 454 FLX Library Quantification Kit (KAPA Biosystems Inc. Woburn, MA) and pooled for sequencing in equal molar quantity.

Sequence data analysis

Raw sequence reads were initially assigned to samples based on multiplex identifier barcodes, trimmed of forward and reverse primer sequences, and filtered for quality and length

(minimum 150 bp) using the QIIME package (v1.6.0) [156-159] . High quality reads were then organized by sample and error-corrected using the Acacia tool (v1.52) [160] , and subsequently screened for chimeras utilizing *de novo* UCHIME (v4.2.40) [161] . Chloroplast DNA was identified and removed using the RDP Bayesian classifier (v2.5) [162] .

The final high-quality contaminant-free dataset was then submitted to the CloVR-16S pipeline (v1.1)[163]for diversity estimation, taxonomic characterization and comparative analysis of sample groups of interest. Sequences were clustered *de novo* into species-level operational taxonomic units (OTU) using UCLUST[164] with a 95% identity threshold[125,165], Taxonomic assignment of OTU representatives was performed using the RDP classifier with a minimum threshold of 0.5. There was no exclusion removal of low frequency OTUs (e.g., singletons). Pipeline runs were executed using CloVR (v2012.11.16) on the DIAG academic cloud (http://diagcomputing.org).

Immunohistochemistry and Immunofluorescence

Immunohistochemistry was performed on tissue fixed in 10% formalin and paraffin-embedded following standard procedures. Sections were de-paraffinized and rehydrated through a xylene, ethanol-water gradient. Ki67 staining was performed on an automated Benchmark ST Staining System using detection reagents from the iView DAB detection kit (Roche) with Ki67 monoclonal antibody (clone 30-9, Roche). Antibody for phospho-Stat3 (clone Tyr705 D3A7, Cell Signaling Technology) was applied at 1:400 dilution overnight following antigen retrieval. The TUNEL assay was performed per the manufacturer's instructions using the fluorescein *in situ* cell death detection kit (Roche). Immunofluorescent staining was performed following standard procedures. Antibodies for Il-6 (ab6672, Abcam), E-cadherin (Mouse IgG2a, BD), and smooth muscle antigen (SMA) (clone 1A4, mouse monoclonal, Sigma) were used at 1:400, 1:100, 1:500 dilutions respectively. Enzymatic

antigen retrieval was performed prior to IL-6 antibody application (15 minutes proteinase K treatment at 37°C); citrate-based antigen retrieval was performed prior to E-cadherin and SMA (microwave boiling in antigen unmasking solution; Vector Laboratories). All antibodies were applied overnight at 4°C, followed by secondary antibody application for one hour at room temperature. All slides were treated with DAPI for 10 minutes and mounted with glass coverslips using Prolong Gold antifade reagent (Life Technologies).

Quantification of Stat3, Proliferation and Apoptosis

Phosphorylated Stat3 (pStat3) was semi-quantitatively assessed using a 0-+3 grading scale in which epithelial cell and immune cell pStat3 are individually assessed [166].

A total of 5 well-oriented crypts were selected from each sample to be scored for Ki67+ cells by two blinded individuals. Positive cells were counted on both sides of each crypt starting at the base and ending at the luminal surface in increments of 15 cells. Each interval was scored as cells positive per 15 cells. The mean number of proliferating cells within each interval was calculated for each analyzed sample, groups were compared using the nonparametric Mann Whitney U test.

Apoptosis scoring was performed by two independent observers. TUNEL positive cells were counted per 1000 epithelial cells in 10 randomly selected fields. Results were graphed as percent positive and groups were compared using the nonparametric Mann Whitney U test.

Immunofluorescence quantification of IL-6 and E cadherin

Pixel intensity per area was measured from selected cell populations (differentiated surface epithelium, crypt epithelium, lamina propria or whole tissue) and background fluorescence was subtracted using imageJ. For differentiated surface epithelial quantification, five measurements were taken from 400x images along 100μm linear distance of surface

epithelium (for a total of 500 linear μm measured per sample). Crypt epithelial

measurements were collected from the base of five crypts. Total lamina propria fluorescence

was quantified in three distinct, representative 400x fields per specimen. Total tissue

fluorescence was measured as total fluorescence intensity upon selection of the entire tissue

specimen at 200x on at least two tissue sections per specimen. Individual specimen

fluorescence intensity values were calculated as the mean of the individual measurements

made for each specimen.

Isolation of colonic epithelial cells

Colonic epithelial cells were isolated using a modified rapid low-temperature method[167].

Briefly, approximately 500 mg of epithelial tissue was washed with ice-cold PBS and divided

into 2-3 mm fragments before transferring to chelating buffer (27 mM trisodium citrate,

5mM Na_2PO4, 8mM KH_2PO4, 1.5mM KCl, 0.5 mM DTT, 55mM D-sorbitol, 44 mM

sucrose, 6mM EDTA, 5mM EGTA, pH 7.3) for 45 minutes at 4°C. Cells were dissociated

by repeated vigorous shaking. After removal of debris, using a 100 μm cell strainer, epithelial

cells were collected by centrifugation at 150 g for 10 minutes and stored at -80°C until

protein extraction.

IL-6 ELISA

Colonic epithelial cell pellets were lysed in cell extraction buffer (Life Technologies) for 30

minutes on ice with vortexing at 10 minute intervals. Cellular debris was removed by

centrifugation at 13,000 rpm for 10 minutes. One microgram of total protein from the clear

lysate was used in the IL-6 ELISA (Sigma) according to the vendor's protocol.

Statistical analysis

Prior to downstream statistical analysis, sequence data were subsampled to equivalent depths (2500 sequences per sample) [158,168] . Unweighted UniFrac distances [169], and principal coordinate analysis plots were computed in QIIME. Additional statistical analyses were performed in R (v2.15.1) and included paired Student's t test, Fisher's exact test and the nonparametric Mann Whitney U test as appropriate. Pairwise beta-diversity comparisons utilized the nonparametric Mann Whitney U test.

Figure 2.1 Detection and quantification of bacterial biofilms on colon tumors. (A) FISH of all bacteria (red) on cancers (top panels), paired normal tissue from patients with CRC (middle panels) and colonoscopy biopsies from healthy individuals without CRC (bottom panels). All were counterstained with the nuclear stain, DAPI (blue). The top white brackets demarcate the mucus layer and the bottom white brackets denote the cytoplasm separating the nucleus (blue) of the colorectal epithelium from the mucus layer. PAS stains (Figure S2) further delineate the mucus layer on these samples. Insets depict 100X closeup showing close contact between bacteria and epithelial cells in Patient A. The pale, non-punctate red staining of the mucus layer in patients without biofilms (Patient B) represents non-specific binding to the mucus layer, which is easily demarcated from the bright red punctate staining of the bacteria infiltrating the mucus layer in patients with biofilms. Scale bar 50 µm. (B) Biofilm depth and density measurements from right CRCs/surgical normal pairs (n=15), right adenomas/surgical normal pairs (n=4), left CRCs/surgical normal pairs (n=15), left adenomas/surgical normal pairs (n=2), and right/left paired normal colonoscopy biopsies from healthy individuals without CRC (n=60). Data displayed as bar and whisker graphs where line designates the median, boxes the 25/75th percentile and whiskers the 95th percentile. (C) Geographical distribution of tumors (CRC, n=30 and adenomas, n=6) with biofilm designation. (D) SEM images. Left. Biofilm on a right colon cancer dominated by filamentous bacteria . Middle. Biofilm-negative left colon cancer where no bacteria are visualized. Right. Image of bacterial contact with host epithelium (white arrow) on biofilm-covered right colon adenoma. Mixed bacterial morphology (rods and cocci, *) is seen. All scale bars are 2 µm.
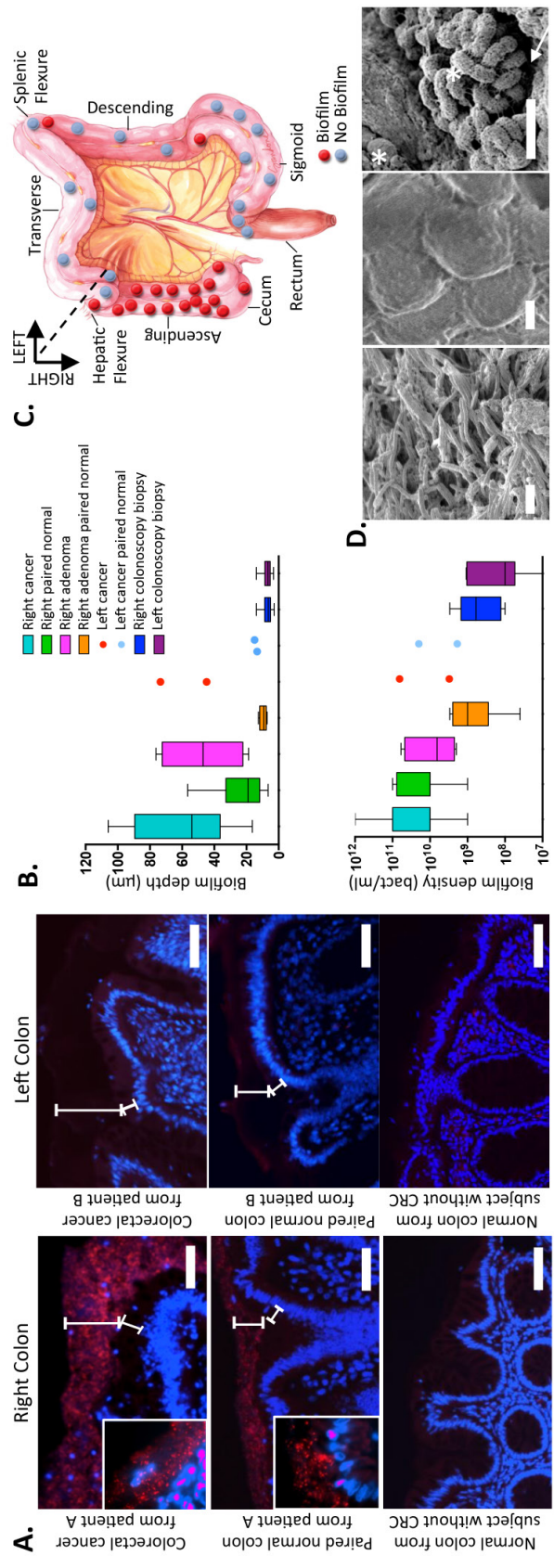
Figure 2.2. FISH and sequencing analysis of tissue reveal invasive polymicrobial biofilms and transitioning microbial populations. (**A**, **B and C**) Multiprobe spectral images of FISH-targeted bacterial groups (40x). *Bacteroidetes* (green), *Lachnospiraceae* (magenta), *Fusobacteria* (cyan), *Enterobacteriaceae* (orange), *B. fragilis* (red) are represented within the biofilms, and tissue autofluorescence is white. (**A**) Multi-group bacterial biofilm with invasion of cancer tissue (white arrows). Dotted line depicts tissue border. Right cancer with *Fusobacteria*, *Bacteroidetes*, *Lachnospiraceae*, and *Enterobacteriaceae*. Dominant group in left cancer is *Bacteroidetes*. Cancer-invading bacteria represent a subset of biofilm community members. (**B**) Bacterial biofilms on paired normal tissue comprised of *Lachnospiraceae*, *Bacteroidetes*, and *Enterobacteriaceae*. (**C**) Thin biofilms detected on right (*Bacteroidetes*, *Lachnospiraceae* and *Enterobacteriaceae*) and left (*Bacteroidetes* and *Lachnospiraceae*) normal colonoscopy biopsies (**D, left**). All bacteria FISH (red) with DAPI nuclear counterstain (blue) of paired normal tissue covered by a biofilm (20x) from a patient with CRC. White arrows mark sites of biofilm infiltration of the epithelium (20x). (**D, right**) Confocal z-stack of tissue bacterial (red) invasion (40X) denoted by white box in D, left. Disordered epithelial cells and leukocytes are visible at the infiltrated sites, while surrounding epithelial cells are intact and ordered. (Scale bars: 50 μm in A, B, C and D) (**E**) Histogram of bacterial classes represented on biofilm positive and negative samples as defined by sequence analysis. Tumor denotes 23 CRCs and 2 adenomas. (**F**) PCoA plot (based on unweighted UniFrac distances) displaying mucosa community structure of all samples (each point reflects an individual sample). Colonoscopy biopsies from healthy individuals without CRC (n=21, red) and paired normal tissues without a biofilm from patients with CRC (n=12, orange) transition to normal tissues with a biofilm from patients with CRC (n=13, green) that cluster more closely to biofilm positive

adenomas (n=2, dark blue squares) and CRCs with (n=12, dark blue) and without (n=11, light blue) biofilms.
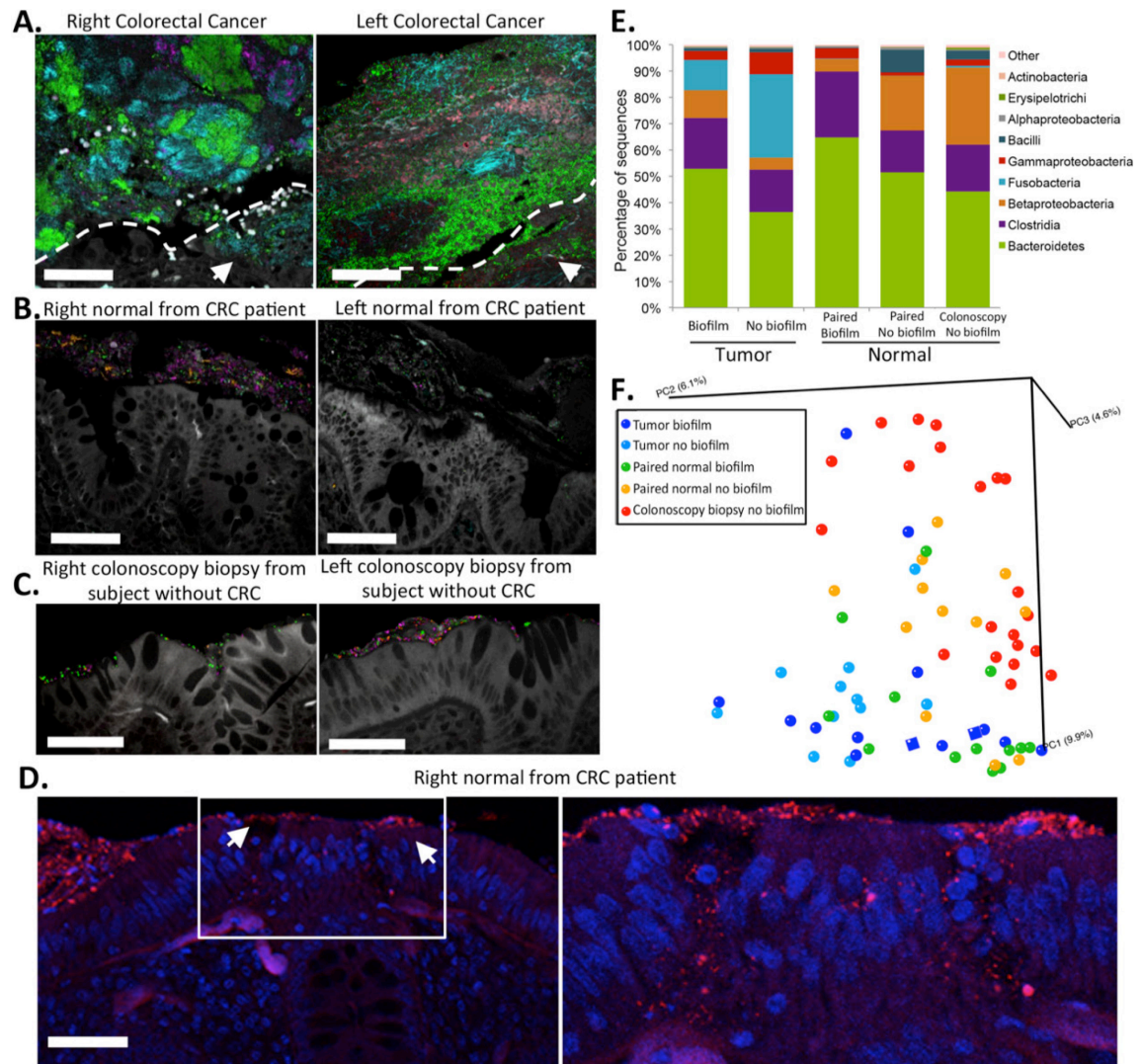
Figure 2.3 Biofilms are associated with changes in E-cadherin, IL-6 and Stat3 activation. (**A** and **B)** Evaluation of E-cadherin and IL-6 by immunofluorescence (green) and activated Stat3 (pStat3, brown nuclei) by immunohistochemistry. Blue, nuclear DAPI counterstain; red, smooth muscle antigen (SMA). Scale bars are 100 μm (E-cadherin) and 50 μm (IL-6, pStat3). (**A**). Normal colonic tissues associated with a biofilm from patients with CRC (left), obtained during surgery, display diminished crypt colonic epithelial cell E-cadherin (white arrows, n=7 biofilm positive or negative tissues) and increased epithelial cell IL-6 (white arrows, n=13 biofilm positive or negative tissues) as well as epithelial cell pStat3 (black arrows, n=16 biofilm positive and n=12 biofilm negative tissues). Normal colonic tissues without a biofilm from patients with CRC (right), likewise obtained during surgery, display intact E-cadherin. IL-6 and pStat3 are detected in the lamina propria. Quantification of crypt cell E-cadherin (fluorescence intensity), epithelial cell IL-6 (fluorescence intensity and isolated colonic epithelial cells (CEC) by ELISA) and epithelial cell pStat3 (immunohistochemistry) are shown in **C-F**, respectively. Data displayed as bar and whisker graphs where line designates the median, boxes the 25/75th percentile and whiskers the 95th percentile (C,D) or mean+/-SD (E,F). See Supplemental Methods for details. (**B**). Biofilm positive colonoscopy biopsies from subjects without CRC (left) display epithelial cell E-cadherin redistribution (inset) and increased tissue IL-6 while biofilm negative colonoscopy biopsies (right) display intact E-cadherin and modest lamina propria IL-6 expression. pStat3 is observed in the lamina propria immune cells in both biofilm positive and biofilm negative colonoscopy biopsies. See Supplemental Figures 3.10 and 3.11 for E-cadherin, IL-6 and pStat3 quantification.
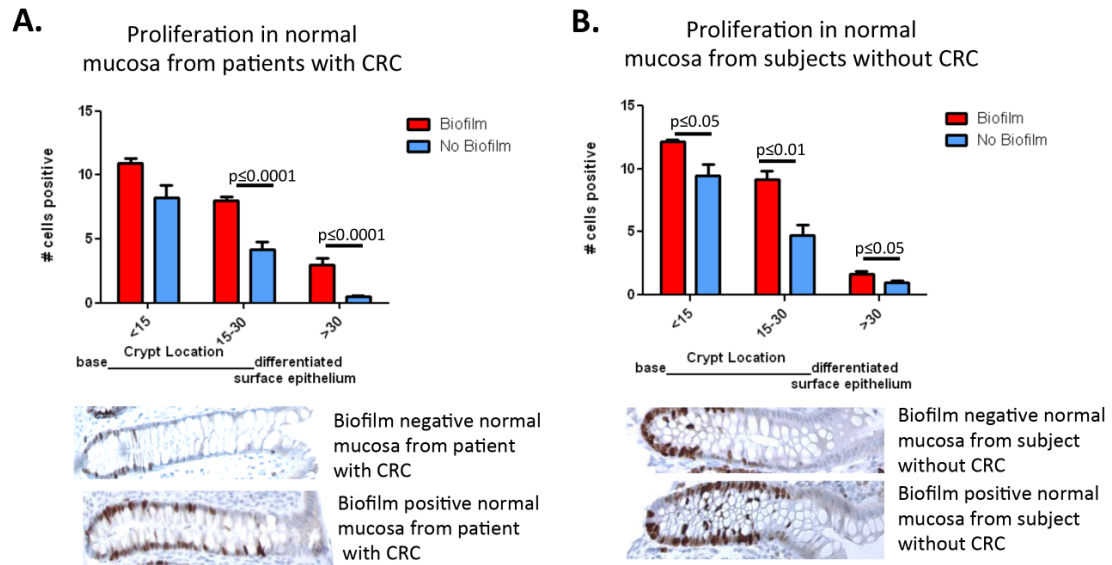
Figure 2.4 Scoring of Ki67 positive cells from the base of the crypt to the luminal surface. Normal tissues from patients with CRC obtained at surgery (**A**) with (n=17) and without (n=18) a biofilm as well as normal mucosa from healthy subjects obtained via colonoscopy (**B**) with (n=7) and without (n=10) a biofilm displayed increased proliferation in a biofilm setting. Data displayed as mean +/- SEM in groups based on distance from crypt base (<15 cells, 15-30 cells, >30 cells).
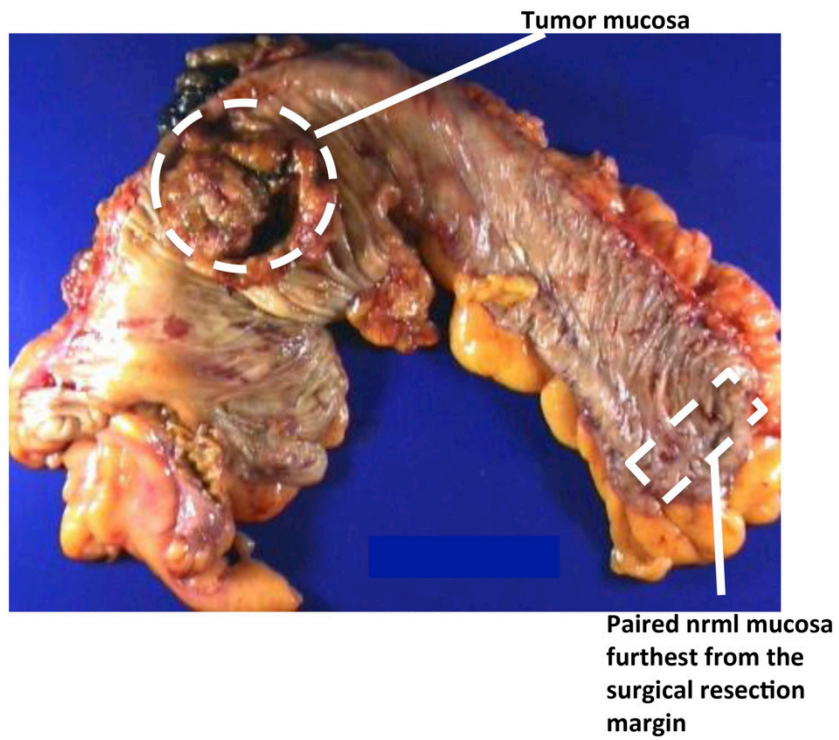
Figure S2.1. Example of mucosal tumor and normal tissue sites selected for analyses of surgically resected colons from CRC or adenoma patients.
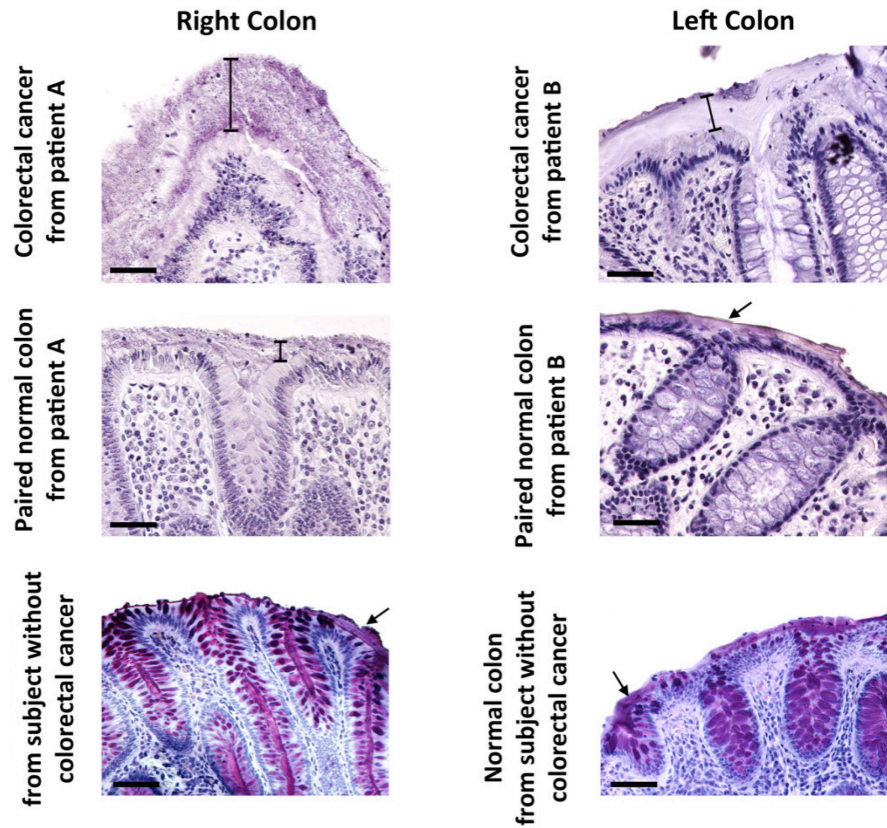
Figure S2.2. PAS-stained histopathology images of cancer and normal tissue pairs from Patient A and Patient B as well as the right and left normal colonoscopy biopsies from healthy individuals shown in Figure 1A. The mucus layer of the epithelium of each image, stained by PAS, is delineated by brackets or arrows.
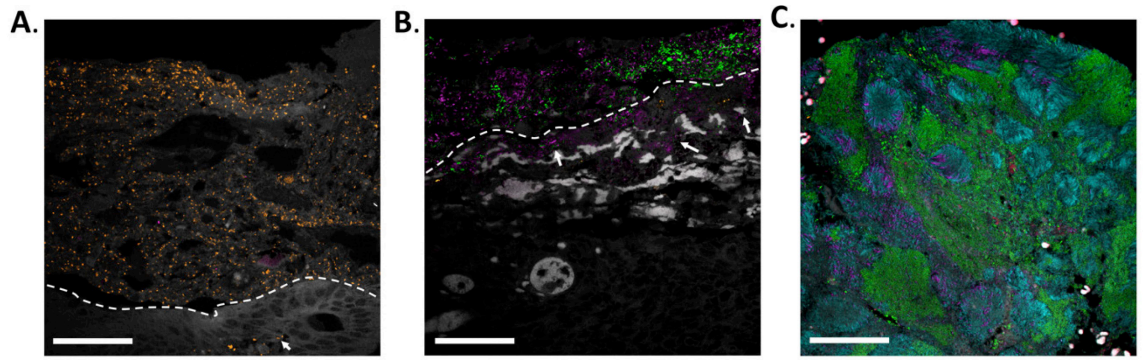
Figure S2.3. Bacterial biofilms detected on CRCs and adenomas have variable compositions. (**A**) Right adenoma biofilm comprised solely of *Enterobacteriaceae* (orange) and *Lachnospiraceae* (magenta). (**B**) Right CRC biofilm composed of *Bacteroidetes* (green) and *Lachnospiraceae* (magenta). (**C**) Right CRC biofilm composed of *Fusobacteria* (cyan), *Bacteroidetes* (green) and *Lachnospiraceae* (magenta). (Scale bar: 50 μm). Dotted white lines depict margin between bacterial biofilm and tumor tissue; white arrows identify bacteria invading into tumor tissue; S3C image is comprised only of biofilm at this magnification.
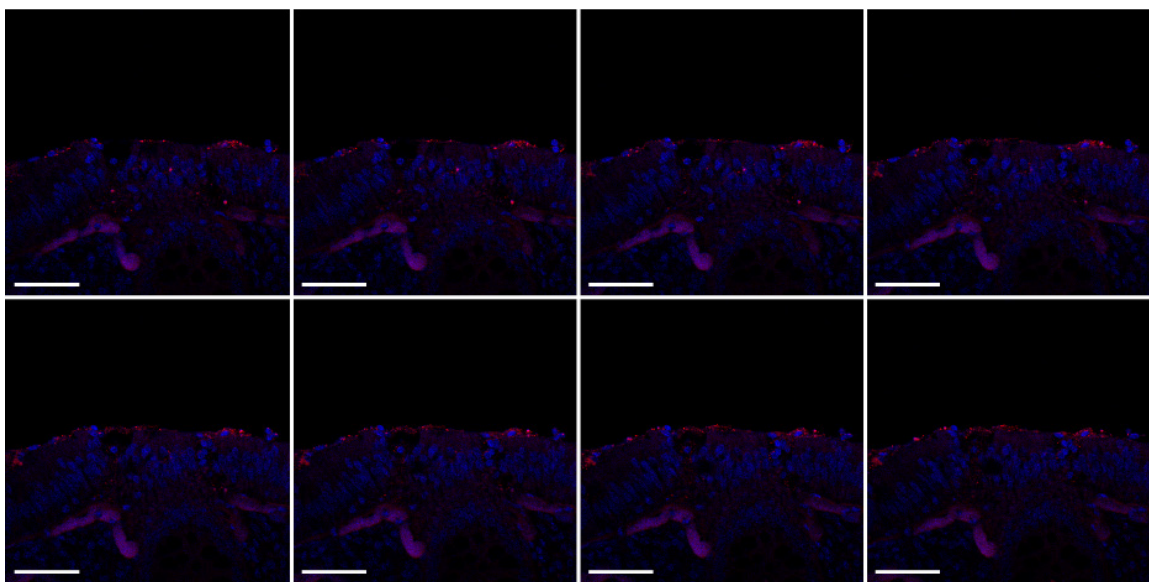
Figure S2.4. A series of eight z-stack slices (through 4 μm) depicting bacterial invasion of normal tissue (epithelial cells and submucosa) from a patient with colorectal cancer (Scale bar: 50 μm).
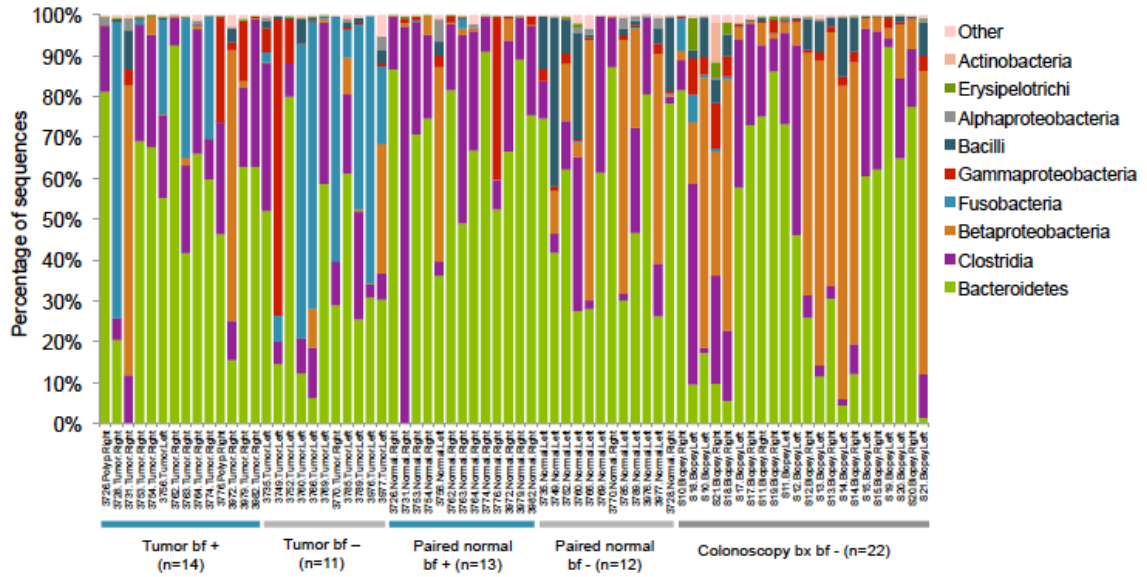
Figure S2.5. Individual subject histograms of bacterial classes grouped by tissue type and biofilm status. Tumors comprised of 23 CRCs and 2 adenomas. Paired normal tissues indicate surgically-resected normal mucosa from tumor host. Colonoscopy biopsies indicate normal mucosa biopsies from individuals undergoing screening colonoscopy.
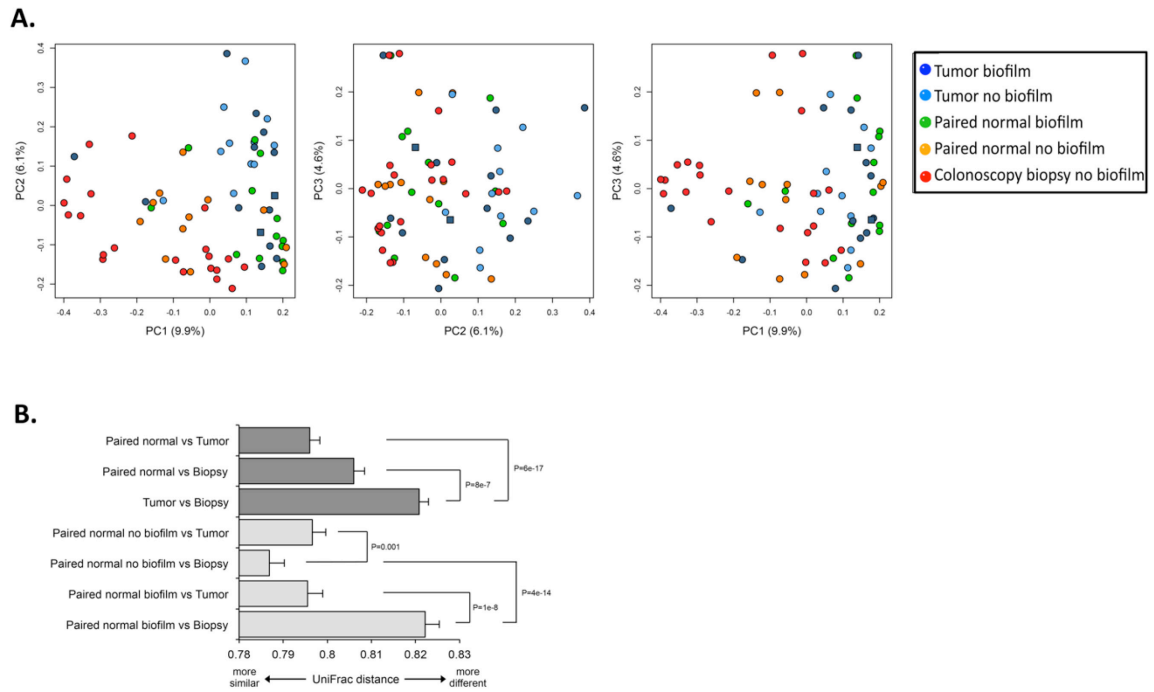
Figure S2.6. (A) Pairwise display of the first three principal coordinate axes of the PCoA.

(B) Unweighted Unifrac distance analysis. Dark-shaded bars display all tissues from tumor

hosts (surgical paired normal or tumor) whether biofilm positive (N=13) or negative (N=12)

and all colonoscopy biopsies (N=21) evaluated by sequence analysis. Light-shaded bars

display a similar analysis subsetted by biofilm status.

A. IL-6 staining **B.** IgG negative control **C.** Secondary antibody only negative control
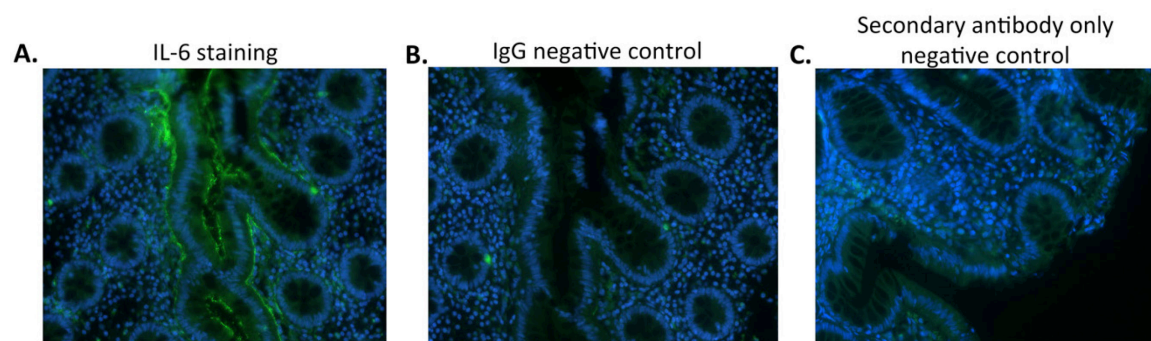
Figure S2.7. Colon mucosal tissue samples showing IL-6 immunofluorescence staining and controls. (A) IL-6 immunofluorescence staining; (B) Parallel section treated with IgG negative control antibody; (C) Sample treated only with secondary antibody

Lamina Propria IL-6
normal tissue from patients with CRC
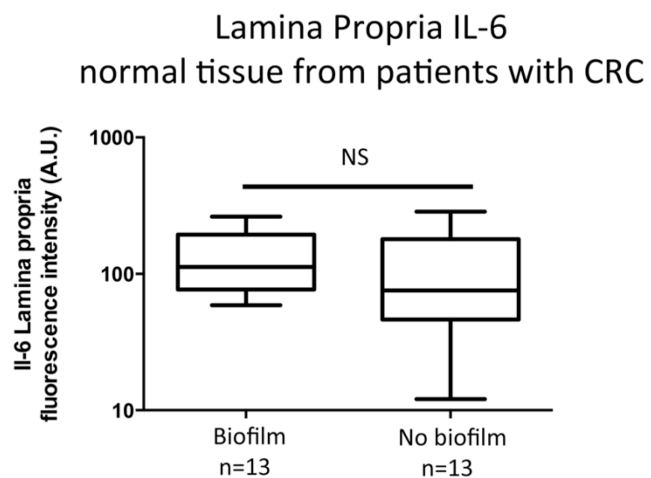
Figure S2.8. IL-6 quantification by immunofluorescence in lamina propria from biofilm positive or biofilm negative normal surgical colon tissues from patients with CRC. Data displayed as bar and whisker graphs where line designates the median, boxes the 25/75th percentile and whiskers the 95th percentile. A.U., Arbitrary Units; NS, nonsignificant.

Differentiated epithelial cell E-cadherin normal tissue from patients with CRC
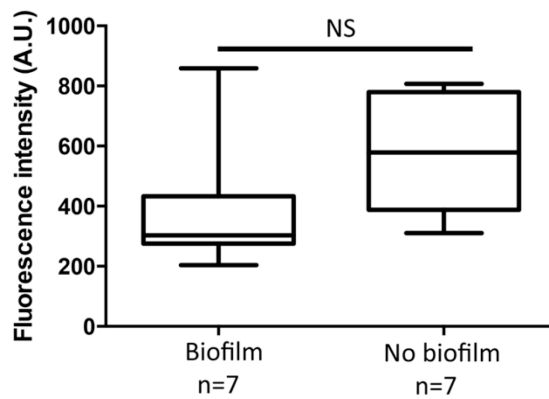
Figure S2.9. Measurement of E-cadherin in differentiated epithelial cells in biofilm positive and biofilm negative normal surgical tissues from patients with CRC. Data displayed as bar and whisker graphs where line designates the median, boxes the 25/75th percentile and whiskers the 95th percentile. A.U., Arbitrary Units; NS, nonsignificant.

A. Paired normal tissue from patients with CRC

B. Colonoscopy biopsies from subjects without CRC
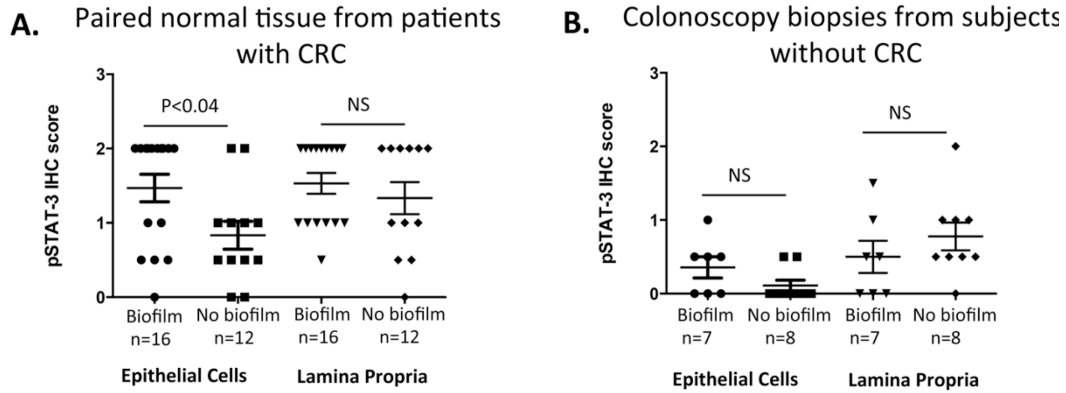
Figure S2.10. Quantification of immunohistochemistry (IHC) for activated Stat3 (pStat3) in biofilm positive or biofilm negative normal colon tissues from patients with CRC (**A**) and biofilm positive and biofilm negative colonoscopy biopsies from subjects without CRC (**B**). Epithelial cell pStat3 was significantly increased in biofilm positive normal tissues from CRC patients. Data are displayed as mean+/-SD. NS, nonsignificant
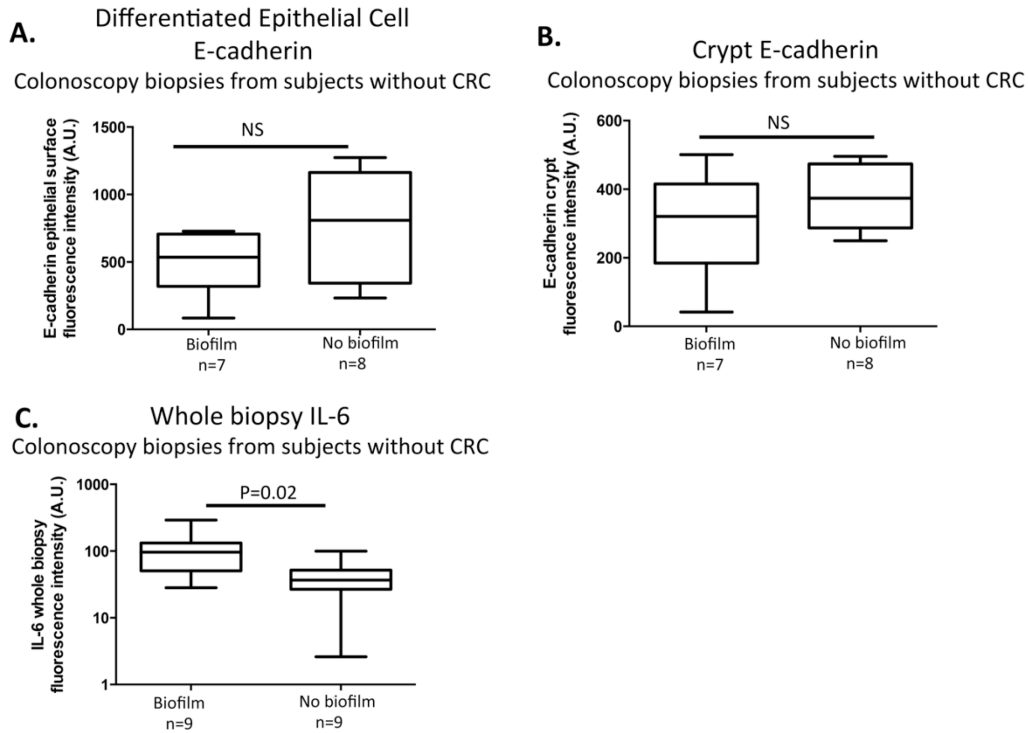
Figure S2.11. Quantification of E-cadherin or IL-6 fluorescence intensity in biofilm positive and biofilm negative colonoscopy biopsies from subjects without CRC. E-cadherin fluorescence intensity was quantified separately in differentiated surface epithelial cells (**A**) and crypt cells (**B**). Total IL-6 fluorescence intensity was quantified in each biopsy specimen (**C**). Total IL-6 fluorescence was significantly higher in biofilm positive compared to biofilm negative biopsy specimens. Data displayed as bar and whisker graphs where line designates the median, boxes the 25/75th percentile and whiskers the 95th percentile. A.U., Arbitrary Units; NS, nonsignificant.
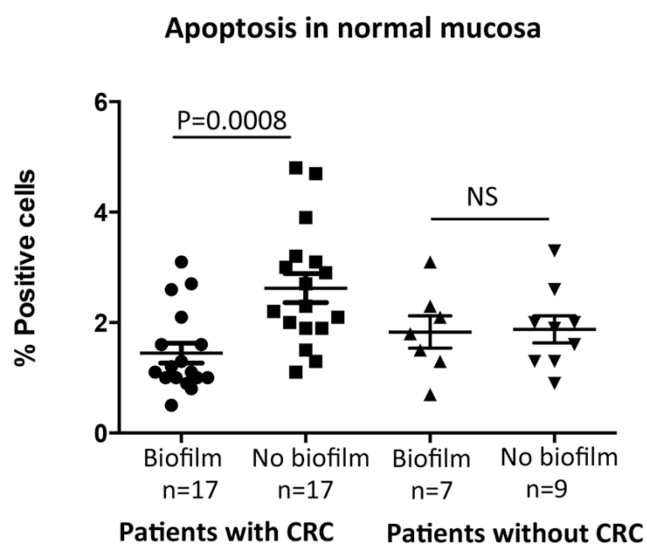
Figure S2.12. Percent of apoptotic cells scored per 1000 epithelial cells counted. Normal surgical tissue from patients with CRC with and without a biofilm, along with normal mucosa from colonoscopy biopsies from healthy individuals with and without a biofilm (subjects without CRC). NS, nonsignificant.

Supplementary Table 2.1

Surgical CRC and Polyp metadata

Patient ID	Patient Type	Age	Sex	Race	Tumor Site	Biofilm	Stage	Size (mm)	Histology
3972	Surgical CRC	78	M	Caucasian	Cecum	Yes	3	40.0	Adenocarcinoma
3979	Surgical CRC	77	F	African American	Cecum	Yes	3	35.0	Adenocarcinoma
3726	Surgical Polyp	50	M	Caucasian	Ascending	Yes	NA	40.0	Tubular adenoma-no dysplasia
3728	Surgical CRC	69	M	Caucasian	Ascending	Yes	1	87.0	Adenocarcinoma
3731	Surgical CRC	74	M	Caucasian	Ascending	Yes	2	30.0	Adenocarcinoma
3741	Surgical CRC	64	M	Caucasian	Ascending	Yes	1	8.0	Adenocarcinoma
3753	Surgical CRC	49	F	African American	Ascending	Yes	4	47.0	Mucinous Adenocarcinoma
3754	Surgical CRC	67	F	African American	Ascending	Yes	2	30.0	Adenocarcinoma
3762	Surgical CRC	73	M	Caucasian	Ascending	Yes	4	54.0	Adenocarcinoma
3763	Surgical CRC	66	F	Caucasian	Ascending	Yes	2	30.0	Adenocarcinoma
3764	Surgical CRC	59	F	Caucasian	Ascending	Yes	4	22.0	Adenocarcinoma
3776	Surgical Polyp	84	F	Caucasian	Ascending	Yes	NA	35.0	Tubular adenoma-no dysplasia
3982	Surgical CRC	62	M	Caucasian	Ascending	Yes	2	45.0	Mucinous Adenocarcinoma
3984 A/B*	Surgical Polyp/ Surgical Polyp	47	M	Caucasian	Ascending/ Ascending	Yes/Yes	NA	33.0/8.0	Tubular adenoma-no dysplasia/ Tubular adenoma-no dysplasia
3987	Surgical CRC	66	F	Caucasian	Ascending	Yes	2	50.0	Mucinous Adenocarcinoma
3986	Surgical CRC	59	M	Caucasian	Hepatic Flexure	No	2	50.0	Adenocarcinoma
3770	Surgical CRC	71	M	Caucasian	Hepatic Flexure	No	1	35.0	Adenocarcinoma
3774	Surgical CRC	45	M	Asian	Hepatic Flexure	Yes	2	45.0	Adenocarcinoma
3752	Surgical CRC	73	F	Caucasian	Transverse	No	2	25.0	Adenocarcinoma
3976	Surgical CRC	52	F	Caucasian	Transverse	No	1	20.0	Mucinous Adenocarcinoma
4017 A/B*	Surgical Polyp/ Surgical Polyp	64	F	African American	Transverse/ Rectosigmoid	No/No	NA	30.0/60.0	Tubulovillous adenoma-no dysplasia/ Tubulovillus adenoma-dysplasia
3769	Surgical CRC	78	F	African American	Splenic Flexure	No	3	60.0	Adenocarcinoma
3992	Surgical CRC	91	F	Caucasian	Splenic Flexure	Yes	2	45.0	Adenocarcinoma
3789	Surgical CRC	55	M	Hispanic	Descending	No	3	50.0	Adenocarcinoma
3988	Surgical CRC	48	M	Caucasian	Descending	No	1	35.0	Mucinous Adenocarcinoma
3749	Surgical CRC	39	M	Caucasian	Sigmoid	No	3	50.0	Adenocarcinoma
3756	Surgical CRC	54	M	Caucasian	Sigmoid	Yes	4	45.0	Adenocarcinoma
3766	Surgical CRC	56	F	Caucasian	Sigmoid	No	4	55.0	Adenocarcinoma
3977	Surgical CRC	38	F	Caucasian	Sigmoid	No	1	50.0	Adenocarcinoma
3760	Surgical CRC	29	F	Caucasian	Rectosigmoid	No	2	80.0	Adenocarcinoma
3785	Surgical CRC	54	F	Caucasian	Rectosigmoid	No	3	40.0	Adenocarcinoma
4009	Surgical CRC	53	M	Caucasian	Rectosigmoid	No	3	86.0	Adenocarcinoma
3735	Surgical CRC	64	M	Caucasian	Rectum	No	3	70.0	Adenocarcinoma
3978	Surgical CRC	90	F	Caucasian	Rectum	No	1	27.0	Mucinous Adenocarcinoma

*A/B-two adenomas were analyzed from a single individual

Supplementary Table 2.2

Surgical CRC and Polyp metadata (Malaysia)

Patient ID	Patient Type	Age	Sex	Race	Tumor Site	Biofilm	Stage	Size (mm)	Histology
S005	Surgical CRC	60	F	Malay	Cecum	Yes	3	ND	Adenocarcinoma
S016	Surgical CRC	78	M	Chinese	Cecum	Yes	3	50	Adenocarcinoma
S003	Surgical CRC	68	F	Malay	Ascending	Yes	2	70	Adenocarcinoma
S018	Surgical CRC	73	F	Chinese	Hepatic Flexure	Yes	2	70	Adenocarcinoma
S025	Surgical CRC	54	F	Malay	Transverse	No	2	60	Adenocarcinoma
S019	Surgical CRC	54	F	Other	Splenic Flexure	Yes	4	80	Adenocarcinoma
S021	Surgical CRC	57	F	Malay	Splenic Flexure	No	3	ND	Adenocarcinoma
S009	Surgical CRC	58	F	Chinese	Descending	No	3	60	Adenocarcinoma
S002	Surgical CRC	77	F	Chinese	Sigmoid	No	3	40	Adenocarcinoma
S006	Surgical CRC	79	M	Indian	Sigmoid	No	2	25	Adenocarcinoma
S010	Surgical CRC	58	F	Malay	Sigmoid	No	2	55	Adenocarcinoma
S011	Surgical CRC	67	M	Chinese	Sigmoid	Yes	2	30	Adenocarcinoma
S013	Surgical CRC	55	M	Malay	Sigmoid	No	3	65	Adenocarcinoma
S014	Surgical CRC	70	F	Chinese	Sigmoid	No	3	20	Adenocarcinoma
S015	Surgical Polyp	72	F	Indian	Sigmoid	No	NA	25	Tubulovillous Adenoma
S026	Surgical CRC	78	F	Chinese	Sigmoid	No	2	80	Adenocarcinoma
S020	Surgical CRC	76	M	Chinese	Rectosigmoid	Yes	2	100	Adenocarcinoma
S023	Surgical CRC	85	F	Chinese	Rectosigmoid	No	2	30	Adenocarcinoma
S024	Surgical CRC	77	M	Malay	Rectosigmoid	No	3	80	Adenocarcinoma
S007	Surgical CRC	71	M	Chinese	Rectum	No	2	22	Adenocarcinoma
S012	Surgical CRC	77	M	Indian	Rectum	No	2	50	Adenocarcinoma
S008	Surgical CRC	72	M	Malay	Rectum	Yes	4	42	Adenocarcinoma

Supplementary Table 2.3

Patient ID	Age	Sex	Race	Bowel Prep	Biopsy Site	Biofilm
1	51	F	African American	GoLytely Prep	Descending	Yes
					Ascending	No
2*	66	F	Caucasian	GoLytely Prep	Descending	No
					Ascending	No
3	66	M	African American	GoLytely Prep	Descending	Yes
					Ascending	Yes
4*	66	M	African American	GoLytely Prep	Descending	No
					Ascending	No
5*	44	F	Caucasian	GoLytely Prep	Descending	No
					Ascending	No
6	52	F	Caucasian	GoLytely Prep	Descending	No
					Ascending	No
7	74	M	Caucasian	Miralax Prep	Descending	No
					Ascending	No
8	51	F	African American	GoLytely & Miralax	Descending	No
					Ascending	Yes
9	73	M	African American	GoLytely Prep	Descending	No
					Ascending	No
10*	40	F	Caucasian	GoLytely Prep	Descending	No
					Ascending	No
11*	42	F	Hispanic	GoLytely Prep	Descending	No
					Ascending	No
12*	43	F	Caucasian	Miralax Prep	Descending	No
					Ascending	No
13	62	M	African American	GoLytely Prep	Descending	No
					Ascending	No
14*	75	M	Caucasian	GaviLyte Prep	Descending	No
					Ascending	No
15	44	M	Caucasian	Miralax Prep	Descending	No
					Ascending	No
17	71	F	African American	GoLytely Prep	Descending	No
					Ascending	No
18	74	F	African American	NuLytely Prep	Descending	No
					Ascending	No
19	62	M	African American	GoLytely Prep	Descending	No
					Ascending	No
20	59	M	African American	GoLytely Prep	Descending	No
					Ascending	No
21*	49	F	Asian	NuLytely Prep	Descending	No
					Ascending	No
22*	47	M	African American	GoLytely Prep	Descending	No
					Ascending	No
23*	78	F	Caucasian	Mag Citrate & Dulcolax Prep	Descending	No
					Ascending	No
24*	47	M	African American	Miralax Prep	Descending	No
					Ascending	No
25	63	M	African American	GoLytely Prep	Descending	No
					Ascending	No
26*	39	F	African American	Miralax Prep	Descending	No
					Ascending	No
27	51	F	African American	GoLytely Prep	Descending	No
					Ascending	No
28	47	F	Caucasian	Miralax Prep	Descending	No
					Ascending	No
29*	64	F	Caucasian	GoLytely Prep	Descending	No
					Ascending	No
30	51	M	Caucasian	Miralax Prep	Descending	No
					Ascending	No
31*	65	M	African American	Miralax Prep	Descending	No
					Ascending	No

*patient underwent diagnostic workup

Patient ID	Age	Sex	Race	Bowel Prep	Biopsy Site	Biofilm
32*	58	F	African American	Miralax Prep	Descending	No
					Ascending	No
33	36	M	Asian	Miralax Prep	Descending	No
					Ascending	No
34	66	F	Caucasian	Miralax Prep	Descending	No
					Ascending	No
35	60	F	Caucasian	Miralax Prep	Descending	No
					Ascending	No
36	79	F	African American	GoLytely Prep	Descending	No
					Ascending	No
37	66	M	Caucasian	Miralax Prep	Descending	No
					Ascending	Yes
38	59	M	Caucasian	Miralax Prep	Descending	Yes
					Ascending	Yes
39	72	F	African American	Miralax Prep	Descending	Yes
					Ascending	Yes
40	58	F	African American	GoLytely Prep	Descending	No
					Ascending	No
41	46	F	Asian	GoLytely Prep	Descending	Yes
					Ascending	Yes
42*	69	M	African American	Miralax Prep	Descending	Yes
					Ascending	Yes
43	62	F	African American	GoLytely Prep	Descending	No
					Ascending	No
44*	47	F	Caucasian	GoLytely Prep	Descending	No
					Ascending	No
45*	73	M	African American	GoLytely Prep	Descending	No
					Ascending	No
46	63	F	African American	GoLytely Prep	Descending	No
					Ascending	No
47*	66	M	Caucasian	Miralax Prep	Descending	No
					Ascending	No
48	56	F	African American	GaviLyte Prep	Descending	No
					Ascending	No
49	63	F	African American	GoLytely Prep	Descending	No
					Ascending	No
50	64	M	Other	GoLytely Prep	Descending	No
					Ascending	No
52*	43	F	Caucasian	NuLytely Prep	Descending	No
					Ascending	No
53	47	M	Caucasian	GoLytely Prep	Descending	No
					Ascending	No
54	55	M	Caucasian	NuLytely Prep	Descending	No
					Ascending	No
55	60	F	African American	GoLytely Prep	Descending	No
					Ascending	No
56	52	F	African American	CoLyte Prep	Descending	No
					Ascending	Yes
57*	52	M	Caucasian	TriLyte Prep	Descending	No
					Ascending	No
58	57	F	African American	Miralax Prep	Descending	No
					Ascending	No
59	77	M	African American	Nulytely & Miralax Prep	Descending	No
					Ascending	No
60*	45	F	Caucasian	TriLyte Prep	Descending	No
					Ascending	No
61	58	M	Caucasian	Miralax Prep	Descending	No
					Ascending	No
62	62	M	African American	GoLytely Prep	Descending	Yes
					Ascending	No

*patient underwent diagnostic workup

Supplementary Table 2.4

Probe target(s)	Probe Name	Fluorophore	Probe Sequence (5'-3')
Kingdom Bacteria except Planctomycetales and Verrucomicrobia	Eub338	Cy3, Alexa 405	GCTGCCTCCCGTAGGAGT
Fusobacteria	Fus714	Alexa 488	GGCTTCCCCATCGGCATT
Prevotella, Bacteroides	PRV392	Rhodamine Red X	GCACGCTACTTGGCTGG
Bacteroidetes (Bacteroides, Parabacteroides, Prevotella)	CFB286	Alexa 514	TCCTCTCAGAACCCCTAC
Betaproteobacteria	Bet42a	Alexa 647	GCCTTCCCACTTCGTTT
Gammaproteobacteria	Gam42a	Alexa 647	GCCTTCCCACATCGTTT
Lachnospiraceae	Lac435	Texas Red X	TCTTCCCTGCTGATAGA
Enterobacteriaceae except Proteus spp	Ent186	Alexa 555	CCCCCWCTTTGGTCTTGC
Bacteroides fragilis	S-S-Bfrag-998-a-A-20	Alexa 633	GTTTCCACATCATTCCACTG
Escherichia coli	Eco1531	HRP-tyramide488	CACCGTAGTGCCTCGTCATCA
Escherichia coli	Eco1161	HRP-tyramide488	GCATAAGCGTCGCTGCCG

Chapter 3

Metabolomics correlate biofilm polyamine metabolite biosynthesis with colon cancer

3.1 Abstract

The human colonic microbiome has been suggested to contribute to the eitology of colorectal cancer (CRC). Experimental evidence supports the potential for single species to act independently as oncogenic agents but also for the microbial consortia and their collective metabolites to influence the initiation or progression of CRC. Here we utilized mass spectrometry to investigate the metabolic contribution of bacterial biofilms adherent to the mucosal epithelium of colorectal cancer. We evaluated both tumor and paired normal tissues from CRC patients, with and without a biofilm, as well as colonoscopy biopsies without a biofilm from healthy subjects. An upregulation of polyamines N^1, N^{12}-diacetylspermine, N^1-acetylspermidine, N^1-acetylspermine and spermidine, previously determined to increase cellular proliferation *in vitro*, were identified in tumor tissues. Furthermore, comparison of colon cancer and paired normal tissues with or without biofilms revealed significant enhancement of N^1, N^{12}-diacetylspermine in both biofilm positive colon cancers and their matched biofilm-positive normal tissues when compared, respectively, to colon cancer and paired normal tissues without biofilms. Antibiotic treatment, which cleared biofilms, decreased N^1, N^{12}-diacetylspermine levels to those seen in biofilm negative colon cancers, suggesting that both host cancer and bacterial biofilm microenvironment contribute to the polyamine metabolite pool. These results demonstrate that colonic mucosal biofilms alter the cancer metabolome to produce a regulator of cellular proliferation and colon cancer growth potentially affecting cancer development and progression.

3.2 Introduction

Colon cancer develops over decades through the step-wise accumulation of genetic mutations [4]. Of the many environmental factors with potential to contribute to the initiation and progression of CRC, accumulating data suggests that the intestinal microbiome is playing a role[136,170]. The colon microbiota exerts the potential to mediate inflammation, DNA damage, and epithelial cell biology [7,8,171]. Recent findings by our group revealed that a subset of tumors, largely cancers from the right ascending colon, and paired normal mucosa are covered with an invasive polymicrobial biofilm that lines the epithelial surface (**CHAPTER 2**) [172]. Principle coordinate analysis of 454 sequencing reads of microbiomes revealed a progressive dysbiosis whereby biofilm-covered paired normal tissues clustered more closely to tumor populations than paired normal tissues without a biofilm. Further, these bacterial biofilms are associated with increased epithelial IL-6, pSTAT3 and crypt proliferation, a pathway well established with a pro-oncogenic role. A viscous inner gel-like mucus layer covers the epithelium of the healthy human colon [13]. Bacterial perturbation of the inner mucus layer, as identified on right-sided CRCs, constitutes a pathogenic biofilm disease-state that facilitates direct contact between the microbial consortia and host epithelial cells. This close proximity of mucosal associated bacteria likely exacerbates the metabolomic potential of these organisms. We hypothesized that the altered microbial structure of this biofilm might affect cancer biology by modulating the colon cancer metabolome, yielding metabolites that promote oncogenesis.

3.3 Results

We investigated the colon tissue metabolome of biofilm-positive and biofilm-negative CRC mucosa along with colonoscopy biopsies from healthy subjects without CRC (Patient

metadata in Tables 3.1 and 3.2). Global changes in metabolic products between cancer and

paired normal tissues were assessed using an untargeted metabolomics approach. This

analysis revealed 304 differentially regulated products, that vast majority of which were

increased in the cancer tissues compared to the paired normal (blue circles) (Fig3.1A).

Specifically, N^1, N^{12}-diacetylspermine was determined to be one of the most upregulated

products with a fold increase of 9.4. The identification of this polyamine was confirmed via

tandem mass spectroscopy (MS) using standards (Fig3.1A, B). Two additional metabolites

belonging to the polyamine class, N^1-acetylspermidine and N^1-acetylspermine, were also

upregulated in cancer tissue samples at 3.7 and 3.6 fold, respectively. Further analysis of the

mass spectra and tandem MS data revealed that a number of additional products could be

classified as phospholipids and fatty acids, in addition to the observation of multiply charged

ions coming from proteins. However, this study focuses on an in-depth analysis of

polyamines as they have been suggested to be pro-oncogenic, both in vitro and in vivo[173]

We next stratified the untargeted metabolomics approach to explore the differential

contribution of biofilms to the metabolomic profile. Samples were assessed for biofilm

presence by fluorescent in situ hybridization (FISH) using the EUB338 probe directed to the

conserved 16S ribosomal RNA bacterial domain. First, we directly compared colon cancer

tissues with and without a biofilm to understand the impact of biofilm on the cancer

metabolome (Fig 3.1C). An additional significant increase in polyamine concentration was

detected in the biofilm positive cancer tissues compared to biofilm negative cancer tissues.

Once again, N^1, N^{12}-diacetylspermine displayed the highest upregulation, 3.8-fold, while N^1-

acetylspermidine and N^1-acetylspermine were increased 1.7 and 2.0-fold. Importantly, our

sample set included two cancer specimens from the left descending colon that were biofilm

positive. These two samples had the highest concentrations of N^1, N^{12}-diacetylspermine of

all the cancer samples examined, confirming that this metabolite is not simply a product of the geographic location of the cancer in the colon but rather specifically related to the presence of a biofilm. All paired normal samples maintain the same biofilm status as their matched tumor. When biofilm positive paired normal samples were compared to their matching biofilm positive tumors N^1, N^{12}-diacetylspermine once again displayed the highest fold change of 62.2 in the cancer tissue (Fig 3.1E). N^1-acetylspermidine, N^1-acetylspermine and spermidine were also upregulated 6.5, 5.8, and 2.3-fold respectively. When the correlative comparison was made between biofilm negative tumor tissues and their respective biofilm negative paired normal tissues, a lower 7.2-fold increase was detected for N^1, N^{12}-diacetylspermine, 3-fold for N^1-acetylspermidine, 3.1-fold for N^1-acetylspermine and 1.4-fold for spermidine (Fig 3.1F). While overall increased polyamine biosynthesis is detected in cancer tissues, this analysis highlights the additional increase associated with presence of a biofilm.

To confirm the increased polyamine metabolite detection in biofilm positive cancer tissues we used LC-QqQ-MS and a quantitative targeted analysis approach to determine absolute concentrations of polyamines in the tissues. We selected the following metabolites within the upregulated polyamine metabolic pathway: spermine, spermidine, N^1-acetylspermine, N^1-acetylspermidine, N^8-acetylspermidine and N^1, N^{12}-diacetylspermine (Fig. 3.2A). This targeted approach with increased accuracy and specificity confirmed a general upregulation of the polyamine metabolites in cancer tissues compared to their paired normal colon tissues, independent of biofilm status (Fig. 3.2B, C.). These more precise measurements highlighted the fold-change significance of N^1, N^{12}-diacetylspermine that had near zero values in the untargeted analyses of normal colon tissues (Fig. 3.1B compared to Fig 3.2B). Further, using this quantitative targeted approach direct comparison of biofilm positive and

negative cancer tissues revealed only significant upregulation of N^1, N^{12}-diacetylspermine in the biofilm positive cancer tissues (Fig. 3.2D). Similar to the untargeted analysis, the two samples from the distal colon once again had the highest fold difference. This reinforces that biofilm status, rather than geographic location in the colon, dictates the detected polyamine upregulation. Using this targeted approach, the polyamines N^1-acetylspermine and N^1-acetylspermidine showed an increased trend in biofilm positive cancers but were not significantly upregulated as the initial untargeted results indicated, stressing the value of combined analyses for verification and validation.

While biofilm association with polyamine upregulation in cancer tissues suggests a potential role in cancer progression, the analysis of biofilm presence on normal tissues can begin to give insight to any pro-oncogenic potential of this phenotype before epithelial cell transformation. Importantly, the normal tissues which were concordantly biofilm negative or positive with their paired cancer tissues, revealed no significant changes in relative abundance of N^1, N^{12}-diacetylspermine and N^1-acetylspermidine by untargeted metabolomics; however targeted analysis revealed that N^1, N^{12}-diacetylspermine and N^1-acetylspermine were significantly upregulated in the biofilm positive normal colon tissues compared to biofilm negative normal tissues obtained from colon cancer hosts (Fig. 3.2E). All but one biofilm negative paired normal tissues had a value of zero for N^1, N^{12}-diacetylspermine, with an average concentration of 4.9 ± 4.3 fmol/mg tissue. In contrast, biofilm positive paired normal tissue averaged 0.6 ± 0.5 pmol/mg.

In the human host, sperimine/spermidine acetyltransferase (SSAT) is required for acetylation of spermine to generate N^1, N^{12}-diacetylspermine, while bacterial species have been verified to contain alternative acetyltransferases [174-177]. Thus, the increased acetylated polyamines could be generated by host cellular proliferation and mucosal repair leading to increased

epithelial/mucosal SSAT expression, or through bacterial acetylation. To begin to delineate the source(s) of increased acetylated polyamine production, mucosal SSAT expression was examined by immunohistochemical staining and quantification (Fig. 3.3A,B, FigS3.1). No significant difference was detected in epithelial cell SSAT between tissues with and without a biofilm (cancer tissues or paired normal tissues) (Fig. 3.3B), suggesting that the upregulation in biofilm-covered tissues is not due to increased mucosal SSAT acetylation of polyamines by the host.

We utilized nanostructure-initiator mass spectrometry (NIMS) imaging to clarify the spatial location of polyamine production on the biofilm positive tissues. This technology correlates the abundance of metabolites to their location within the tissue using a matrix-free nanostructured silicon surface which has initiator materials trapped inside nanopores, metabolites are desorbed from the tissue and placed onto the surface by laser irradiation[178,179]. Figure 3.3C shows NIMS metabolite intensity images for the polyamines in paired biofilm positive normal and cancer tissues. A consecutive tissue slice was stained for histology to verify tissue architecture, cancer, and normal cells (Fig. 3.3D). These results reveal enhanced detection of the acetylated metabolites at the mucosal edge of the cancer tissue further suggesting that the microbial biofilm could be contributing to the signal detected by MS. Overall levels of *in situ* polyamine concentrations (relative intensity in the total tissue section), consistent with the MS data, are higher in the cancer tissue than in normal tissue (Fig 3.3C).

To further delineate the microbial vs. host source of the upregulated polyamine metabolites identified, we utilized tissue samples (normal and cancer tissues, three left-sided, six right-sided pairs) collected from nine colon cancer patients treated with oral antibiotics 24 hours prior to surgery. Although CRCs collected from the right colon are predicted to be biofilm

positive based on our previous report (**CHAPTER 2,** [172]), FISH analyses of right tumor and paired normal tissues from patients treated with antibiotics revealed no biofilms (Fig.3.4A), indicating that oral antibiotics are effective at clearing the microbial biofilm population within the mucus layer. Additionally, microbial culture revealed little to no anaerobic or aerobic microbial growth on nutrient rich agar, suggesting that oral antibiotic treatment is effective at considerably lowering the cultivatable microbial load (Fig 3.4B). Considering that all right colon (cecum and ascending colon) cancers screened (n=17) were found to contain a biofilm, and the majority (88%, n=17) of left colon (transverse, distal and rectum) cancers were found to be biofilm negative, these samples offered a unique opportunity to clarify the potential microbial contribution independent of geographic location. Comparison of right and left antibiotic treated cancer samples revealed no increased metabolites; therefore, cancer tissues lacking a microbial presence no longer retain the significant increase in polyamine metabolites (Fig S3.2A). Consistent with the non-antibiotic treated samples, when all antibiotic treated cancers are compared to their paired normal tissues N^1-acetylspermine, N^1-acetylspermidine and N^1, N^{12}-diacetylspermine are increased in the cancer tissue (Fig S3.2B, Fig 3.2C). Importantly, proximal cancers from antibiotic treated patients had significantly less N^1, N^{12}-diacetylspermine than the biofilm positive cancers (Fig 3.4C). Furthermore, there are no detectable polyamine metabolite differences between antibiotic treated proximal cancers and biofilm negative cancer tissues from patients that did not receive antibiotics prior to surgery (Fig 3.4D). Collectively, these data reveal that antibiotic-treated cancers, even from the biofilm-prone right colon, are more metabolically similar to biofilm negative cancers than biofilm covered cancers. Ultimately suggesting that both host cells and microbial biofilms contribute to the global upregulation of polyamine metabolites in colon cancer.

To test the specificity of the polyamine metabolite changes to the cancer host, we examined proximal and distal colon biopsies collected from individuals undergoing routine screening colonoscopies. None of these biopsies exhibited biofilms, and these individuals did not have colon cancer or inflammatory colonic disease. No acetylated polyamines were detected using targeted metabolomics on biopsies from healthy individuals. Further, proximal and distal biopsies did not have significantly different levels of spermine or spermidine, confirming that there are no baseline differences in polyamine levels due to geographical colon location (Fig S3.2C). When the colonoscopy biopsies from healthy individuals were compared to the surgically resected paired normal tissues from CRC patients, spermine was significantly increased in the normal tissues from the cancer host (Fig 3.4E,F). Further, N^1-acetylspermidine levels were significantly increased in the biofilm-covered paired normal tissues from the cancer host compared to the biopsies from healthy individuals (Fig 3.4E). These results confirm past observations of cancer field-effect, whereby histologically normal tissue from a cancer host exhibits some cellular alterations, such as acetylated polyamines, far from the cancer site[180].

3.4 Discussion

Polyamines are essential metabolites necessary for core physiological processes, including cellular proliferation[173]. In the human colon, both host cells and the gut microbiota are capable of polyamine synthesis. Several lines of evidence have reported that bacteria have evolved mechanisms to capitalize on the presence of polyamine molecules to increase virulence and optimize their fitness within the host[181]. In vitro studies have shown that increased levels of polyamines lead to enhanced proliferation of several mammalian cell lines, while a microbial report suggests that bacterial polyamine production drives biofilm

formation[173,182,183]. This study demonstrates, for the first time, that there is a direct correlation between bacterial biofilm formation on tissues from the cancer host and upregulation of N^1, N^{12}-diacetylspermine, a polyamine that may affect the growth of both cancer cells and the associated biofilm. Although biofilms are found to be largely restricted to right-sided cancers, the two highest concentrations of N^1, N^{12}-diacetylspermine detected are from of biofilm positive left-sided cancers. This finding highlights the role of the biofilm, rather than the geographic location within the colon, in regulation of this polyamine metabolism. Association between increased polyamine levels and cancers have been reported in both animals and humans [173,184]. Increases in N^1, N^{12}-diacetylspermine, has been detected in several cancers including CRC[185,186], however, this is the first indication that the microbial metabolome is in part responsible for the production of this pro-oncogenic metabolite in a cancer setting. Notably, no significant difference in epithelial SSAT expression was detected between tissues with and without a biofilm, as measured by histological scoring. Further, NIMS analysis localizes a significant polyamine concentration at the mucosal surface where the biofilm resides. While we cannot definitively attribute the increase in N^1, N^{12}-diacetylspermine to the bacteria comprising the biofilm, the data in this report strongly suggests a role for the mucosal-associated community. We have recently demonstrated that the presence of a biofilm correlates with the pro-carcinogenic state of increased epithelial IL-6, pSTAT3 and epithelial proliferation (**CHAPTER 2**, [172]).

Collectively, we propose a model whereby host and bacterial products act together to promote biofilm formation and cellular proliferation, creating conditions conducive to oncogenic transformation in colonic epithelial cells. Consistent with this hypothesis, studies have shown that ornithine decarboxylase (ODC), the first enzyme in the pathway leading to

polyamine synthesis, is affected by microbiota in human cancer cell lines [187]. Furthermore, *Helicobacter pylori* can upregulate c-MYC leading to the activation of ODC[188]. Alternatively, the increased cellular proliferation of biofilm-covered epithelium could increase the available pool of extracellular polyamines (that bacteria have evolved specialized transporters to take up) [181,183]. The upregulation of polyamines can enhance cancer growth, invasion and metastisis [182].

A limitation of this work is that the temporal sequence of biofilm formation and cancer initiation has not been addressed. Therefore, it remains unclear whether biofilms directly induce tumorigenesis or if they are merely a consequence. To begin to approach this difficult question, a prospective study analyzing biofilm presence in healthy individuals undergoing routine colonoscopy would be required. These samples would give us the capacity to identify mechanisms that biofilms may exert in the absence of a cancer environment. Unfortunately, no biofilm-covered biopsies from healthy individuals were assessed in this analysis. Nevertheless, the data presented here adds to the growing body of evidence supporting different etiologies of disease between right and left cancers, which differ in their molecular and metabolic characteristics[65,189]. Furthermore, treatment implications may be extracted from this data. Both animal models and clinical trials utilizing inhibitors that block the polyamine metabolic pathway have resulted in unclear findings. However, focused targeting taking into consideration both polyamine production and biofilm interactions may prove a more successful strategy.

3.5 Materials and Methods

Materials

Spermine, spermidine, N^1-acetylspermine and N^8-acetylspermidine were purchased from Sigma Aldrich (St. Louis, MO). N^1-acetylspermidine and N^1, N^{12}-diacetylspermine dihydrochloride were kind gifts from Frank J. Gonzalez, National Cancer Institute, Bethesda, MD and Masao Kawakita, Tokyo Metropolitan Institute of Medical Science, Toyko, Japan, respectively. The anti-SSAT antibody was provided by Robert A Casero, Johns Hopkins Medical Institutions, Baltimore, MD. All other chemicals were of the highest chemical grade and purchased from Sigma Aldrich.

Sample Collection

Colon cancers and paired histologically normal tissues were collected from patients undergoing surgery at the Johns Hopkins Hospital. Tumor and normal tissues not required for pathologic diagnosis were preserved in Carnoy's fixative or snap frozen in liquid nitrogen for analysis. Patients who received pre-operative radiation, chemotherapy or had a personal history of CRC were excluded. Pre-operative intravenous antibiotics were administered in all cases (cefotetan or clindamycin/gentamycin). A subset of patients received oral antibiotics (neomycin and erythromycin) the day prior to surgery.

Healthy control patients undergoing screening colonoscopy were recruited and signed informed consent. All patients underwent a standard mechanical bowel preparation. Mucosal biopsies from grossly normal colon were taken from the ascending (right) and descending (left) colon during the colonoscopy. Mucosal biopsies were rapidly preserved in Carnoy's fixative or snap frozen in liquid nitrogen for analysis. Patients who had a personal history of

CRC, inflammatory bowel disease or were treated with antibiotics within the past three months were excluded. Patient metadata are in Tables 1 and 2.

This study was approved by the Johns Hopkins Institutional Review Board.

Fluorescent in situ hybridization

The universal bacterial probe, EUB338 (5'GCTGCCTCCCGTAGGAGT3'), and nonsense probe NON338 (5'ACTCCTACGGGAGGCAGC), were synthesized and conjugated at the 5' end to Cy3 (Eub338) or Alexa488 (NON338) (Invitrogen Life Technologies). Universal probe was applied to 5 μm thick Carnoy's-fixed paraffin-embedded tissue sections. The nonsense probe, NON338, was also applied to test for nonspecific binding. Successive sections were stained with Periodic acid Schiff (PAS) to confirm mucus presence and preservation.

Slides were de-waxed following standard procedures and subjected to 10 minutes of 10 mg/ml lysozyme in Tris buffer, followed by three rinses in Tris buffer. Eub338 oligonucliotide probe was applied to slides at a concentration of 2 pmol/μl in prewarmed hybridization buffer (900 mM NaCl, 20 mM Tris pH 7.5, 0.01% SDS, 20% formamide). Slides were incubated at 46°C in a humid chamber for 2 hours, and washed at 48°C for 15 minutes in wash buffer (215 mM NaCl, 20 mM Tris pH 7.5, 5 mM EDTA). Coverslips were mounted using ProLong Gold antifade reagent (Life Technologies) and slides were imaged using a Nikon E800 and imaged NIS elements solftware.

Sample Preparation for Metabolomics

For each sample, 10 mg of tissue was weighed and added to 1.5 mL centrifuge tubes containing 600μl ice cold acetone. Samples were vortexed for 30 seconds, snap frozen in liquid nitrogen for 1 minute, thawed for 3 minutes and sonicated for 15 minutes at 50°C.

Freeze-thaw cycles were repeated two more times before storing samples at -20°C for 1 hour. Samples were centrifuged at 13,000 rpm for 15 minutes and the supernatant transferred to a new 1.5 mL centrifuge tube for storage at -20°C. The pellet was resuspended in 400 μL ice cold methanol/water/formic acid (86.5/12.5/1.0 v/v/v), vortexed for 30 seconds and sonicated for 15 minutes at 50°C. The pellet samples were stored at -20°C for 1 hour followed by centrifugation (13,000 rpm for 15 minutes). The supernatant was pooled with the supernatant collected earlier and dried down in a Speedvac for 4 hours. The samples were resuspended in 100 μl acetonitrile/water/isopropanol (50/40/10 v/v), sonicated for 5 minutes at 50°C and stored at 4°C for 1 hour. The samples were finally centrifuged at 13,000 rpm for 15 minutes and the supernatants transferred to glass HPLC vials for LC-MS analysis. These methods recover both hydrophobic and hydrophilic metabolites from the samples.

Untargeted Metabolomics

Samples were randomized and analyzed by high performance liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry (HPLC-ESI-QTOFMS). Samples (8 μL) were injected onto a reversed-phase 150 x 1.0 mm Zorbax 5μm C18 column (Agilent Technologies, Santa Clara, CA) using an Agilent Technologies series 1200 HPLC with a gradient mobile phase of 0.1% formic acid (solution A) and acetonitrile containing 0.1% formic acid (solution B) at a flow rate of 20 μL/minute: 2% B for 5 minutes to 98% B at 50 minutes, held for 10 minutes at 98% B then re-equilibration at 2% B. MS was performed on an Agilent Technologies 6538 UHD Accurate Mass Q-TOF. The samples were analyzed in ESI positive mode. LC/MS data were processed using XCMS Online. XCMS applies a nonlinear retention time correction, performs peak-picking, feature identification and matches peaks across runs[190]. It reports integrated areas of each

detected peak in individual samples and calculates the Welch's t test for two sample groups. For this study paired and unpaired non-parametric tests were carried out (Wilcoxon Rank Sum and Mann-Whitney). Features were listed in a feature table and as an interactive cloud plot, containing their integrated intensities (extracted ion chromatographic peak areas) observed fold-changes across the two sample groups, and *p*-values for each sample[191]. The default XCMS parameter set for HPLC-UHD-QTOFMS was used with tolerance for database search set to 30 ppm. Integration of METLIN to XCMS Online allowed for putative identification of metabolites. Identifications were then made by comparing retention time and tandem MS fragmentation patterns to the sample and a standard compound. **Tandem MS experiments** were carried out with the collision energy set to 20 eV and caused the fragmentation of the metabolites into a number of fragments specific for the metabolite. This fragmentation pattern combined with the retention time comparison to a standard allows for accurate identification. The full dataset is available as a public share on XCMS Online.

Targeted Metabolomics of Polyamines

For method development a number of column chemistries were tried and tested in acidic conditions to retain and resolve the polyamines including the Zorbax C18 (Agilent Technologies), Aminopropyl (Phenomonex, Torrance, CA) and ZIC-HILIC (SeQuant, Umea, Sweden); however the Scherzo SM-C18 column 150 x 0.5 mm 3μm (Imtakt, Philadelphia, PA) gave the optimal results. Samples (8 μL) were injected onto the column using an Agilent Technologies series 1200 HPLC with a gradient mobile phase of 5 mM ammonium acetate (solution A) and 50 mM acetate and acetonitrile (50/50 v/v) (solution B) at a flow rate of 20 μL/minute: 2% B for 5 minutes, to 17% B at 11 minutes, to 98% B at 13.5 minutes, held for 5 minutes at 98% B then re-equilibration at 2% B. Targeted analysis

for quantitation of the polyamines were measured by using the above column conditions and selected reaction monitoring triple quadrupole mass spectrometry (Agilent 6410 QqQ-MS). The following quantifier and qualifier transitions were used for each metabolite: spermine: 203.2 -> 112.1, 203.2 -> 129.1; spermidine: 146.2 -> 112.1, 146.2 -> 72.1; N^1-acetylspermine: 245.2 -> 100.1, 245.2 -> 112.1; N^1-acetylspermidine: 188.2 -> 100.1, 188.2 -> 72.1; N^8-acetylspermidine: 188.2 -> 114.1, 188.2 -> 72.1; N^1, N^{12}-diacetylspermine: 287.2 -> 100.1, 287.2 -> 171.1. The fragmentor voltage and collision energies were as follows: spermine: 110 V, 6 V (quantifier), 18 V (qualifier); spermidine: 107 V, 10 V (quantifier), 14 V (qualifier); N^1-acetylspermine: 119 V, 18 V (quantifier), 18 V (qualifier); N^1-acetylspermidine: 107 V, 14 V (quantifier), 18 V (qualifier); N^8-acetylspermidine: 119 V, 14 V (quantifier), 22 V (qualifier); N^1, N^{12}-diacetylspermine: 113 V, 14 V (quantifier), 22 V (qualifier).

NIMS analysis

NIMS substrates were prepared as previously described [179]. In brief, p-type silicon wafers, 500 to 550 μm thick with 0.01 to 0.02 Ω cm resistivity (Silicon Quest International, Santa Clara, CA) were cut into 33 mm² pieces. The wafers were soaked in Piranha solution (sulfuric acid and hydrogen peroxide (2:1)) overnight, washed thoroughly with nanopure water and then dried using nitrogen gas. Etching was carried out by clamping the wafer in a Teflon chamber. Gold foil was used for the anode and a platinum loop as the cathode; a 25% ethanolic hydrogen fluoride solution was then added to the chamber. A BIO-RAD PowerPack1000 (Hercules, CA, USA) was connected and run at a constant-current mode (300 mA) for 30 minutes. The etched wafers were washed in methanol and evaporated to dryness using nitrogen gas. Bis(heptadecafluoro-1,1,2,2tetrahydrodecyl)tetramethyldisiloxane (Gelest, Morrisville, PA, USA) (100 μL) was applied to the surface of the chip and allowed

to sit at room temperature for 1 h before using nitrogen gas to remove excess from the surface. Tissue-Tek® Optimal Cutting Temperature (OCT) medium (Sakura Finetek, Torrance, CA) -embedded cancers were cut to 1-2 μm slices using a microtome and placed on top of the NIMS surface. A consecutive slice (5μm) was taken, applied to a Superfrost Plus microscope slide (Fisher Scientific, Signal Hill, CA) and stored at -80°C for histology. The NIMS chip was transferred to a room temperature vacuum for drying. Tissues were visibly dry within one minute. NIMS imaging data was acquired at 50 μm intervals using an AB/SCIEX 5800 TOF/TOF mass spectrometer in positive-mode. Images were analyzed using TissueView Software (Version 1.0) with a m/z range 50-400, bin size of 4 with 10819 data points)

SSAT immunohistochemical staining

Immunohistochemistry was performed on tissue fixed in 10% formalin and paraffin-embedded following standard procedures. Sections were de-paraffinized and rehydrated through a xylene, ethanol-water gradient. Antigen unmasking was performed by steaming in high temperature target retrieval solution (Target Retrieval Solution; Dako) for 45 minutes. Endogenous peroxidase activity was quenched by incubation with peroxidase block for 5 minutes at room temperature. Slides were incubated with an anti-SSAT antibody (Dr. Casero, 1:6000) overnight at 4°C. A horseradish peroxidase-labeled secondary antibody (PowerVision; Leica Microsystems) was applied for 30 minutes at room temperature. Signal detection was performed using 3,3'-diamino-benzidine (DAB) as the chromagen. Slides were counterstained with hematoxylin, dehydrated and mounted with glass coverslips using Permount (Fisher Scientific).

Microbial Culture

Tissue was collected following surgical resection for microbial cultivation in aerobic and anaerobic conditions. Anaerobic tissue specimens collected in specialized transport media (Anaerobe systems) were washed twice with 0.016% DTT in saline prior to hand homogenization in saline under anaerobic conditions. Tissue homogenate was diluted (10^0-10^6) and plated on pre-reduced non-selective Brucella blood agar (Bru) plates. Plates were stored under anaerobic conditions at 37°C until colony forming unit counts could be obtained (24-72 hours).

Figure 3.1. Untargeted liquid chromatography quadrupole time-of-flight mass spectrometry (LC-QTOFMS) metabolomics. **(A)** Cloud plot generated by XCMS Online showing dysregulated features between colon cancer and matched paired normal tissues (N=30, two-tailed Wilcoxon matched-pairs signed rank test). Total ion chromatograms (TICs) for each sample can be seen on the plot; features whose intensity are increased in cancer tissues are shown on the top part of the plot as blue circles and features whose intensity decreases in cancer tissues are shown on the bottom part of the plot as green circles. Larger and brighter circles (features) correspond to larger fold changes and higher p-values respectively. **(B)** Corresponding relative abundance of N^1, N^{12}-diacetylspermine integrated from extracted ion chromatograms (EICs) (two-tailed Wilcoxon test), **(C)** Cloud plot comparing cancer tissue from biofilm negative patients (lower part of plot) to biofilm positive cancers from CRC patients (upper part of plot) (two-tailed Mann-Whitney test) (n=11 biofilm negative cancers, n=8 biofilm positive cancers). **(D)** Corresponding relative abundance of N^1, N^{12}-diacetylspermine integrated from EICs (two-tailed Mann-Whitney test) **p<0.01. Also shown is a tandem MS spectrum of N^1, N^{12}-diacetylspermine in colon cancer samples and standard compounds. **(E)** biofilm positive normal tissues compared to their paired biofilm positive cancer tissues (N=7 tissue pairs examined) and **(F)** biofilm negative normal tissues compared to their paired biofilm negative cancer tissues (N=11 tissue pairs examined).

A. Normal vs. Cancer Tissues

B. N^1,N^{12}-diacetylspermine
$p<0.0001$
Normal   Tumor

C. Biofilm Positive Cancer vs.
Biofilm Negative Cancer Tissue

D. N^1,N^{12}-diacetylspermine
$p<0.01$
Tumor Biofilm Negative   Tumor Biofilm Positive

E. Biofilm Positive
Normal vs. Cancer Tissue

F. Biofilm Negative
Normal vs. Cancer Tissue
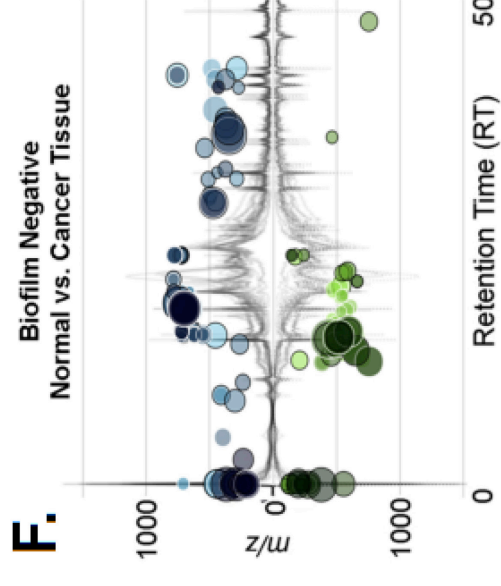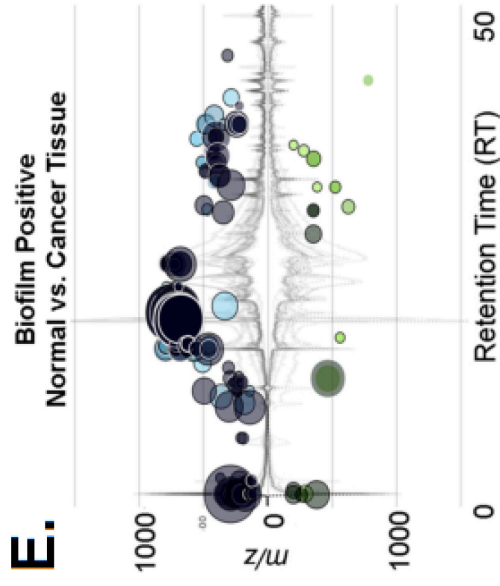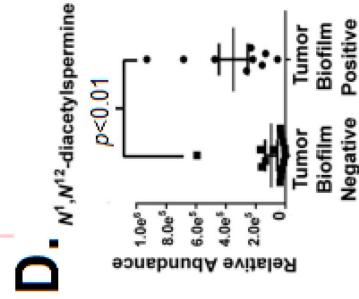
Figure 3.2. Biofilm effects on metabolites in colon tissues. **(A)** Scheme of polyamine metabolism. Enzymes involved include polyamine oxidase (PAO), spermidine/spermine N^1-acetyltransferase (SSAT), ornithine decarboxylase (ODC), histone acetyltransferase (HAT), spermidine synthase (SRM), spermine oxidase (SMO), spermine synthase (SMS). **(B-E)** Targeted liquid chromatography triple quadrupole mass spectrometry (LC-QqQ-MS) metabolomics. Absolute concentrations of spermine, spermidine, N^1-acetylspermine, N^1-acetylspermidine and N^1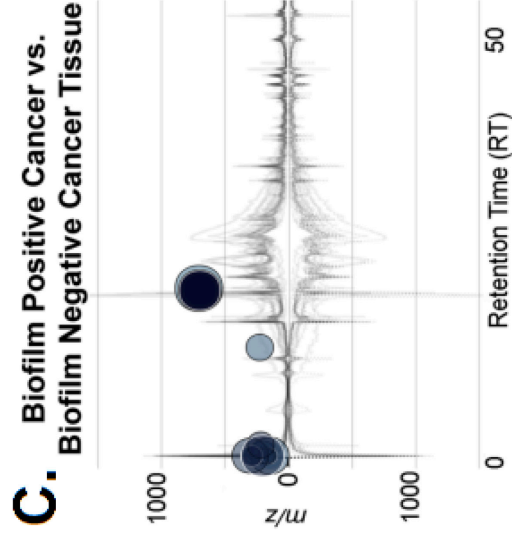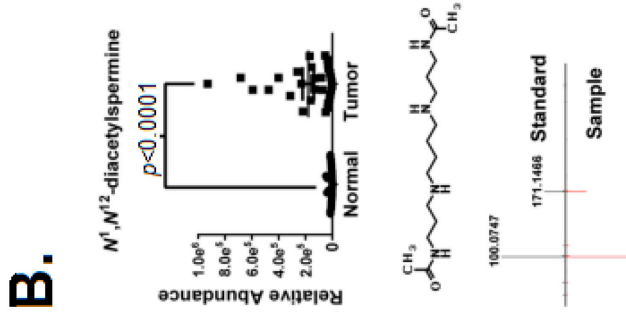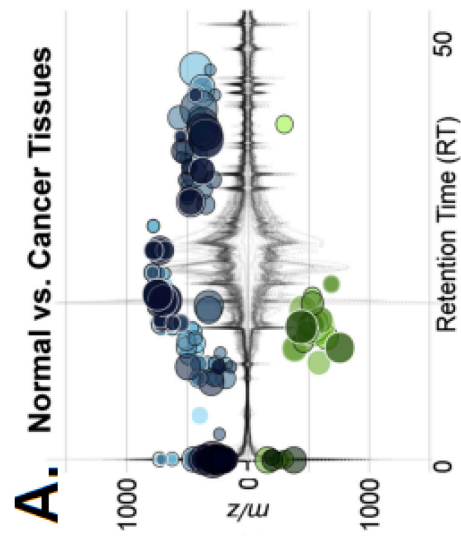, N^{12}-diacetylspermine in: **(B)** paired normal and cancer tissues with biofilms (two-tailed Wilcoxon matched-pairs signed rank test, n=8 pairs) **(C)** paired normal and cancer tissues lacking biofilms (two-tailed Wilcoxon matched-pairs signed rank test, n=11 pairs) **(D)** cancers with or without biofilms (two-tailed Mann-Whitney test, n=9 biofilm positive and n=11 biofilm negative) **(E)** paired normal tissues with or without biofilms from CRC patients (two-tailed Mann-Whitney test n=8 biofilm positive and n=11 biofilm negative). Left-sided biofilm positive samples are indicated by empty red circle (cancers) or empty blue square (paired normal tissues) symbols. Error bars are SEM, *p<0.05, **p<0.01, ***p<0.001, n.s = not significant.
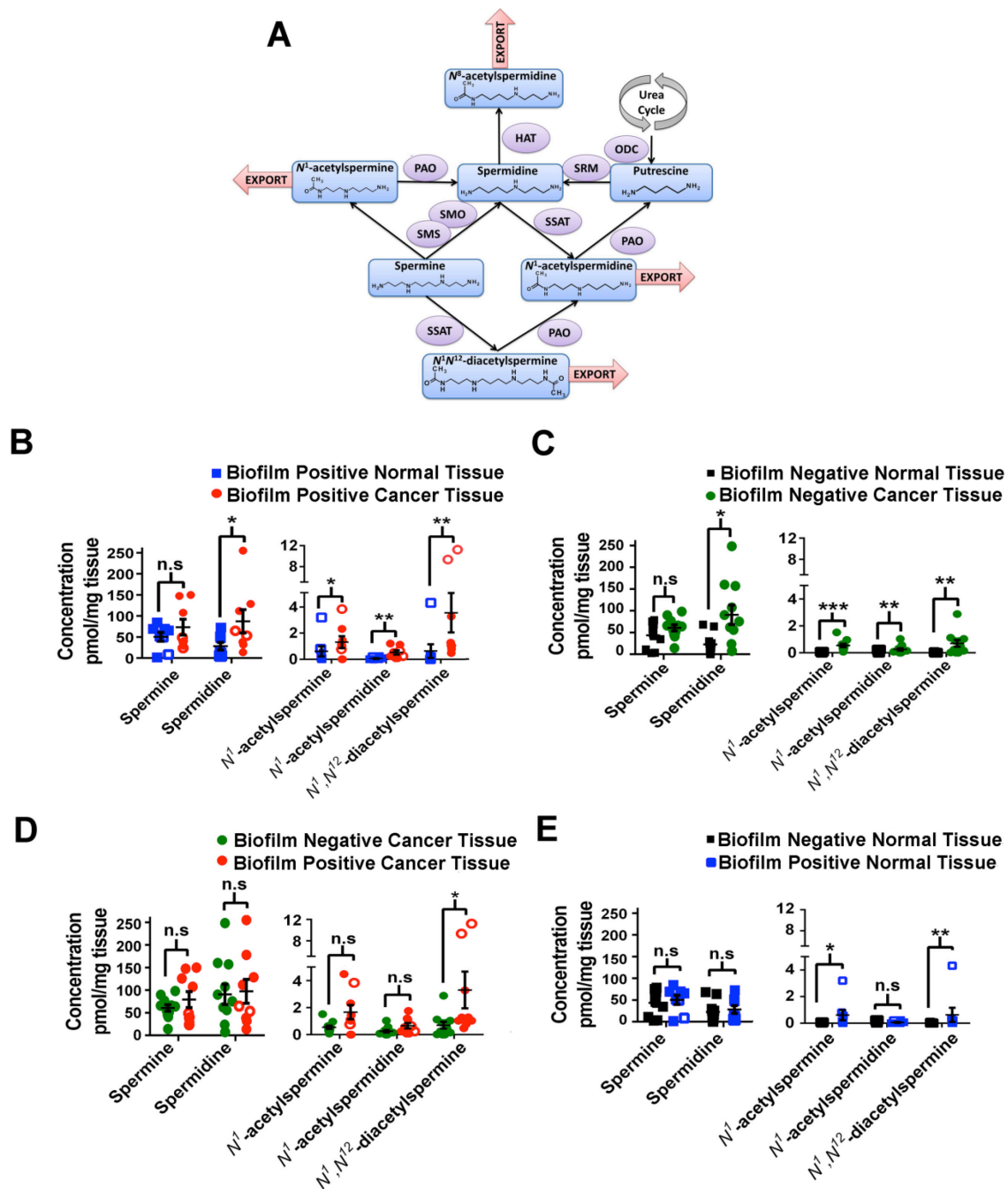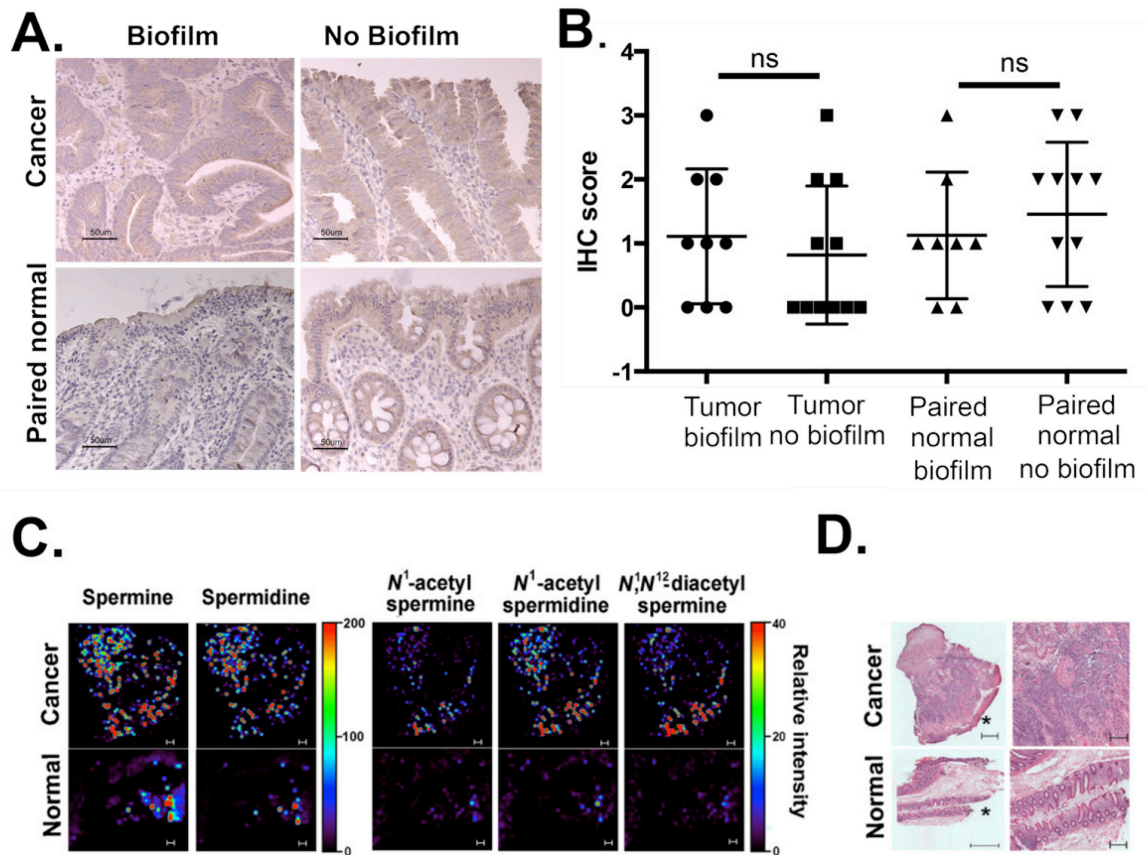
Figure 3.3. Localization and quantification of SSAT and polyamines **(A)** SSAT IHC of cancer and paired normal tissues from patients with and without a biofilm **(**scale bar 50um**)** **(B)** Scoring of epithelial SSAT IHC in Cancers and paired normal tissues with (n=9) and without (n=11) a biofilm **(C)** Nanostructure-initiator mass spectrometry (NIMS) imaging on paired biofilm positive normal and cancer tissues showing spatial specificity of polyamines in tissues. Scale bars approximately 100μm. **(D)** Hematoxylin and eosin staining of a consecutive section of tissues displayed in **(C)** showing cancer and normal colon tissue orientation. *mucosal edge of cancer and normal tissue. Scale bars: 500 μm left column; 200 μm right column.

Figure 3.4. Antibiotic-treatment clears biofilm from the mucosa and decreases polyamines. **(A)** FISH of all bacteria (red) displaying a bacterial biofilm on right cancer and paired normal tissue from a non-antibiotic-treated CRC patient, and no bacterial presence on a right cancer and paired normal from an antibiotic-treated patient. Tissues were counterstained with DAPI (scale bar 10 um). **(B)** Microbial culture data from patients treated with antibiotics (n=8) compared to non-antibiotic treated CRC patients (n=8). **(C-F)** Absolute concentrations of spermine, spermidine, N^1-acetylspermine, N^1-acetylspermidine and N^1, N^{12}-diacetylspermine in **(C)** right cancers from antibiotic-treated CRC patients (n=6) and cancers with biofilms (n=9) from non-antibiotic-treated colon cancer patients, **(D)** right cancers from antibiotic-treated colon cancer patients (n=6) and cancers without biofilms (n=11) from non-antibiotic-treated colon cancer patients, **(E)** colonoscopy biopsies from healthy subjects (n=8 biopsies, 4 each from the right and left colon) and paired normal tissues with biofilms from colon cancer patients (n=8), **(F)** colonoscopy biopsies from healthy individuals (n=8 biopsies, 4 each from the right and left colon) and paired normal tissues without biofilms from colon cancer patients (n=11). Left-sided biofilm positive samples are indicated by empty red circl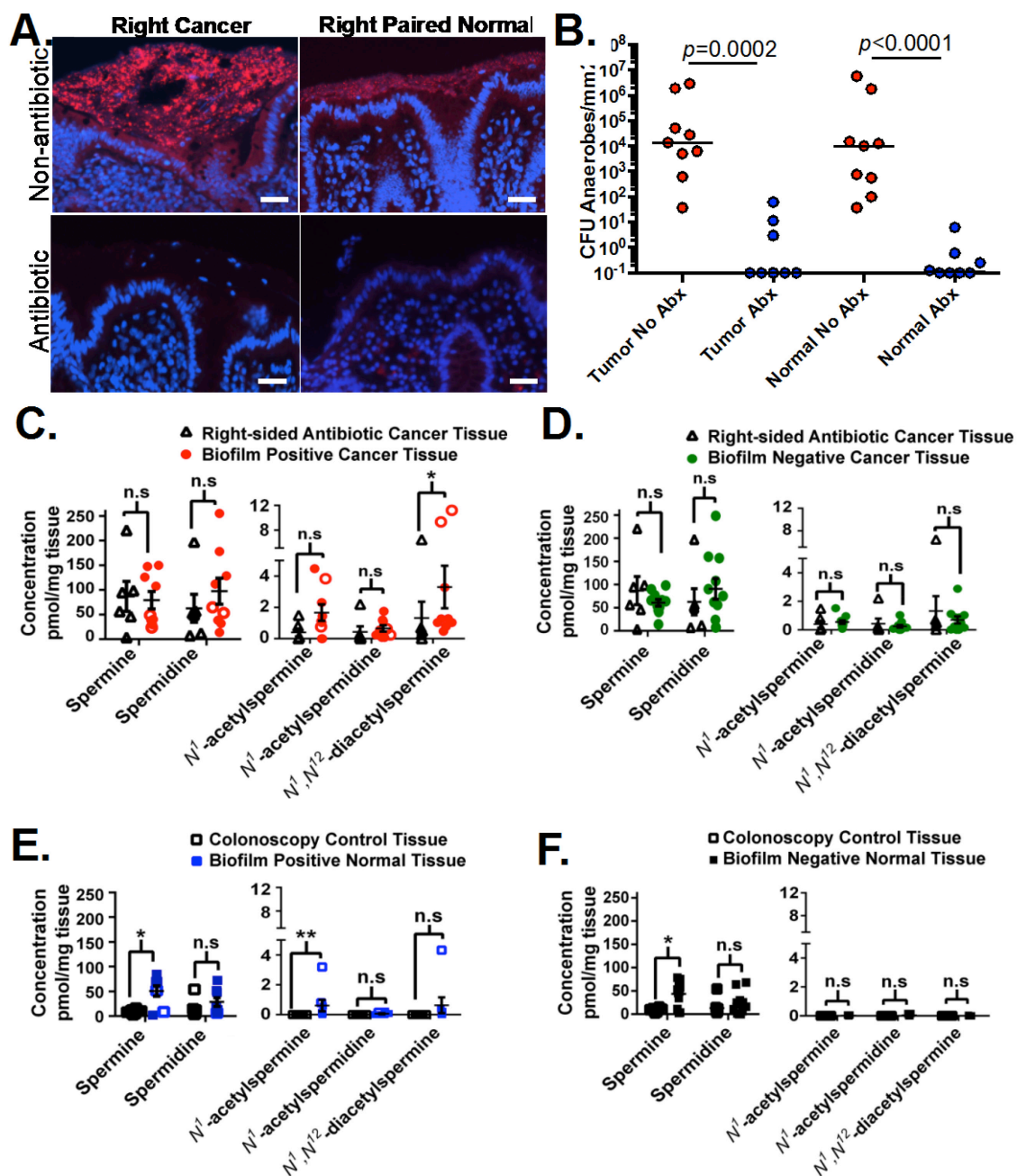e (cancers) or empty blue square (normal tissues) symbols. All were analyzed by two-tailed Mann-Whitney test, error bars are SEM, *p<0.05, **p<0.01, n.s = not significant.
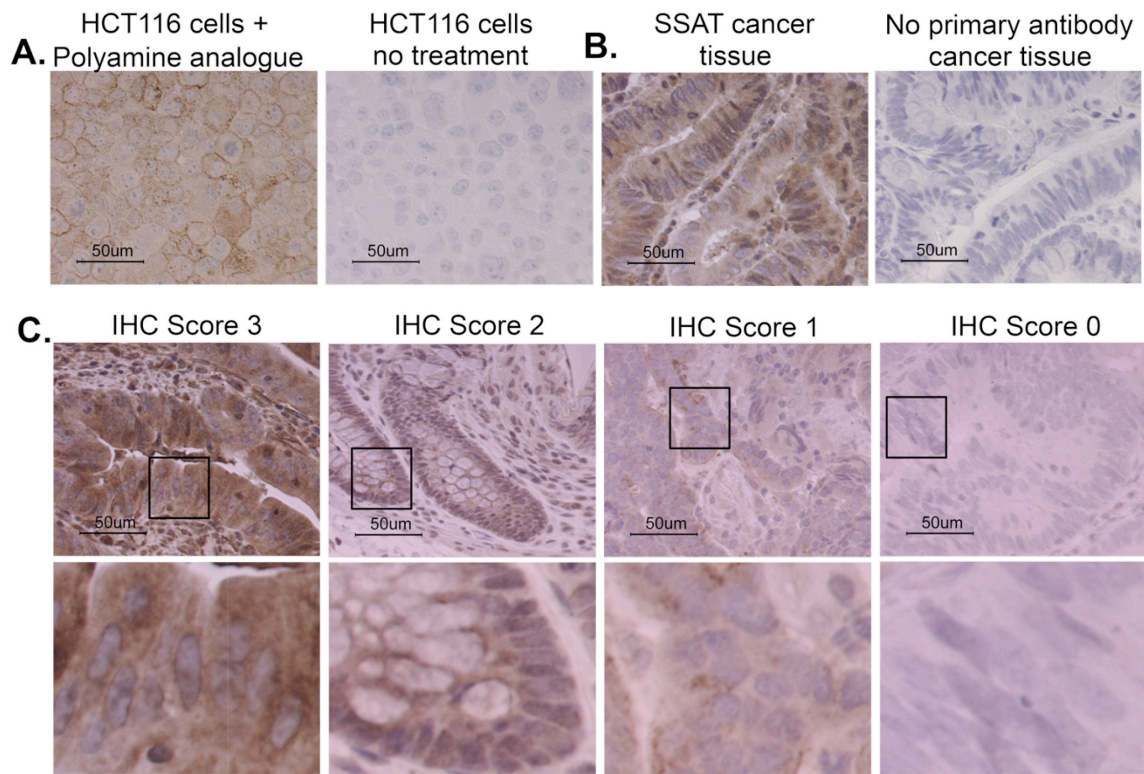
Figure S3.1. SSAT immunohistochemical analysis of CRC tissues. Images of control staining of **(A)** HCT116 cells stimulated with the polyamine analogue N1,N11-bis(ethyl) norpermine (first panel) displaying positive SSAT staining compared with unstimulated HCT116 (second panel) cells lacking any SSAT. **(B)** SSAT staining of cancer tissue with (first panel) and without (second panel) primary antibody. Immunohistochemical 0-3 scoring system **(C)** with representative cytoplasmic staining intensities. Selected inserts are displayed in the bottom panel. All images were captured at 400x.
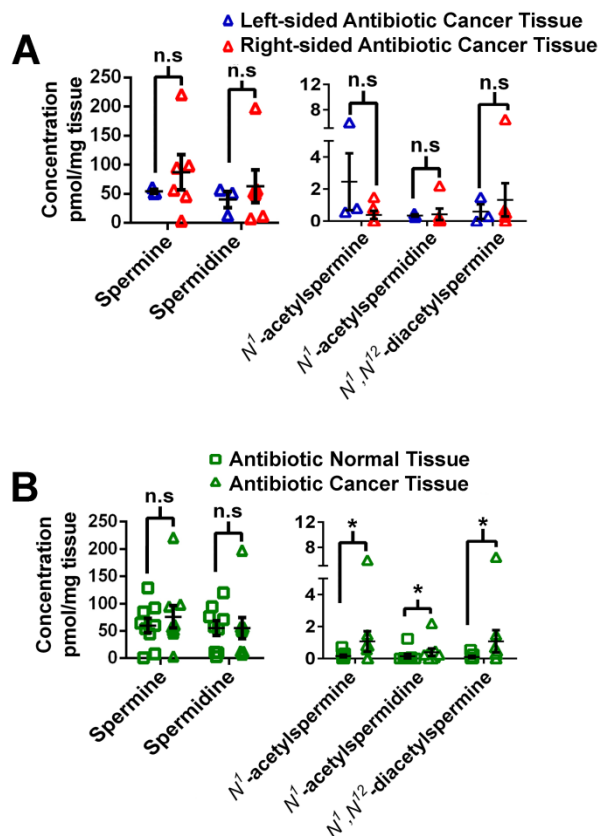
Figure S3.2. Further comparison of antibiotic-treated cancer tissues and normal colonoscopy biopsies from healthy individuals. Absolute concentrations of spermine, spermidine, N^1-acetylspermine, N^1-acetylspermidine and N^1, N^{12}-diacetylspermine in **(A)** left- and right-sided cancers (n=9, 3 left-sided, 6 right-sided) from antibiotic-treated colon cancer patients, **(B)** surgically-resected normal and cancers (n=9 tissue pairs) from antibiotic-treated colon cancer patients, and **(C)** left- and right-sided normal colonoscopy biopsies from healthy individuals (n=4 right and left colon biopsy pairs from 4 patients). Two-tailed Wilcoxon matched-pairs signed rank test. Error bars are SEM, *p<0.05, **p<0.01, ***p<0.001, n.s = not significant.

Table 3.1. Colorectal cancer patient metadata

Patient ID	Patient Type	Age	Sex	Race	Tumor Site	Biofilm	Stage	Tumor Size (mm)	Histology
3711*	Surgical CRC	45	M	Caucasian	Cecum	ND**	2	145.0	Adenocarcinoma
3727	Surgical CRC	44	F	African American	Cecum	ND**	3	50.0	Adenocarcinoma
3773*	Surgical CRC	58	F	African American	Cecum	No	4	35.0	Adenocarcinoma
3979	Surgical CRC	77	F	African American	Cecum	Yes	3	35.0	Adenocarcinoma
3989*	Surgical CRC	41	F	Caucasian	Cecum	No	4	85.0	Mucinous Adenocarcinoma
3728	Surgical CRC	69	M	Caucasian	Ascending	Yes	1	87.0	Adenocarcinoma
3731	Surgical CRC	74	M	Caucasian	Ascending	Yes	2	30.0	Adenocarcinoma
3744*	Surgical CRC	68	M	Caucasian	Ascending	No	4	80.0	Adenocarcinoma
3753	Surgical CRC	49	F	African American	Ascending	Yes	4	47.0	Mucinous Adenocarcinoma
3754	Surgical CRC	67	F	African American	Ascending	Yes	2	30.0	Adenocarcinoma
3762	Surgical CRC	73	M	Caucasian	Ascending	Yes	4	54.0	Adenocarcinoma
3779 B*	Surgical Polyp	74	F	Caucasian	Ascending	No	NA	42.0	Tubulovillous adenoma
3780*	Surgical CRC	78	F	Caucasian	Ascending	No	4	40.0	Adenocarcinoma
3991*	Surgical Polyp	70	M	Caucasian	Ascending	No	NA	31.0	Tubular adenoma
3994*	Surgical CRC	82	F	Caucasian	Ascending	No	2	30.0	Mucinous Adenocarcinoma
3987	Surgical CRC	66	F	Caucasian	Ascending	Yes	2	50.0	Mucinous Adenocarcinoma
3770	Surgical CRC	71	M	Caucasian	Hepatic Flexure	No	1	35.0	Adenocarcinoma
3774	Surgical CRC	45	M	Asian	Hepatic Flexure	Yes	2	45.0	Adenocarcinoma
3752	Surgical CRC	73	F	Caucasian	Transverse	No	2	25.0	Adenocarcinoma
3976	Surgical CRC	52	F	Caucasian	Transverse	No	1	20.0	Mucinous Adenocarcinoma
3768	Surgical CRC	45	M	African American	Transverse	No	4	42.0	Adenocarcinoma
3718	Surgical CRC	60	M	Caucasian	Splenic Flexure	ND**	2	35.0	Adenocarcinoma
3769	Surgical CRC	78	F	African American	Splenic Flexure	No	3	60.0	Adenocarcinoma
3992	Surgical CRC	91	F	Caucasian	Splenic Flexure	Yes	2	45.0	Adenocarcinoma
3749	Surgical CRC	39	M	Caucasian	Sigmoid	No	3	50.0	Adenocarcinoma
3756	Surgical CRC	54	M	Caucasian	Sigmoid	Yes	4	45.0	Adenocarcinoma
3766	Surgical CRC	56	F	Caucasian	Sigmoid	No	4	55.0	Adenocarcinoma
3786*	Surgical CRC	56	M	Caucasian	Sigmoid	ND**	4	10.0	Adenocarcinoma
3977	Surgical CRC	38	F	Caucasian	Sigmoid	No	1	50.0	Adenocarcinoma
3712	Surgical CRC	79	F	Caucasian	Rectosigmoid	ND**	3	65.0	Adenocarcinoma
3721	Surgical CRC	55	F	Caucasian	Rectosigmoid	ND**	2	50.0	Adenocarcinoma
3736	Surgical CRC	56	M	African American	Rectosigmoid	ND**	3	53.0	Adenocarcinoma
3760	Surgical CRC	29	F	Caucasian	Rectosigmoid	No	2	80.0	Adenocarcinoma
3788*	Surgical CRC	52	M	Caucasian	Rectosigmoid	No	4	40.0	Adenocarcinoma
3719	Surgical CRC	31	F	Asian	Rectum	ND**	2	65.0	Adenocarcinoma
3735	Surgical CRC	64	M	Caucasian	Rectum	No	3	70.0	Adenocarcinoma
3759*	Surgical CRC	87	F	Caucasian	Rectum	No	4	45.0	Adenocarcinoma
3978	Surgical CRC	90	F	Caucasian	Rectum	No	1	27.0	Mucinous Adenocarcinoma

*Patient received antibiotics day prior to surgery
**Biofilm indeterminate (not fixed in Carnoy's for mucus optimization)

Table 3.2. Healthy subject metadata

Patient ID	Age	Sex	Race	Bowel Prep	Biopsy Site	Biofilm
52	43	F	Caucasian	NuLytely Prep	Descending	No
					Ascending	No
53	47	M	Caucasian	GoLytely Prep	Descending	No
					Ascending	No
54	55	M	Caucasian	NuLytely Prep	Descending	No
					Ascending	No
55	60	F	African American	GoLytely Prep	Descending	No
					Ascending	No

Chapter 4

Microbial associations
of familial polyposis

4.1 Abstract

Familial adenomatous polyposis (FAP) is a dominantly inherited condition characterized by a germline mutation in the adenomatous polyposis coli (*APC*) tumor suppressor gene that confers a 100% lifetime risk of developing colorectal cancer (CRC). It remains unclear what causes the accumulation of additional mutations yielding CRC, however, recent evidence has emphasized contributions by the gut microbiome. Abnormalities have been reported regarding the colonic microbiome of individuals with sporadic CRC; however, limited microbiome analyses of individuals with hereditary CRC exist. We investigated the composition and microbial organization of the FAP microbiome. Sequence analysis of mucosal samples revealed enrichment of *Bacteroides* on FAP polyps when compared to healthy control subjects. Further, fluorescent in situ hybridization (FISH) of mucosal tissues revealed conserved and distinct bacterial biofilms throughout the colon, comprised predominately of *Escherichia coli* and *Bacteroides fragilis*, adherent to both polyps and histologically normal tissue. These findings were supported by semi-quantitative microbial culture data revealing increased detection of *E. coli* and *B. fragilis* on mucosal tissue from FAP patients compared to healthy subjects. Further, the oncogenic toxins Colibactin (from the *E. coli pks* pathogenicity island) and *Bacteroides fragilis* toxin (BFT), known to be expressed by these two organisms, were highly enriched in FAP patients (68%, n=25 and 60%, n=25, for *pks+ E. coli* and ETBF, respectively) when compared to healthy subjects without FAP (22%, n=23 and 30%, n=23, for *pks+ E. coli* and ETBF, respectively). Interestingly, if any individual, FAP or a healthy subject, was colonized by one of these organisms they likely harbored the other (co-colonization 52%, n=25, for FAP and 22%, n=23 for healthy subjects). Preliminary mouse studies of co-colonization indicate that these two organisms may act additively to increase DNA damage and colitis in WT mice. These data suggest that

co-colonization by *pks+ E. coli* and ETBF may be associated with a pro-oncogenic state in FAP patients.

4.2 Introduction

Colorectal cancer (CRC) is one of the most common malignancies worldwide, with an annual incidence of 1.2 million and mortality of over 600,000 [1]. It is widely understood that CRC develops through a sequential accumulation of mutations that facilitate the transition from normal mucosa, to adenoma, to adenocarcinoma. Greater than 90% of cases occur in individuals with little or no genetic risk, however, a significant minority (approximately 5% of cases) occur as a result of an inherited mutation [4]. One such hereditary condition, Familial Adenomatous Polyposis coli (FAP), is caused by germline mutations in the *APC* gene. Over 700 different mutation sites have been identified, leading to variable onset development of multiple adenomas (polyps) throughout the colon and rectum. Individuals with this genetic defect are born with their first mutation in the transition to CRC, and as additional mutations accumulate, hundreds to thousands of colorectal polyps develop. Environmental factors contributing to these mutations are an area of immense interest and may be similar to those in the development of sporadic CRC. In particular, bacterial initiators and promoters of CRC progression have long been proposed [136].

The human colon coexists in close proximity to trillions of bacteria that collectively constitute the colonic microbiome. Luminal bacteria are segregated from the host by a dense mucus layer that extends throughout the large intestine, preventing direct contact with epithelial cells [133]. In a healthy state this protective coating serves to promote tolerance of foreign antigens, limiting mucosal inflammatory responses; In contrast, persistent bacterial

breaches of the colonic mucus layer with, in some cases, biofilm formation fosters chronic mucosal inflammation [134]. Colonic biofilms characterize the disease states of Crohn's, ulcerative colitis, self-limiting colitis and have recently been associated with sporadic colorectal cancer (**CHAPTER**,[172]). The work presented in this chapter tests the hypothesis that FAP patients would share a similar dysbiosis in microbial associations.

4.3 Results

We investigated the intestinal microbiome composition and spatial arrangement in patients with known familial polyp syndromes undergoing surgery at Johns Hopkins Hospital. All patients had a phenotype consistent with familial polyposis and a subset underwent genetic counseling and mutational analysis confirming a mutation in the *APC* gene (Table S4.1). To conduct our analyses, tissue not needed for clinical evaluation was collected from sites throughout the colon (polyp and grossly normal tissue) and preserved in Carnoy's fixative (n=6 cases, n=22 controls) or snap frozen (n=23 cases, n=22 controls).

Microbial 454 sequence analysis was conducted using DNA extracted from the mucosal surface of polyps and paired normal tissue from four FAP patients harboring three different germline mutations (1 MYH-attenuated phenotype, 2 APC, and 1 Juvenile Polyposis(JP)) (Table S4.1). Results were compared with a parallel sequence analysis (previously reported) of colonoscopy biopsies collected from 10 healthy volunteers as well as tumor and paired normal tissue collected from 28 patients with sporadic CRC (**CHAPTER 2**, [172]) (Fig 4.1A). Despite the limited number of samples analyzed, notable trends were observed. FAP patients were found to have bacterial populations of predominantly *Bacteroides* on polyp mucosa. The *Bacteroides* levels were significantly increased when compared to sporadic CRC patients (tumor or normal mucosa) and normal mucosal biopsies collected from healthy

subjects (Fig 4.1 B). Further, specific species enrichment of *Prevotella stercorea* was identified on FAP polyp mucosa as compared to mucosal biopsies from healthy subjects (Fig 4.1). FAP polyps displayed a significant decrease in *Proteobacteria* (both *Gammaproteobateria* and *Betaproteobacteria*) when compared to normal mucosal biopsies from healthy controls (Fig 4.1E). There were no differentially abundant taxa identified at any taxonomic level between polyps and their respective paired normal tissues from any FAP patient. This is in contrast to sporadic CRC tumor epithelium that is characterized by a progressive microbial dysbiosis when compared to paired normal mucosa; namely sequences of sporadic CRC mucosa is a complete subset of its paired normal in 52% of individuals, with particular enrichment of *Fusobacteria*. In our sample set, only the JP patient had detectable *Fusobacteria* (*F. necrophorum*) present on a polyp (>63% of total reads were *F. necrophorum*) and normal flanking tissue (>16% of total reads were *F. necrophorum*) collected from the left colon, whereas *Fusobacteria* was largely absent from the polyp and normal flanking tissue in the right colon of this patient (<0.3% of sequencing reads) (Fig4.1A). No detectable *Fusobacterium* reads were identified on the mucosa (polyp or normal) or in the stool of the three additional FAP patients.

Of interest, stool was collected and sequenced from two sites (ascending and descending) from the FAP patient with the MYH mutation along with sequencing of mucosa from each site. Analysis of the two stool samples from this individual revealed that the luminal stool population was dynamic along the GI tract. Strikingly, the local stool microbiome reflected the mucosal population from the site of collection rather than stool collected from a different position along the gastrointestinal tract; stool from the ascending colon was more similar to the ascending mucosa than to stool collected in the descending colon and vice versa (Fig.4.1A).

In addition to sequence analysis, the spatial arrangement of bacteria on the colonic epithelium is also of interest as our group recently identified invasive polymicrobial bacterial biofilms on sporadic CRC (**CHAPTER 2**, [172]). In sporadic CRC, biofilms were detected on all tumors and paired normal tissues from the right colon (proximal to the hepatic flexure), as well as a subset of tumors and paired normal tissues from the left colon (distal to the hepatic flexure). Further, biofilms were associated with decreased crypt E-cadherin, increased epithelial cell IL-6, pSTAT-3, polyamine metabolites and proliferation, all potential contributors to malignant transformation (**CHAPTERS 2 & 3**, [172]). Herein we screened surgically resected tissue collected along the axis of the colon (cecum to rectum) from six FAP patients (Table 4.1) to evaluate bacterial biofilms. Polyps and macroscopically normal tissue were labeled with a FISH probe (EUB338) designed to recognize 16s ribosomal RNA of all bacteria. Four of the six FAP patients exhibited a bacterial biofilms scattered along the colonic axis (Fig 4.2A,B, FigS4.2). Unlike biofilms detected on sporadic CRC that are continuous along the mucosal surface, FAP tissue displayed patchy biofilm formation with bacterial mucus invasion spanning 200-700 linear um on approximately 70% of surgically resected colon specimens from four of the six patients. Biofilms were not restricted to polyps, nor did they display geographic preference within the colon as observed in sporadic CRC (Table 4.1, FigS4.2, FigS4.3). Of note, biofilms were not detected on tissues (polyp or paired normal) from the individual with juvenile polyposis, a condition where polyps never transition to cancer (FigS4.3). Biofilms were also not detected on the colon of an individual that received oral antibiotics 24 hours before surgery, a treatment known to clear mucosal associated biofilms (**CHAPTER 2**)(FigS4.3).

Specimens found to contain a biofilm were further screened by additional probes designed to recognize the major phyla detected in biofilms of sporadic CRC: *Bacteroides/Prevotella*,

Proteobacteria, *Lachnospiraceae*, and *Fusobacteria* (Table S4.3). Interestingly, the biofilms of FAP

patients were comprised most prominently (~70%) of *Proteobacteria* (Fig4.2B), a minority

member of the FAP microbiome based on the 454 sequence analysis. The major group

identified through sequence analysis, *Bacteroides*, was also detected, however it was found to

make up 5-25% of the biofilms (Table 4.1). Neither *Fusobacteria* nor *Lachnospiraceae* were

detected by FISH. Through additional, more specific probe sets (Table S4.3), the

predominant biofilm members were identified as *E. coli* and *B. fragilis* (Fig 4.2B). Similar to

the biofilms on the normal tissue of sporadic CRC patients, epithelial cell invasion by the

biofilm community members was observed sporadically on all patients harboring a biofilm

(Fig 4.2C).

Semi-quantitative microbial culture analysis was carried out to determine the overall relative

abundance of *E. coli* and *B. fragilis* on mucosal FAP samples. Mucosal tissue was

homogenized and serial dilutions were cultured on selective Bacteroides Bile Esculin agar

(*Bacteroides*) or MacConkey plates (Lactose-fermenting *Enterobacteriaceae*). Colony forming

units (CFU) per mm^2 FAP tissue were compared to CFU per mm^2 on tissues collected from

control subjects. From each FAP patient, polyps and corresponding normal tissue, and from

control subjects tissue from both the left and right side colon were processed and the

combined results of each patient were analyzed. Consistent with FISH analysis, cultivatable

Bacteroides and Lactose-fermenting *Enterobacteriacae* were significantly increased on FAP

tissues when compared to control subjects regardless of the tissue collection site (Fig 4.2D).

Strong experimental evidence exists supporting the oncogenic potential of molecular

subtypes of both *E. coli* and *B. fragilis* [47,192]; the two dominant biofilm members identified

in direct contact with host epithelial cells in our FAP patients. *E. coli* containing the *pks*

pathogenecity island, which encodes the genes responsible for synthesis of the colibactin

genotoxin, induces DNA damage *in vitro* and *in vivo* along with tumorigenesis in AOM/IL10 deficient mice[47]. Additionally, enterotoxigenic *Bacteroides fragilis* (ETBF) induces multiple tumors in the distal colon of APC mice, a CRC mouse model that carries a truncation mutation in APC [192].

In addition to experimental evidence, human epidemiological studies have associated both organisms (ETBF and *pks+ E. coli*) with inflammatory bowel disease and sporadic colorectal cancer [47,54,193,194]. Thus, we screened banked frozen mucosal tissues from 25 FAP patients (2 polyps and 2 paired normal tissues per patient when available) and 22 healthy subjects (1 ascending and 1 descending biopsy per subject) for the presence of *pks+ E. coli* and ETBF. The mucosa of FAP patients was significantly associated with *pks+ E. coli* (68%, n=25 patients) and ETBF (60%, n=25) when compared to mucosa of healthy subjects (22% *pks+ E. coli* and 30% ETBF+, n=23) (Table 4.2). There was no preferential association between the presence of ETBF or *pks+ E. coli* on polyp mucosa vs. paired normal mucosa from FAP patients (Fig S4.3). Typically, when a mucosal sample was identified as having *pks+ E. coli* or ETBF, the additional mucosal specimen(s) from that subject was also positive (73% for *pks+ E. coli*, 59% for ETBF), similar to our recently published results for BFT in sporadic CRC patients [194] . Even more noteworthy, if a patient had *pks+ E. coli*, there was a high likelihood of also carrying ETBF; this was the case in both FAP patients and healthy control subjects (Table 4.2).

The high frequency of double colonization with *pks+ E. coli* and ETBF, independent of disease state, highlights the importance of understanding the potential effects on the host of simultaneously harboring these two oncogenic organisms. To this end, we characterized specific pathogen free (SPF) C57BL/6J mice co-inoculated with a *pks+ E. coli* isolated from an FAP patient (PL1) and a laboratory ETBF strain (086-5443-2-2) used in past tumor

mouse studies. Initially, three FAP *pks+E.coli* isolates and the canonical NC101 strain

utilized in murine tumor experiments were characterized in C57BL/6J. Histopathology was

assessed at seven days post inoculation to detect acute colitis and the most virulent *pks+ E.*

coli isolate (PL1) was selected for subsequent co-colonization experiments (Fig S4.3). Co-

colonization experiments were analyzed at seven days post-colonization. Mice were assessed

for DNA damage and inflammatory cytokines as compared with singly infected mice

colonized by just ETBF or *pks+ E. coli* as well as sham mice. Co-colonized mice had

markedly increased DNA damage along the entire length of the colon, when compared with

singly colonized mice as detected by phosphorylated histone variant H2AX (γH2AX) (Fig.

4.3A). Quantification of percentage of positive cells per crypt revealed that γH2AX was

significantly increased when compared to mice singly infected with either *pks+ E. coli* or

ETBF (Fig 4.3 B). Further, analysis of inflammatory cytokines revealed significant

upregulation of iNOS and IL1-beta mRNA in co-infected mice when compared to singly

infected mice (Fig 4.3 C, D).

4.4 Discussion

Here we found that individuals with hereditary cancer, FAP, harbor a unique microbiome

with altered epithelial associations. Individuals with FAP harbor bacterial biofilms along the

length of the colon comprised predominantly of *E. coli* and *B. fragilis* in direct contact with

the mucosa, with intermittent invasion of epithelial cells. Microbial culture analysis of

mucosal samples from FAP vs. normal control subjects supports increased *Bacteroides* and

Enterobacteriaceae populations on FAP mucosa.

While, sequencing results were not consistent with the FISH and culture data, the sample

size and tissue processing methodology could explain the differences observed. Further,

oncogenic members of these species, *pks+ E. coli* and ETBF, were significantly associated

with FAP patients when compared to healthy subjects. The close proximity of biofilm members to host epithelium highlights the potential for this population to transfer oncogenic toxins directly to the epithelium. The effect of co-colonization with these two organisms revealed a synergistic increase in DNA damage and inflammatory cytokines. Together these data support the notion that FAP patients harbor mucosal-associated members of the microbiota may act in consort to increase DNA damage and inflammation. Analyses of the CRC microbiome have revealed distinct microbial communities when compared to healthy individuals. While specific species have been associated with CRC there has been no consensus what microbial members are important. Speculations that polymicrobial contributions drive CRC have been proposed, suggesting that combinations of specific species may act in consort to promote tumorigenesis [170,195,196]. To our knowledge this is the first data, that links the carriage of one oncogenic organism with another in a cancer host. The preliminary analysis of co-infection with more than one putative oncogenic agent in mice strongly suggests that two organisms may act together to lead to increased pathology. Further exploration of the tumorigenic potential of co-colonization with these two organisms in murine models of carcinogenesis is underway.

4.5 Materials and Methods

Patient selection and Sample acquisition

Polyps and paired normal tissues were collected from patients with Familial Polyposis undergoing surgery at Johns Hopkins Hospital. All patients had a phenotype consistent with familial polyposis and a subset underwent genetic counseling and mutational analysis. Two bowel preparations were routinely used and recorded (mechanical bowel preparation [Miralax™], or Fleet Phospo-soda™ enema), pre-operative intravenous antibiotics were

administered in all cases (cefotetan or clindamycin/gentamycin) immediately preceding surgery. One patient received pre-operative oral antibiotics as noted in the patient metadata (Table S4.1). Normal control mucosal biopsies and surgically resected colon from grossly normal tissue were collected from individuals without a history of CRC, inflammatory bowel disease, or antibiotic usage within three months (designated herein as control subjects) (Table S4.2). All tissue not needed for pathologic diagnosis was rapidly preserved in Carnoy's solution, RNAlater, anaerobic transport media or snap frozen for subsequent analysis.

This study was approved by the Johns Hopkins Institutional Review Board. All samples were obtained in accordance with the Health Insurance Portability and Accountability Act (HIPAA).

Sample Preparation for Sequencing

Samples were prepared as previously described (**CHAPTER 2**, [172]). Briefly, mucosal samples collected from FAP polyp, FAP paired normal tissues, surgically resected control normal tissues, and colonoscopy biopsies were collected in the pathology or endoscopy suites and immediately placed in RNAlater (Qiagen Inc. Germantown, MD) and stored at -80°C. Tissue samples were incubated at 95°C in ASL buffer with frequent vortexing to remove bacteria from the epithelial surface. Following the dislodging of mucosal associated bacteria, supernatant was removed and cells were thoroughly lysed using pressure lysis and DNA was extracted using the QIAamp DNA Stool Kit (Qiagen). The V3-V5 region of bacterial 16S rDNA was amplified and sequenced following the procedures described by the Human Microbiome Project standard protocol (http://www.hmpdacc.org/doc/16S_Sequencing_SOP_4.2.2.pdf). Briefly, the V3-V5 region of 16S rDNA was amplified with PCR primers (357F 5' CCTACGGGAGGCAGCAG 3'

and 926R 5' CCGTCAATTCMTTTRAGT 3') that were appended with Roche 454 Titanium

FLX library adapter sequences. All B-adapter primers were identical, while A-adapter

primers also contained a unique barcode of 5-10 nucleotides to allow indexing of individual

samples. Each sample was PCR amplified and purified. Purified DNAs were quantified

using the 454 FLX Library Quantification Kit (KAPA Biosystems Inc. Woburn, MA) and

pooled for sequencing in equal molar quantity.

Sequence data analysis

Raw sequence reads were analyzed as previously described (**CHAPTER 2**, [172]). Briefly,

samples were identified using their barcodes, trimmed and filtered for quality and length

(minimum 150 bp) using the QIIME package (v1.6.0). High quality reads were organized by

sample and error-corrected using the Acacia tool (v1.52), *de novo* UCHIME (v4.2.40), and

RDP Bayesian classifier (v2.5). The final high-quality contaminant-free dataset was then

submitted to the CloVR-16S pipeline (v1.1) for diversity estimation, taxonomic

characterization and comparative analysis of sample groups of interest. Downstream analysis

included clustering of sequences into species-level OTUs (95% identity threshold),

taxonomic assignment of OTU representatives and beta-diversity estimation. Pipeline runs

were executed using CloVR (v2012.11.16) on the DIAG academic cloud

(http://diagcomputing.org).

Fluorescent in situ hybridization

Carnoy's fixed, paraffin-embedded tissues were sectioned to 5 μm thickness and de-waxed

following standard procedures. Sections were stained with Periodic acid Schiff (PAS) to

confirm mucus presence and preservation and successive sections were hybridized with the

Eub338 universal bacterial probe and with a nonsense probe to test for nonspecific binding of probes. Slides were imaged using a Nikon E800 microscope with NIS elements software.

Samples that were determined to have a bacterial presence by universal probe were next analyzed by a more specific probe set synthesized and conjugated at the 5' end to the fluorophores listed in Table S3 (Life Technologies). Probes were applied to slides at a concentration of 2 pmol/ul of each probe in prewarmed hybridization buffer (900 mM NaCl, 20 mM Tris pH 7.5, 0.01% SDS, 20% formamide). Slides were incubated at 46°C in a humid chamber for 2 hours, and washed at 48°C for 15 minutes in wash buffer (215 mM NaCl, 20 mM Tris pH 7.5, 5 mM EDTA). Slides were mounted using ProLongGold antifade reagent (Life Technologies).

Biofilm quantification

Bacterial biofilms were quantified for longitudinal distance along epithelium, depth, and density using slides hybridized with the universal bacterial probe (Table S4.3). When present, up to five biofilm measurements were taken of the entire longitudinal length of biofilm along the surface of the epithelium (some samples did not have five patches of biofilm, in these cases all present biofilms were measured). The average of five measurements of biofilm depth was taken from five tissue samples per patient. Relative biofilm species quantification was performed using tissues hybridized with the universal bacterial probe along with *B. fragilis* and *E. coli* species specific probes (Table S4.3). One specimen per sample was selected for *E. coli* and *B. fragilis* relative quantification. Images were taken at 100x magnification and individual bacterial cells (all bacteria, *E. coli*, and *B. fragilis*) in a 10x10 μm space were counted. Five 10x10 μm boxes were counted per patient to determine the relative biofilm composition as a percentage.

Microbial Culture

Enterobacteriaceae and Bacteroides

In the pathology suite, tissue for culturing (8mm punches) was rapidly enclosed in anaerobic transport media to maintain the anaerobic environment required for *Bacteroides*. Tissue of both polyps and normal tissue of FAP patients and left and right-sided tissue of surgery controls was washed and homogenized in an anaerobic hood as described previously [194]. In short, after homogenization, serial dilutions (10^0-10^6) of the tissue homogenate were plated on Bacteroides Bile Esculin (BBE) agar for selective culturing of *Bacteroides*. Subsequently, BBE plates were incubated in an anaerobic incubator for 48 hours and single colony forming units (CFU) were assessed on plates with separate single colonies. Similarly, serial dilutions of the tissue homogenates were plated on MacConkey agar and incubated for 24 hours in an aerobic incubator. Subsequently, the CFU of single pink colonies (Lac^+ *Enterobacteriaceae*) was assessed. CFUs of both polyp and normal tissue of FAP patients was compared to CFUs of left and right-sided control tissue for both *Bacteroides* and *Enterobacteriaceae* with Mann-Whitney U statistics.

Detection of *pks E. coli* and ETBF

Tissue stored at -80°C was utilized for microbial culture and selective amplification and identification of *E. coli* and *B. fragilis* isolates. Two to four mucosal samples per patient were available for microbiology culture analysis (metadata table S4.1 and S4.2). An approximately 3mm diameter punch of mucosal sample from surgically-resected control tissue, FAP polyp, FAP paired normal, or colonoscopy biopsy was placed in tryptic soy broth (TSB) or peptone yeast glucose bile broth (PYGB) and grown in aerobic or anaerobic conditions, respectively, at 37°C for 48 hours. Microbial growth was pelleted and an aliquot was preserved for PCR

detection of *clbB* and *E. coli* (TSB culture) and *bft* and *B. fragilis* (PYGB culture). The remaining pellet was diluted and plated on semi-selective agar for single colony identification; aerobic TSB culture was plated on MacConkey plates and anaerobic PYGB culture was plated on BBE plates. A total of fifty Lac + or bile-esculin + colonies were selected for PCR from each sample. PCR detection was performed using *clbB* primers (Forward: GCA ACA TAC TCG CCC AGA CT, Reverse: TCT CAA GGC GTT GTT GTT TG) or *bft* primers (Forward: GCG AAC TCG GTT TAT GCA GT, Reverse: GTT GTA GAC ATC CCA CTG GC).

Mouse Experiments

Specific pathogen free (SPF) C57BL/6J (Jackson Laboratories, Bar Harbor, ME) mice were utilized for colonization experiments. Four-week-old mice were given water containing 500 mg/L cefoxitin for 48 hours, and inoculated by oral gavage with 10^8 ETBF, 10^8 *pks*+ *E. coli*, or a mixture containing 10^8 of each strain 24 hours after antibiotic water was removed. Colonization was confirmed by collection and cultivation of stool on selective media (MacConkey plates or BHI plates (with 10μg/ml clindamycin and 200 μg/ml gentamicin)) 48 hours after inoculation.

Immunohistochemistry

Formalin-fixed (10%), paraffin embedded tissues were sectioned (5μm) and stained. Slides were de-paraffinized and rehydrated following standard procedures. Slides were steamed in citrate buffer for 45 minutes, and allowed to cool to room temperature, followed by blocking of endogenous peroxidase activity for 10 minutes. Slides were blocked for 30 minutes in 10% normal goat serum, followed by primary antibody application overnight (1/500 rabbit anti-γH2AX [Bethyl Laboratories, IHC00008]). Slides were incubated with HRP for 30

minutes followed by DAB chromogen for 10 minutes. All sections were counterstained with hematoxylin prior to mounting.

Immunohistochemistry quantification

Nine crypts were selected for γH2AX quantification; three each from the proximal, middle, and distal mouse colon. Positive cells (containing 3 or more nuclear foci) were counted along with total number of cells in the crypt and a resulting percentage was determined. Cells were counted in a blinded manner.

Quantitative real-time PCR

An approximately 200 mg segment of distal mouse colon was processed for RNA isolation immediately following removal of the colon. Tissue was homogenized by bead beating in buffer ALS (Qiagen) and then run through a tissuelyzer column (Qiagen). The resultant solution was utilized for RNA extraction with RNeasy kit according to the manufacturer's recommended procedures. Transcription to complementary DNA was carried out using superscript III (Invitrogen). All qPCRs were carried out in triplicate with TaqMan primer/probes for IL-17a, IL-6, IFNγ, Nos2 and18s (as reference gene), and TaqMan 2x mastermix (Applied Biosystems). PCR conditions were for 48°C for 30 minutes, 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. The level of target mRNA was determined by the delta delta CT method.

Figure 4.1 16s sequence analysis of FAP patients. (**A**) Histogram of bacterial species detected on mucosal tissues and stool from 4 FAP patients harboring three different germline mutations. Asterisks denote stool samples collected from the left or right colon; bacterial stool populations are dynamic sharing more similarities with the local mucosa than stool in a different location along the colonic axis. *Fusobacterium necrophorum* was the only *Fusobacterium* species detected and was identified on only one of the four patients (**B, C and D**) Relative abundance of *Bacteroides*, *Prevotella steracorrera* and *Gammaproteobacteria* on FAP mucosa (n=4 patients) compared with mucosa of healthy subjects (n=10 subjects) and sporadic CRC (n=25 patients). (**B**) Trend of *Bacteroides* enrichment on FAP polyp mucosa, though not significant, when compared to Sporadic CRC (p=0.09) and healthy subjects (p=0.11). (**C**) *Prevotella stercorea* on FAP polyp mucosa compared with sporadic CRC (p=0.027) and healthy subjects (p=0.0012). (**D**) *Gammaproteobacteria* population detected by sequencing compared with sporadic CRC (p=0.029) and healthy subjects (p=0.016). Significance calculated with the nonparametric Mann-Whitney test.

A.
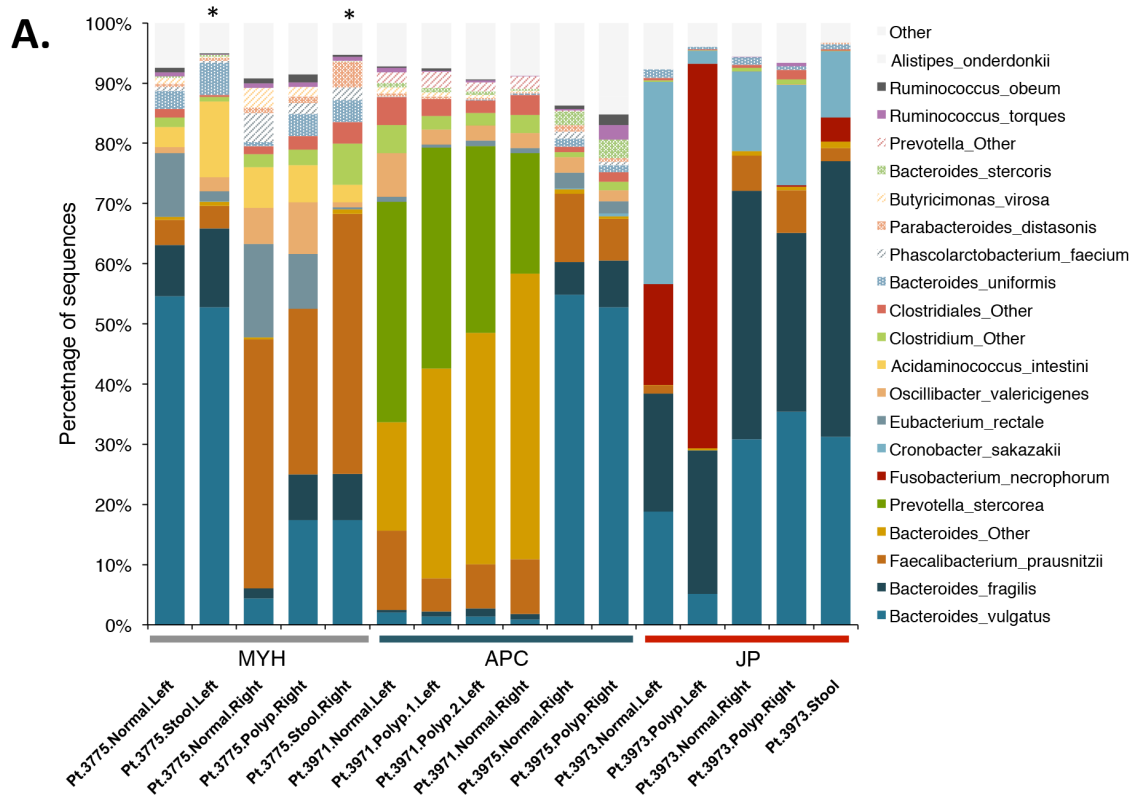
B. *Bacteroides*
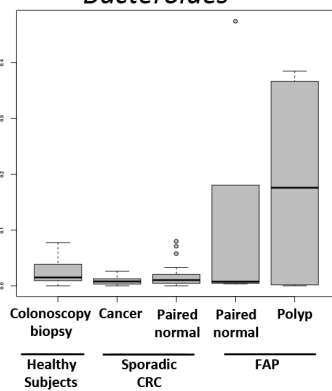
C. *Prevotella stercorea*
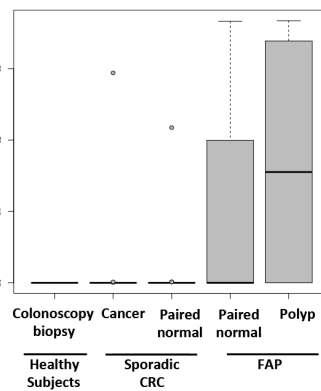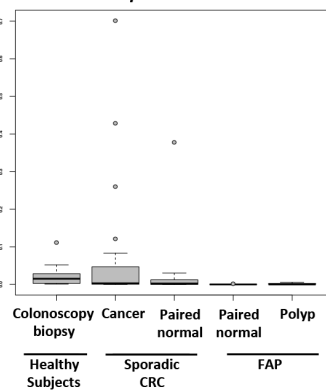
D. *Gammaproteobacteria*

Figure 4.2. FISH and microbiology culture analysis of FAP mucosal tissues. (**A**) FISH of all bacteria (red) biofilms on the mucosal surface of FAP polyp and paired normal tissues counterstained with DAPI nuclear stain (blue). The majority of the biofilm composition (**B**) was identified as *B. fragilis* (green) and *E. coli* (red) using species-specific probes. Images obtained at 40x magnification, scale bars 50 μm. (**C**) *Enterobacteriaceae* (yellow) and *E. coli* (red) FISH probes on paired normal FAP tissue (100x) revealing invasion of epithelial cells at the base of a crypt. Right panels with insets of Enterobacteriaceae (top right panel) in yellow, E. coli (middle right panel) in red and overlay (bottom right panel) confirming identification of the invasive species. Scale bar represents 20 μm. Semi-quantitative culture analysis (**D**) *E. coli* and (**E**) *B. fragilis* mucosal populations on mucosa of FAP patients (n=6 patients, 14 tissue specimens) compared to control subjects (n=11 subjects, 11 tissue specimens). Data displayed as colony forming units (CFU) per mm^2 of tissue cultured and significance calculated using non-parametric Mann-Whitney test.

A. Polyp | Paired Normal

B.

C.

D.

E.

Figure 4.3 Acute colitis induced by *pks+E.coli* and ETBF. (**A**) γ-H2AX IHC of DNA damage in the proximal and distal colon of wild-type mice infected with *pks+E.coli* (left panels) ETBF (middle panels) or co-infected with *pks+E.coli* and ETBF. *pks+E.coli* DNA damage is detected mostly at the differentiated surface epithelium, ETBF-induced DNA damage is mostly detected at the base of the crypt, while co-colonized mice have significantly increased DNA damage throughout the crypt along the entire colonic axis. (**B**) Quantification of DNA damage displayed as percentage of positive cells per crypt (n=6 mice per group). (**C** and **D**) Quantitative PCR of inflammatory cytokines in distal colon of singly and co-infected mice, data displayed as fold increase above sham (n=5 mice per group). (**C**) iNOS expression in the distal colon of co-infected mice when compared with mice singly infected with *pks+E.coli*. (**D**) IL-1 beta expression in the distal colon of co-infected mice compared with mice singly infected with ETBF or *pks+E.coli*. p-values were calculated with the nonparametric Mann-Whitney test.
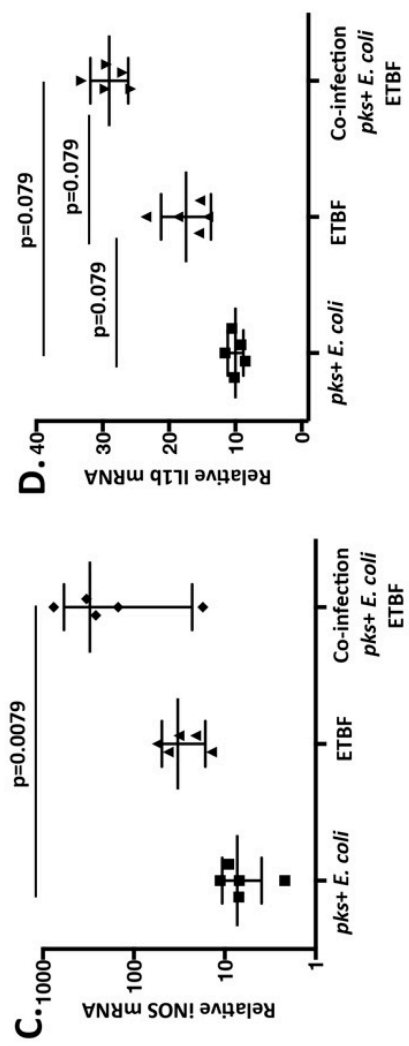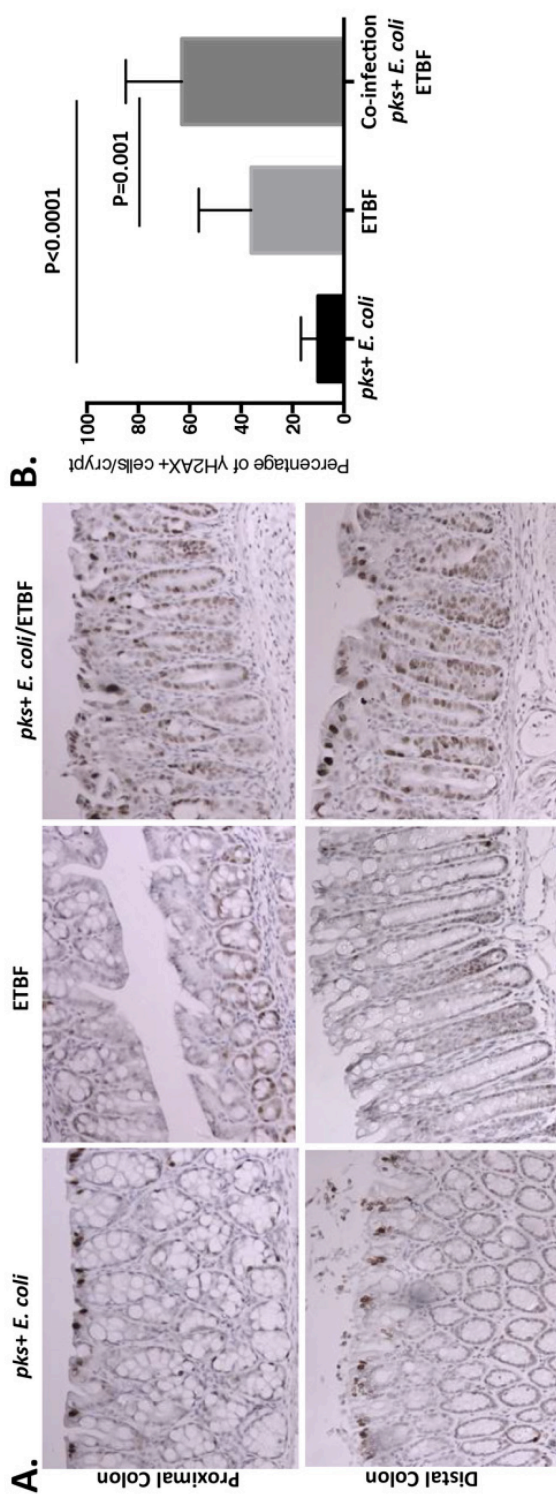
Table 4.1 FISH analysis

Patient ID	Mutation	Tissue samples collected	Tissue samples containing a biofilm	Biofilm location	E. coli (% of biofilm)	B. fragilis (% of biofilm)
3775*		23	16, 69%	3 Right (1 polyp, 2 paired normal), 13 Left (1 polyps, 12 paired normal)	49%	9%
3971*[a]		22	14, 61%	4 Right (1 polyp, 3 paired normal), 11 Left (2 polyps, 9 paired normal)	51%	29%
3975		7	5, 71%	1 Right (polyp), 4 Left (2 polyps, 2 paired normal)	66%	7%
3995		6	6. 100%	3 Right (1 polyp, 2 paired normal, 3 Left (2 polyps, 1 paired normal)	59%	5%
3973		8	0	NA	NA	NA
3983[b]		4	0	NA	NA	NA

*attenuated phenotype
a Juvenille polyposis
b Patient received oral antibiotics 24 hours prior to surgery

Table 4.2 Presence of *pks+ E. coli* and ETBF in FAP patients and Controls

	Total	*pks+ E. coli*	ETBF+	*pks+ E. coli*, ETBF+	*pks- E. coli*, ETBF-
FAP	25	17	15	13	6
Controls	23	5	7	5	16
p value		0.0017	0.0487	0.0401	0.0033
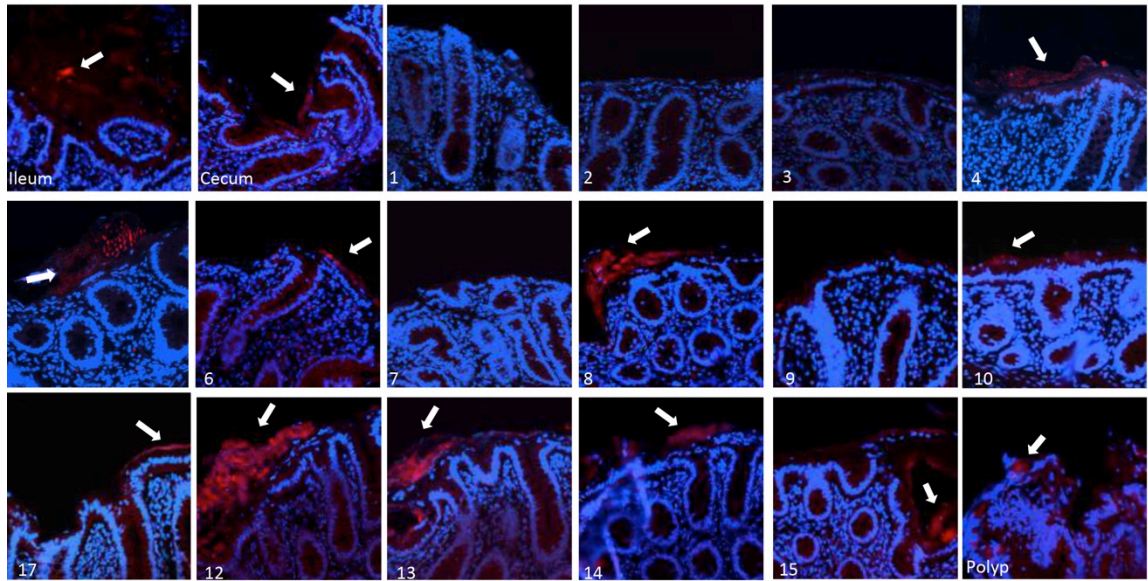
Figure S4.1. FISH of all bacteria (red) on colon specimens collected from an individual with

FAP. Tissue specimens were collected approximately every 3-5 centimeters starting in the

right colon (sample 1) and ending in the rectum (sample 15). Patchy biofilms detected

throughout the colon on both polyp and grossly normal tissues.

Figure S4.2 Biofilm characterization of FAP colons. Specimens from six prospectively collected FAP colons were available for FISH analysis. Four individuals (3775, 3975, 3995 and 3971) contained a bioifim, while the JP and antibiotic treated patient had no biofilms. FISH of all bacteria top panels, displaying representative biofilms from each patient and species specific FISH probes (below) of *E. coli* and *B. fragilis* biofilm composition from each patient. Bottom panel displays colon specimen sites with biofilm designations (red=biofilm, blue=no biofilm).

3973 — Juvenille polyposis

3983 — Oral antibiotics APC mutation

3971 — Attenuated phenotype

3995 — APC Mutation

3975 — APC Mutation

3775 — Attenuated phenotype MYH Mutation

Biofilm positive
Biofilm Negative
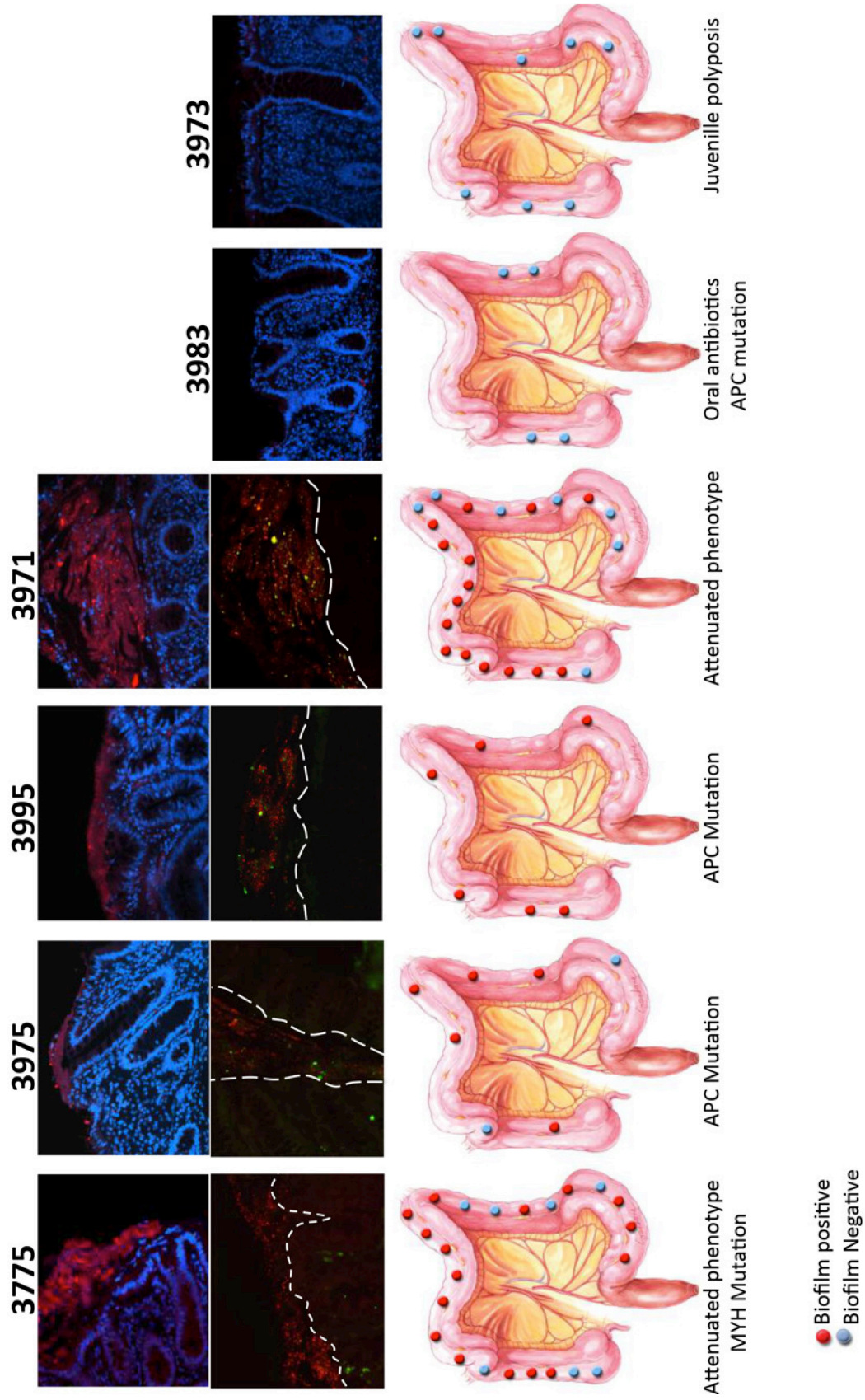
Figure S4.3 Histopathology of acute colitis induced by *pks+E.coli* isolates. H&E stained distal colon of wild-type mice seven days post-inoculation. Variable colitis was detected among the strains with increasing pathology from NC101< isolate1< isolate2< isolate3. A sham colon is displayed in the far left panel for comparison.

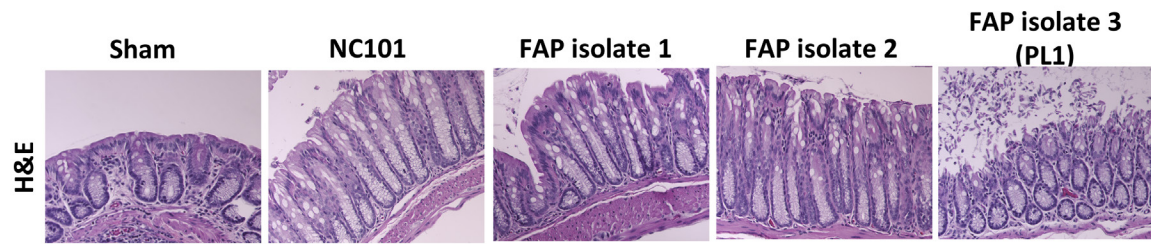| Supplementary Table 4.1 FAP patient metadata						
Patient ID	Age	Sex	Race	Operation	Mutation	Analyses
365	39	M	Caucasian	Procolectomy	ND	Micro
1420	16	F	Caucasian	Colectomy	ND	Micro
1679	27	F	Caucasian	Colectomy	APC	Micro
2017	8	F	Caucasian	Procolectomy	ND	Micro
2215	16	M	Caucasian	Procolectomy	ND	Micro
2544	7	F	Caucasian	Procolectomy	ND	Micro
2732	25	F	Caucasian	Procolectomy	ND	Micro
2735	35	M	Caucasian	Procolectomy	ND	Micro
2891	65	F	African American	Whipple	ND	Micro
2904	18	F	Caucasian	Procolectomy	ND	Micro
2927	22	M	Caucasian	Procolectomy	ND	Micro
3024	53	M	Caucasian	Procolectomy	ND	Micro
3026	38	M	Caucasian	Procolectomy	ND	Micro
3037	9	M	Caucasian	Procolectomy	ND	Micro
3166	43	F	Caucasian	Procolectomy	ND	Micro
3174	17	F	Caucasian	Colectomy	ND	Micro
3233	27	M	African American	Procolectomy	ND	Micro
3345	35	F	Caucasian	Colectomy	ND	Micro
3381	25	M	Caucasian	Procolectomy	ND	Micro
3775*	51	F	Caucasian	Colectomy	MYH	Seq, FISH, Micro
3971*[a]	27	M	African American	Procolectomy	ND	Seq, FISH, Micro
3973	27	F	Caucasian	Colectomy	ND	Seq, FISH, Micro
3975	50	M	Caucasian	Colectomy	APC	Seq, FISH, Micro
3983[b]	51	F	African American	Colectomy	APC	FISH, Micro
3995	39	M	Caucasian	Procolectomy	APC	FISH, Micro

Abbreviations: ND, not determined;

*attenuated phenotype
a Juvenille polyposis
b Patient received oral antibiotics 24 hours prior to surgery

Supplementary Table 4.2 Control Subjects

Patient ID	Age	Sex	Race	Analyses
3714	33	F	Caucasian	Micro
3723	49	F	Caucasian	Micro
3724	66	M	Hispanic	Micro
3730	61	F	Caucasian	Micro
3734	57	F	Caucasian	Micro
3737	52	F	Caucasian	Micro
S55	60	F	African American	Micro
S56	52	F	African American	Micro
S57	52	M	Caucasian	Micro
S58	57	F	African American	Micro
S59	77	M	African American	Micro
S60	45	F	Caucasian	Micro
S61	58	M	Caucasian	Micro
S62	64	M	African American	Micro
S63	74	F	African American	Micro
S64	61	M	African American	Micro
S65	67	F	African American	Micro
S66	49	F	African American	Micro
S67	67	F	Caucasian	Micro
S68	47	F	African American	Micro
S69	57	M	Caucasian	Micro
S70	59	F	Caucasian	Micro
S71	52	F	African American	Micro

Supplementary Table 4.3 FISH probes

Probe target(s)	Probe Name	Fluorophore	Probe Sequence (5'-3')
Kingdom Bacteria except Planctomycetales and Verrucomicrobia	Eub338	Cy3, Alexa 405	GCTGCCTCCCGTAGGAGT
Fusobacteria	Fus714	Alexa 488	GGCTTCCCCATCGGCATT
Prevotella, Bacteroides	PRV392	Rhodamine Red X	GCACGCTACTTGGCTGG
Bacteroidetes (Bacteroides, Parabacteroides, Prevotella)	CFB286	Alexa 514	TCCTCTCAGAACCCCTAC
Betaproteobacteria	Bet42a	Alexa 647	GCCTTCCCACTTCGTTT
Gammaproteobacteria	Gam42a	Alexa 647	GCCTTCCCACATCGTTT
Lachnospiraceae	Lac435	Texas Red X	TCTTCCCTGCTGATAGA
Enterobacteriaceae except gut Proteus spp	Ent186	Alexa 555	CCCCCWCTTTGGTCTTGC
Bacteroides fragilis	S-S-Bfrag-998-a-A-20	Alexa 633	GTTTCCACATCATTCCACTG
Escherichia coli	Eco1531	HRP-tyramide488	CACCGTAGTGCCTGGTCATCA
Escherichia coli	Eco1161	HRP-tyramide488	GCATAAGCGTCGCTGCCG

Chapter 5

General Discussion

Sporadic colorectal cancer is ultimately a genetic disease, where gene alterations and chromosomal instability are central to the stepwise progression towards neoplasia [4,5]. This complex process is undoubtedly the result of numerous influences ranging from age, gender, nutritional intake, physical activity and host genetic background, to the diverse and variable colonic microbiome. Epidemiological and experimental evidence discussed here strongly suggest a role for several bacterial agents in CRC. However, traditional bacteriological approaches are built on the assumption that an etiologic pathogen can be isolated, cultured and identified, and that pathogenesis can be explained through confirmation of disease. Throughout the 19th century and beyond, these concepts, grounded in Koch's postulates, have proven to be crucial in the identification of countless infectious pathogens, including the etiologic agent of gastric cancer, *H. pylori* [197]. Yet unlike the archetypal infectious disease consisting of a single causative agent, the colon houses a variety of commensal organisms, many of which have been implicated, both alone and in consort, to contribute to the genesis of colon cancer. The challenge of traditional epidemiological approaches to identify links between bacterial agents and CRC is further hampered by the long length of time between initiation and detectable carcinogenesis. Searching for the responsible agent(s) among the multiple constituents of the colonic flora presents a challenging prospect, since it is possible that the critical inciting microbial agent or composition is no longer present at the time of disease discovery. As such, we are then potentially reliant on detection of an immune signature to the microbe or microbiota to provide the epidemiologic link to CRC. The work presented in this thesis supports the potential for bacterial alpha bugs or drivers in the context of the aggregate flora to shape the microbial community yielding a procarcinogenic environment [195,196]. This emphasizes the need for detailed knowledge about specific microbes as well as alterations of whole microbial communities under diseased

and healthy states to better understand the etiology of CRC. The advent of next generation sequencing technologies have facilitated these types of studies that can take into account the community of a specimen, many of which were discussed here (**CHAPTER 1**, Table 1.2). However, limitations in the experimental evidence to date include small sampling numbers and limited control populations for comparison. Furthermore, information regarding host genetics is also necessary; as revealed by numerous mouse studies, commensal bacteria have pathogenic capabilities in the context of genetic abnormalities in the host.

While advances have been made in the early stages of characterizing what species are present on tumors and their flanking tissues, the data presented here is the first attempt to determine the spatial organization of those microbes with respect to the host epithelium. The spatial arrangement of the bacterial community dictates both microbe-microbe interactions and microbe-host interactions. Proximity to the host epithelium facilitates the way in which microbes are recognized and responded to by the host innate and adaptive immune system [198]. Our systematic studies of the distribution of microbes along both the length of the colon, as well as a cross-sectional characterization of the lumen and mucus layer members are essential to further elucidate the role of specific bacterial community members in the cancerous disease state.

As the field moves forward, several types of evidence will be needed to link the microbiota to human CRC [199]. Prospectively conducted studies, initiated at a time point before the onset of disease, and with relevant samples (blood, tissue and stool) for analysis would be ideal. Capturing information about the microbiome structure and composition in the early stages of disease initiation and throughout disease development would be invaluable. Ideally, the detection of microbiome dysbiosis or exposure to specific putative etiologic agents before disease development would help to address the cause or consequence

conundrum. However, population-based microbiome studies are both cost-prohibitive and impractical for evaluating long-term (20-40 years in the case of colon cancer) disease development. Attention to designing control groups and using varied controls is important as well to help determine if a microbe or a microbiota composition exhibits a strong, consistent association with human CRC. We should seek to detect an immunologic response to the purported microbial etiologies of CRC. It was the combined criteria of either detection of *H pylori* or an immune response to *H pylori* that provided crucial data to define *H pylori* as the cause of most gastric cancer [197]. Murine models of colon oncogenesis will likely provide key insights into molecules and mediators with translational importance to understanding how the microbiota contributes to human CRC. Ultimately, elimination of the inciting microbe or restructuring of the microbiome whether by diet, probiotics, antibiotics or vaccination with subsequent prevention of CRC is required for definitive declaration of disease association. While these criteria are stringent and create a necessarily high bar for investigators to reach, there has never been more interest in understanding the microbial inciters of human CRC.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D: Global cancer statistics. *CA Cancer. J. Clin.* 61(2), 69-90 (2011).

2. Center MM, Jemal A, Ward E: International trends in colorectal cancer incidence rates. *Cancer Epidemiol. Biomarkers Prev.* 18(6), 1688-1694 (2009).

3. Center MM, Jemal A, Smith RA, Ward E: Worldwide variations in colorectal cancer. *CA Cancer. J. Clin.* 59(6), 366-378 (2009).

4. Fearon ER, Vogelstein B: A genetic model for colorectal tumorigenesis. *Cell* 61(5), 759-767 (1990).

5. Vogelstein B, Kinzler KW: The multistep nature of cancer. *Trends Genet.* 9(4), 138-141 (1993).

6. Eckburg PB, Bik EM, Bernstein CN *et al.*: Diversity of the human intestinal microbial flora. *Science* 308(5728), 1635-1638 (2005).

7. Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL: An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* 122(1), 107-118 (2005).

8. Reikvam DH, Erofeev A, Sandvik A *et al.*: Depletion of murine intestinal microbiota: effects on gut mucosa and epithelial gene expression. *PLoS One* 6(3), e17996 (2011).

9. Kleessen B, Kroesen AJ, Buhr HJ, Blaut M: Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. *Scand. J. Gastroenterol.* 37(9), 1034-1041 (2002).

10. Rowland IR: The role of the gastrointestinal microbiota in colorectal cancer. *Curr. Pharm. Des.* 15(13), 1524-1527 (2009).

11. Sheng YH, Hasnain SZ, Florin TH, McGuckin MA: Mucins in inflammatory bowel diseases and colorectal cancer. *J. Gastroenterol. Hepatol.* 27(1), 28-38 (2012).

12. McGuckin MA, Linden SK, Sutton P, Florin TH: Mucin dynamics and enteric pathogens. *Nat. Rev. Microbiol.* 9(4), 265-278 (2011).

13. Johansson ME, Larsson JM, Hansson GC: The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc. Natl. Acad. Sci. U. S. A.* 108 Suppl 1, 4659-4665 (2011).

14. Matsuo K, Ota H, Akamatsu T, Sugiyama A, Katsuyama T: Histochemistry of the surface mucous gel layer of the human colon. *Gut* 40(6), 782-789 (1997).

15. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI: The human microbiome project. *Nature* 449(7164), 804-810 (2007).

16. Ahmed S, Macfarlane GT, Fite A, McBain AJ, Gilbert P, Macfarlane S: Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. *Appl. Environ. Microbiol.* 73(22), 7435-7442 (2007).

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

18. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR: Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. U. S. A.* 104(34), 13780-13785 (2007).

19. Swidsinski A, Loening-Baucke V, Herber A: Mucosal flora in Crohn's disease and ulcerative colitis - an overview. *J. Physiol. Pharmacol.* 60 Suppl 6, 61-71 (2009).

20. Tarmin L, Yin J, Harpaz N *et al.*: Adenomatous polyposis coli gene mutations in ulcerative colitis-associated dysplasias and cancers versus sporadic colon neoplasms. *Cancer Res.* 55(10), 2035-2038 (1995).

21. Xie J, Itzkowitz SH: Cancer in inflammatory bowel disease. *World J. Gastroenterol.* 14(3), 378-389 (2008).

22. Marchesi JR, Dutilh BE, Hall N *et al.*: Towards the human colorectal cancer microbiome. *PLoS One* 6(5), e20447 (2011).

23. Kostic AD, Gevers D, Pedamallu CS *et al.*: Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Res.* 22(2), 292-298 (2012).

24. Balkwill F, Mantovani A: Inflammation and cancer: back to Virchow? *Lancet* 357(9255), 539-545 (2001).

25. Grivennikov SI, Greten FR, Karin M: Immunity, inflammation, and cancer. *Cell* 140(6), 883-899 (2010).

26. Coussens LM, Werb Z: Inflammation and cancer. *Nature* 420(6917), 860-867 (2002).

27. Karin M, Greten FR: NF-kappaB: linking inflammation and immunity to cancer development and progression. *Nat. Rev. Immunol.* 5(10), 749-759 (2005).

28. Nath G, Gulati AK, Shukla VK: Role of bacteria in carcinogenesis, with special reference to carcinoma of the gallbladder. *World J. Gastroenterol.* 16(43), 5395-5404 (2010).

29. Grivennikov S, Karin E, Terzic J *et al.*: IL-6 and Stat3 are required for survival of intestinal epithelial cells and development of colitis-associated cancer. *Cancer. Cell.* 15(2), 103-113 (2009).

30. Bollrath J, Phesse TJ, von Burstin VA *et al.*: gp130-mediated Stat3 activation in enterocytes regulates cell survival and cell-cycle progression during colitis-associated tumorigenesis. *Cancer. Cell.* 15(2), 91-102 (2009).

31. Li Y, de Haar C, Chen M *et al.*: Disease-related expression of the IL6/STAT3/SOCS3 signalling pathway in ulcerative colitis and ulcerative colitis-related carcinogenesis. *Gut* 59(2), 227-235 (2010).

32. Goodwin AC, Destefano Shields CE, Wu S *et al.*: Polyamine catabolism contributes to enterotoxigenic Bacteroides fragilis-induced colon tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* 108(37), 15354-15359 (2011).

33. Waris G, Ahsan H: Reactive oxygen species: role in the development of cancer and various chronic conditions. *J. Carcinog.* 5, 14 (2006).

34. Lax AJ: Opinion: Bacterial toxins and cancer--a case to answer? *Nat. Rev. Microbiol.* 3(4), 343-349 (2005).

35. Huang JQ, Zheng GF, Sumanac K, Irvine EJ, Hunt RH: Meta-analysis of the relationship between cagA seropositivity and gastric cancer. *Gastroenterology* 125(6), 1636-1644 (2003).

36. Peek RM,Jr, Blaser MJ: Helicobacter pylori and gastrointestinal tract adenocarcinomas. *Nat. Rev. Cancer.* 2(1), 28-37 (2002).

37. Hatakeyama M, Higashi H: Helicobacter pylori CagA: a new paradigm for bacterial carcinogenesis. *Cancer. Sci.* 96(12), 835-843 (2005).

38. Brandt S, Kwok T, Hartig R, Konig W, Backert S: NF-kappaB activation and potentiation of proinflammatory responses by the Helicobacter pylori CagA protein. *Proc. Natl. Acad. Sci. U. S. A.* 102(26), 9300-9305 (2005).

39. Lamb A, Yang XD, Tsang YH *et al.*: Helicobacter pylori CagA activates NF-kappaB by targeting TAK1 for TRAF6-mediated Lys 63 ubiquitination. *EMBO Rep.* 10(11), 1242-1249 (2009).

40. Franco AT, Israel DA, Washington MK *et al.*: Activation of beta-catenin by carcinogenic Helicobacter pylori. *Proc. Natl. Acad. Sci. U. S. A.* 102(30), 10646-10651 (2005).

41. Buti L, Spooner E, Van der Veen AG, Rappuoli R, Covacci A, Ploegh HL: Helicobacter pylori cytotoxin-associated gene A (CagA) subverts the apoptosis-stimulating protein of p53 (ASPP2) tumor suppressor pathway of the host. *Proc. Natl. Acad. Sci. U. S. A.* 108(22), 9238-9243 (2011).

42. Miehlke S, Kirsch C, Agha-Amiri K *et al.*: The Helicobacter pylori vacA s1, m1 genotype and cagA is associated with gastric carcinoma in Germany. *Int. J. Cancer* 87(3), 322-327 (2000).

43. Jinadasa RN, Bloom SE, Weiss RS, Duhamel GE: Cytolethal distending toxin: a conserved bacterial genotoxin that blocks cell cycle progression, leading to apoptosis of a broad range of mammalian cell lineages. *Microbiology* 157(Pt 7), 1851-1875 (2011).

44. Elwell C, Chao K, Patel K, Dreyfus L: Escherichia coli CdtB mediates cytolethal distending toxin cell cycle arrest. *Infect. Immun.* 69(5), 3418-3422 (2001).

45. Nougayrede JP, Homburg S, Taieb F *et al.*: Escherichia coli induces DNA double-strand breaks in eukaryotic cells. *Science* 313(5788), 848-851 (2006).

46. Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrede JP: Escherichia coli induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 107(25), 11537-11542 (2010).

47. Arthur JC, Perez-Chanona E, Muhlbauer M *et al.*: Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* 338(6103), 120-123 (2012).

48. Arthur JC, Gharaibeh RZ, Muhlbauer M *et al.*: Microbial genomic analysis reveals the essential role of inflammation in bacteria-induced colorectal cancer. *Nat. Commun.* 5, 4724 (2014).

49. Moore WE, Holdeman LV: Human fecal flora: the normal flora of 20 Japanese-Hawaiians. *Appl. Microbiol.* 27(5), 961-979 (1974).

50. Polk BF, Kasper DL: Bacteroides fragilis subspecies in clinical isolates. *Ann. Intern. Med.* 86(5), 569-571 (1977).

51. Wu S, Rhee KJ, Albesiano E *et al.*: A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* 15(9), 1016-1022 (2009).

52. Rhee KJ, Wu S, Wu X *et al.*: Induction of persistent colitis by a human commensal, enterotoxigenic Bacteroides fragilis, in wild-type C57BL/6 mice. *Infect. Immun.* 77(4), 1708-1718 (2009).

53. Basset C, Holton J, Bazeos A, Vaira D, Bloom S: Are Helicobacter species and enterotoxigenic Bacteroides fragilis involved in inflammatory bowel disease? *Dig. Dis. Sci.* 49(9), 1425-1432 (2004).

54. Prindiville TP, Sheikh RA, Cohen SH, Tang YJ, Cantrell MC, Silva J,Jr: Bacteroides fragilis enterotoxin gene sequences in patients with inflammatory bowel disease. *Emerg. Infect. Dis.* 6(2), 171-174 (2000).

55. Sears CL: Enterotoxigenic Bacteroides fragilis: a rogue among symbiotes. *Clin. Microbiol. Rev.* 22(2), 349-69, Table of Contents (2009).

56. Toprak NU, Yagci A, Gulluoglu BM *et al.*: A possible role of Bacteroides fragilis enterotoxin in the aetiology of colorectal cancer. *Clin. Microbiol. Infect.* 12(8), 782-786 (2006).

57. Zitomersky NL, Coyne MJ, Comstock LE: Longitudinal analysis of the prevalence, maintenance, and IgA response to species of the order Bacteroidales in the human gut. *Infect. Immun.* 79(5), 2012-2020 (2011).

58. Franco AA, Mundy LM, Trucksis M, Wu S, Kaper JB, Sears CL: Cloning and characterization of the Bacteroides fragilis metalloprotease toxin gene. *Infect. Immun.* 65(3), 1007-1013 (1997).

59. Franco AA: The Bacteroides fragilis pathogenicity island is contained in a putative novel conjugative transposon. *J. Bacteriol.* 186(18), 6077-6092 (2004).

60. Wu S, Lim KC, Huang J, Saidi RF, Sears CL: Bacteroides fragilis enterotoxin cleaves the zonula adherens protein, E-cadherin. *Proc. Natl. Acad. Sci. U. S. A.* 95(25), 14979-14984 (1998).

61. Wu S, Morin PJ, Maouyo D, Sears CL: Bacteroides fragilis enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology* 124(2), 392-400 (2003).

62. Wu S, Powell J, Mathioudakis N, Kane S, Fernandez E, Sears CL: Bacteroides fragilis enterotoxin induces intestinal epithelial cell secretion of interleukin-8 through mitogen-activated protein kinases and a tyrosine kinase-regulated nuclear factor-kappaB pathway. *Infect. Immun.* 72(10), 5832-5839 (2004).

63. Su LK, Kinzler KW, Vogelstein B *et al.*: Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science* 256(5057), 668-670 (1992).

64. Fodde R, Edelmann W, Yang K *et al.*: A targeted chain-termination mutation in the mouse Apc gene results in multiple intestinal tumors. *Proc. Natl. Acad. Sci. U. S. A.* 91(19), 8969-8973 (1994).

65. Cancer Genome Atlas Network: Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487(7407), 330-337 (2012).

66. Liu J, Duan Y, Cheng X *et al.*: IL-17 is associated with poor prognosis and promotes angiogenesis via stimulating VEGF production of cancer cells in colorectal carcinoma. *Biochem. Biophys. Res. Commun.* 407(2), 348-354 (2011).

67. Chae WJ, Gibson TF, Zelterman D, Hao L, Henegariu O, Bothwell AL: Ablation of IL-17A abrogates progression of spontaneous intestinal tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* 107(12), 5540-5544 (2010).

68. Su X, Ye J, Hsueh EC, Zhang Y, Hoft DF, Peng G: Tumor microenvironments direct the recruitment and expansion of human Th17 cells. *J. Immunol.* 184(3), 1630-1641 (2010).

69. Sobhani I, Tap J, Roudot-Thoraval F *et al.*: Microbial dysbiosis in colorectal cancer (CRC) patients. *PLoS One* 6(1), e16393 (2011).

70. Morikawa T, Baba Y, Yamauchi M *et al.*: STAT3 expression, molecular features, inflammation patterns, and prognosis in a database of 724 colorectal cancers. *Clin. Cancer Res.* 17(6), 1452-1462 (2011).

71. Moser AR, Pitot HC, Dove WF: A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science* 247(4940), 322-324 (1990).

72. Oshima M, Oshima H, Kitagawa K, Kobayashi M, Itakura C, Taketo M: Loss of Apc heterozygosity and abnormal tissue building in nascent intestinal polyps in mice carrying a truncated Apc gene. *Proc. Natl. Acad. Sci. U. S. A.* 92(10), 4482-4486 (1995).

73. Velcich A, Yang W, Heyer J *et al.*: Colorectal cancer in mice genetically deficient in the mucin Muc2. *Science* 295(5560), 1726-1729 (2002).

74. Berg DJ, Davidson N, Kuhn R *et al.*: Enterocolitis and colon cancer in interleukin-10-deficient mice are associated with aberrant cytokine production and CD4(+) TH1-like responses. *J. Clin. Invest.* 98(4), 1010-1020 (1996).

75. Zhu Y, Richardson JA, Parada LF, Graff JM: Smad3 mutant mice develop metastatic colorectal cancer. *Cell* 94(6), 703-714 (1998).

76. Rudolph U, Finegold MJ, Rich SS *et al.*: Ulcerative colitis and adenocarcinoma of the colon in G alpha i2-deficient mice. *Nat. Genet.* 10(2), 143-150 (1995).

77. Takaku K, Oshima M, Miyoshi H, Matsui M, Seldin MF, Taketo MM: Intestinal tumorigenesis in compound mutant mice of both Dpc4 (Smad4) and Apc genes. *Cell* 92(5), 645-656 (1998).

78. Funabashi H, Uchida K, Kado S, Matsuoka Y, Ohwaki M: Establishment of a Tcrb and Trp53 genes deficient mouse strain as an animal model for spontaneous colorectal cancer. *Exp. Anim.* 50(1), 41-47 (2001).

79. Chu FF, Esworthy RS, Chu PG *et al.*: Bacteria-induced intestinal cancer in mice with disrupted Gpx1 and Gpx2 genes. *Cancer Res.* 64(3), 962-968 (2004).

80. Engle SJ, Ormsby I, Pawlowski S *et al.*: Elimination of colon cancer in germ-free transforming growth factor beta 1-deficient mice. *Cancer Res.* 62(22), 6362-6366 (2002).

81. Taketo MM, Edelmann W: Mouse models of colon cancer. *Gastroenterology* 136(3), 780-798 (2009).

82. Balish E, Warner T: Enterococcus faecalis induces inflammatory bowel disease in interleukin-10 knockout mice. *Am. J. Pathol.* 160(6), 2253-2257 (2002).

83. Dove WF, Clipson L, Gould KA *et al.*: Intestinal neoplasia in the ApcMin mouse: independence from the microbial and natural killer (beige locus) status. *Cancer Res.* 57(5), 812-814 (1997).

84. Li Y, Kundu P, Seow SW *et al.*: Gut microbiota accelerate tumor growth via c-jun and STAT3 phosphorylation in APCMin/+ mice. *Carcinogenesis* 33(6), 1231-1238 (2012).

85. Kado S, Uchida K, Funabashi H *et al.*: Intestinal microflora are necessary for development of spontaneous adenocarcinoma of the large intestine in T-cell receptor beta chain and p53 double-knockout mice. *Cancer Res.* 61(6), 2395-2398 (2001).

86. Uronis JM, Muhlbauer M, Herfarth HH, Rubinas TC, Jones GS, Jobin C: Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. *PLoS One* 4(6), e6026 (2009).

87. Johansson ME, Gustafsson JK, Sjoberg KE *et al.*: Bacteria penetrate the inner mucus layer before inflammation in the dextran sulfate colitis model. *PLoS One* 5(8), e12238 (2010).

88. Perse M, Cerar A: Dextran sodium sulphate colitis mouse model: traps and tricks. *J. Biomed. Biotechnol.* 2012, 718617 (2012).

89. Elinav E, Strowig T, Kau AL *et al.*: NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. *Cell* 145(5), 745-757 (2011).

90. Wu GD, Chen J, Hoffmann C *et al.*: Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334(6052), 105-108 (2011).

91. Garrett WS, Lord GM, Punit S *et al.*: Communicable ulcerative colitis induced by T-bet deficiency in the innate immune system. *Cell* 131(1), 33-45 (2007).

92. Vijay-Kumar M, Aitken JD, Carvalho FA *et al.*: Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science* 328(5975), 228-231 (2010).

93. Hans W, Scholmerich J, Gross V, Falk W: The role of the resident intestinal flora in acute and chronic dextran sulfate sodium-induced colitis in mice. *Eur. J. Gastroenterol. Hepatol.* 12(3), 267-273 (2000).

94. Rath HC, Schultz M, Freitag R *et al.*: Different subsets of enteric bacteria induce and perpetuate experimental colitis in rats and mice. *Infect. Immun.* 69(4), 2277-2285 (2001).

95. Fukuda M, Kanauchi O, Araki Y *et al.*: Prebiotic treatment of experimental colitis with germinated barley foodstuff: a comparison with probiotic or antibiotic treatment. *Int. J. Mol. Med.* 9(1), 65-70 (2002).

96. Kitajima S, Morimoto M, Sagara E, Shimizu C, Ikeda Y: Dextran sodium sulfate-induced colitis in germ-free IQI/Jic mice. *Exp. Anim.* 50(5), 387-395 (2001).

97. Ellmerich S, Scholler M, Duranton B *et al.*: Promotion of intestinal carcinogenesis by Streptococcus bovis. *Carcinogenesis* 21(4), 753-756 (2000).

98. Boleij A, Tjalsma H: Gut bacteria in health and disease: a survey on the interface between intestinal microbiology and colorectal cancer. *Biol. Rev. Camb. Philos. Soc.* 87(3), 701-730 (2012).

99. Fox JG, Ge Z, Whary MT, Erdman SE, Horwitz BH: Helicobacter hepaticus infection in mice: models for understanding lower bowel inflammation and cancer. *Mucosal Immunol.* 4(1), 22-30 (2011).

100. Erdman SE, Rao VP, Poutahidis T *et al.*: Nitric oxide and TNF-alpha trigger colonic inflammation and carcinogenesis in Helicobacter hepaticus-infected, Rag2-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* 106(4), 1027-1032 (2009).

101. Mangerich A, Knutson CG, Parry NM *et al.*: Infection-induced colitis in mice causes dynamic and tissue-specific changes in stress response and DNA damage leading to colon cancer. *Proc. Natl. Acad. Sci. U. S. A.* 109(27), E1820-9 (2012).

102. Balagopal A, Philp FH, Astemborski J *et al.*: Human immunodeficiency virus-related microbial translocation and progression of hepatitis C. *Gastroenterology* 135(1), 226-233 (2008).

103. Klatt NR, Funderburg NT, Brenchley JM: Microbial translocation, immune activation, and HIV disease. *Trends Microbiol.* (2012).

104. Huycke MM, Moore D, Joyce W *et al.*: Extracellular superoxide production by Enterococcus faecalis requires demethylmenaquinone and is attenuated by functional terminal quinol oxidases. *Mol. Microbiol.* 42(3), 729-740 (2001).

105. Wang X, Allen TD, May RJ, Lightfoot S, Houchen CW, Huycke MM: Enterococcus faecalis induces aneuploidy and tetraploidy in colonic epithelial cells through a bystander effect. *Cancer Res.* 68(23), 9909-9917 (2008).

106. Ruiz PA, Shkoda A, Kim SC, Sartor RB, Haller D: IL-10 gene-deficient mice lack TGF-beta/Smad signaling and fail to inhibit proinflammatory gene expression in intestinal epithelial cells after the colonization with colitogenic Enterococcus faecalis. *J. Immunol.* 174(5), 2990-2999 (2005).

107. Winters MD, Schlinke TL, Joyce WA, Glore SR, Huycke MM: Prospective case-cohort study of intestinal colonization with enterococci that produce extracellular superoxide and the risk for colorectal adenomas or cancer. *Am. J. Gastroenterol.* 93(12), 2491-2500 (1998).

108. Castellarin M, Warren RL, Freeman JD *et al.*: Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res.* 22(2), 299-306 (2012).

109. Kostic AD, Chun E, Robertson L *et al.*: Fusobacterium nucleatum Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment. *Cell. Host Microbe* 14(2), 207-215 (2013).

110. Ahn J, Sinha R, Pei Z *et al.*: Human gut microbiome and risk for colorectal cancer. *J. Natl. Cancer Inst.* 105(24), 1907-1911 (2013).

111. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW: Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. *Cell. Host Microbe* 14(2), 195-206 (2013).

112. Mundy R, MacDonald TT, Dougan G, Frankel G, Wiles S: Citrobacter rodentium of mice and man. *Cell. Microbiol.* 7(12), 1697-1706 (2005).

113. Barthold SW, Osbaldiston GW, Jonas AM: Dietary, bacterial, and host genetic interactions in the pathogenesis of transmissible murine colonic hyperplasia. *Lab. Anim. Sci.* 27(6), 938-945 (1977).

114. Newman JV, Kosaka T, Sheppard BJ, Fox JG, Schauer DB: Bacterial infection promotes colon tumorigenesis in Apc(Min/+) mice. *J. Infect. Dis.* 184(2), 227-230 (2001).

115. Maddocks OD, Short AJ, Donnenberg MS, Bader S, Harrison DJ: Attaching and effacing Escherichia coli downregulate DNA mismatch repair protein in vitro and are associated with colorectal adenocarcinomas in humans. *PLoS One* 4(5), e5517 (2009).

116. Keusch GT: Opportunistic infections in colon carcinoma. *Am. J. Clin. Nutr.* 27(12), 1481-1485 (1974).

117. Kok H, Jureen R, Soon CY, Tey BH: Colon cancer presenting as Streptococcus gallolyticus infective endocarditis. *Singapore Med. J.* 48(2), e43-5 (2007).

118. Klein RS, Recco RA, Catalano MT, Edberg SC, Casey JI, Steigbigel NH: Association of Streptococcus bovis with carcinoma of the colon. *N. Engl. J. Med.* 297(15), 800-802 (1977).

119. Boleij A, van Gelder MM, Swinkels DW, Tjalsma H: Clinical Importance of Streptococcus gallolyticus infection among colorectal cancer patients: systematic review and meta-analysis. *Clin. Infect. Dis.* 53(9), 870-878 (2011).

120. Boleij A, Muytjens CM, Bukhari SI *et al.*: Novel clues on the specific association of Streptococcus gallolyticus subsp gallolyticus with colorectal cancer. *J. Infect. Dis.* 203(8), 1101-1109 (2011).

121. Corredoira-Sanchez J, Garcia-Garrote F, Rabunal R *et al.*: Association between bacteremia due to Streptococcus gallolyticus subsp. gallolyticus (Streptococcus bovis I) and colorectal neoplasia: a case-control study. *Clin. Infect. Dis.* 55(4), 491-496 (2012).

122. Schlegel L, Grimont F, Ageron E, Grimont PA, Bouvet A: Reappraisal of the taxonomy of the Streptococcus bovis/Streptococcus equinus complex and related species: description of Streptococcus gallolyticus subsp. gallolyticus subsp. nov., S. gallolyticus subsp. macedonicus subsp. nov. and S. gallolyticus subsp. pasteurianus subsp. nov. *Int. J. Syst. Evol. Microbiol.* 53(Pt 3), 631-645 (2003).

123. Seder CW, Kramer M, Long G, Uzieblo MR, Shanley CJ, Bove P: Clostridium septicum aortitis: Report of two cases and review of the literature. *J. Vasc. Surg.* 49(5), 1304-1309 (2009).

124. Wang T, Cai G, Qiu Y *et al.*: Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME J.* 6(2), 320-329 (2012).

125. 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(6), e39743 (2012).

126. Swidsinski A, Khilkin M, Kerjaschki D *et al.*: Association between intraepithelial Escherichia coli and colorectal cancer. *Gastroenterology* 115(2), 281-286 (1998).

127. Cho JH: The genetics and immunopathogenesis of inflammatory bowel disease. *Nat. Rev. Immunol.* 8(6), 458-466 (2008).

128. Podolsky DK: Inflammatory bowel disease. *N. Engl. J. Med.* 347(6), 417-429 (2002).

129. Singh PK, Schaefer AL, Parsek MR, Moninger TO, Welsh MJ, Greenberg EP: Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* 407(6805), 762-764 (2000).

130. Swidsinski A, Weber J, Loening-Baucke V, Hale LP, Lochs H: Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J. Clin. Microbiol.* 43(7), 3380-3389 (2005).

131. Costerton JW, Stewart PS, Greenberg EP: Bacterial biofilms: a common cause of persistent infections. *Science* 284(5418), 1318-1322 (1999).

132. Hall-Stoodley L, Costerton JW, Stoodley P: Bacterial biofilms: from the natural environment to infectious diseases. *Nat. Rev. Microbiol.* 2(2), 95-108 (2004).

133. Johansson ME, Larsson JM, Hansson GC: The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc. Natl. Acad. Sci. U. S. A.* 108 Suppl 1, 4659-4665 (2011).

134. Swidsinski A, Loening-Baucke V, Theissig F *et al.*: Comparative study of the intestinal mucus barrier in normal and inflamed colon. *Gut* 56(3), 343-350 (2007).

135. McCoy AN, Araujo-Perez F, Azcarate-Peril A, Yeh JJ, Sandler RS, Keku TO: Fusobacterium is associated with colorectal adenomas. *PLoS One* 8(1), e53653 (2013).

136. Dejea C, Wick E, Sears CL: Bacterial oncogenesis in the colon. *Future Microbiol.* 8(4), 445-460 (2013).

137. Grivennikov SI, Wang K, Mucida D *et al.*: Adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature* 491(7423), 254-258 (2012).

138. Soler AP, Miller RD, Laughlin KV, Carp NZ, Klurfeld DM, Mullin JM: Increased tight junctional permeability is associated with the development of colon cancer. *Carcinogenesis* 20(8), 1425-1431 (1999).

139. Bromberg J, Wang TC: Inflammation and cancer: IL-6 and STAT3 complete the link. *Cancer. Cell.* 15(2), 79-80 (2009).

140. Wang L, Yi T, Kortylewski M, Pardoll DM, Zeng D, Yu H: IL-17 can promote tumor growth through an IL-6-Stat3 signaling pathway. *J. Exp. Med.* 206(7), 1457-1464 (2009).

141. Tosolini M, Kirilovsky A, Mlecnik B *et al.*: Clinical impact of different classes of infiltrating T cytotoxic and helper cells (Th1, th2, treg, th17) in patients with colorectal cancer. *Cancer Res.* 71(4), 1263-1271 (2011).

142. Colditz GA, Wolin KY, Gehlert S: Applying what we know to accelerate cancer prevention. *Sci. Transl. Med.* 4(127), 127rv4 (2012).

143. Key TJ, Schatzkin A, Willett WC, Allen NE, Spencer EA, Travis RC: Diet, nutrition and the prevention of cancer. *Public Health Nutr.* 7(1A), 187-200 (2004).

144. Swidsinski A, Weber J, Loening-Baucke V, Hale LP, Lochs H: Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J. Clin. Microbiol.* 43(7), 3380-3389 (2005).

145. Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA: Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* 56(6), 1919-1925 (1990).

146. Valm AM, Mark Welch JL, Rieken CW *et al.*: Systems-level analysis of microbial community organization through combinatorial labeling and spectral imaging. *Proc. Natl. Acad. Sci. U. S. A.* 108(10), 4152-4157 (2011).

147. Diaz PI, Chalmers NI, Rickard AH *et al.*: Molecular characterization of subject-specific oral microflora during initial colonization of enamel. *Appl. Environ. Microbiol.* 72(4), 2837-2848 (2006).

148. Weller R, Glockner FO, Amann R: 16S rRNA-targeted oligonucleotide probes for the in situ detection of members of the phylum Cytophaga-Flavobacterium-Bacteroides. *Syst. Appl. Microbiol.* 23(1), 107-114 (2000).

149. Kong Y, He M, McAlister T, Seviour R, Forster R: Quantitative fluorescence in situ hybridization of microbial communities in the rumens of cattle fed different diets. *Appl. Environ. Microbiol.* 76(20), 6933-6938 (2010).

150. Kempf VA, Trebesius K, Autenrieth IB: Fluorescent In situ hybridization allows rapid identification of microorganisms in blood cultures. *J. Clin. Microbiol.* 38(2), 830-838 (2000).

151. Neef A, Amann R, Schleifer K: Detection of Microbial Cells in Aerosols Using Nucleic Acid Probes. *Syst. Appl. Microbiol.* 18(1), 113-122 (1995).

152. Wallner G, Amann R, Beisker W: Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* 14(2), 136-143 (1993).

153. Rigottier-Gois L, Rochet V, Garrec N, Suau A, Dore J: Enumeration of Bacteroides species in human faeces by fluorescent in situ hybridisation combined with flow cytometry using 16S rRNA probes. *Syst. Appl. Microbiol.* 26(1), 110-118 (2003).

154. Paster BJ, Bartoszyk IM, Dewhirst FE: Identification of oral streptococci using PCR-based, reverse-capture, checkerboard hybridization. *Methods Cell Sci* 20, 223 (1998).

155. Manz W, Amann R, Ludwig W, Wagner M, Schleifer K: Phylogenetic Oligodeoxynucleotide Probes for the Major Subclasses of Proteobacteria: Problems and Solutions. *Syst. Appl. Microbiol.* 15(4), 593-600 (1992).

156. Kuczynski J, Stombaugh J, Walters WA, Gonzalez A, Caporaso JG, Knight R: Using QIIME to analyze 16S rRNA gene sequences from microbial communities. *Curr. Protoc. Microbiol.* Chapter 1, Unit 1E.5. (2012).

157. Caporaso JG, Kuczynski J, Stombaugh J *et al.*: QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7(5), 335-336 (2010).

158. Muegge BD, Kuczynski J, Knights D *et al.*: Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* 332(6032), 970-974 (2011).

159. Koren O, Goodrich JK, Cullender TC *et al.*: Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell* 150(3), 470-480 (2012).

160. Bragg L, Stone G, Imelfort M, Hugenholtz P, Tyson GW: Fast, accurate error-correction of amplicon pyrosequences using Acacia. *Nat. Methods* 9(5), 425-426 (2012).

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

162. Wang Q, Garrity GM, Tiedje JM, Cole JR: Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73(16), 5261-5267 (2007).

163. Angiuoli SV, Matalka M, Gussman A *et al.*: CloVR: a virtual machine for automated and portable sequence analysis from the desktop using cloud computing. *BMC Bioinformatics* 12, 356-2105-12-356 (2011).

164. Edgar RC: Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19), 2460-2461 (2010).

165. White JR, Navlakha S, Nagarajan N, Ghodsi MR, Kingsford C, Pop M: Alignment and clustering of phylogenetic markers--implications for microbial diversity studies. *BMC Bioinformatics* 11, 152-2105-11-152 (2010).

166. Wick EC, LeBlanc RE, Ortega G *et al.*: Shift from pStat6 to pStat3 predominance is associated with inflammatory bowel disease-associated dysplasia. *Inflamm. Bowel Dis.* 18(7), 1267-1274 (2012).

167. Flint N, Cove FL, Evans GS: A low-temperature method for the isolation of small-intestinal epithelium along the crypt-villus axis. *Biochem. J.* 280 (Pt 2)(Pt 2), 331-334 (1991).

168. Lozupone CA, Stombaugh J, Gonzalez A *et al.*: Meta-analyses of studies of the human microbiota. *Genome Res.* 23(10), 1704-1714 (2013).

169. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R: UniFrac: an effective distance metric for microbial community comparison. *ISME J.* 5(2), 169-172 (2011).

170. Sears CL, Garrett WS: Microbes, microbiota, and colon cancer. *Cell. Host Microbe* 15(3), 317-328 (2014).

171. Shah P, Swiatlo E: A multifaceted role for polyamines in bacterial pathogens. *Mol. Microbiol.* 68(1), 4-16 (2008).

172. Dejea CM, Wick EC, Hechenbleikner EM *et al.*: Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. U. S. A.* 111(51), 18321-18326 (2014).

173. Gerner EW, Meyskens FL,Jr: Polyamines and cancer: old molecules, new understanding. *Nat. Rev. Cancer.* 4(10), 781-792 (2004).

174. Gerner EW, Kurtts TA, Fuller DJ, Casero RA,Jr: Stress induction of the spermidine/spermine N1-acetyltransferase by a post-transcriptional mechanism in mammalian cells. *Biochem. J.* 294 (Pt 2)(Pt 2), 491-495 (1993).

175. Woolridge DP, Martinez JD, Stringer DE, Gerner EW: Characterization of a novel spermidine/spermine acetyltransferase, BltD, from Bacillus subtilis. *Biochem. J.* 340 (Pt 3)(Pt 3), 753-758 (1999).

176. Cook T, Roos D, Morada M *et al.*: Divergent polyamine metabolism in the Apicomplexa. *Microbiology* 153(Pt 4), 1123-1130 (2007).

177. Bacchi CJ, Rattendi D, Faciane E *et al.*: Polyamine metabolism in a member of the phylum Microspora (Encephalitozoon cuniculi): effects of polyamine analogues. *Microbiology* 150(Pt 5), 1215-1224 (2004).

178. Northen TR, Yanes O, Northen MT *et al.*: Clathrate nanostructures for mass spectrometry. *Nature* 449(7165), 1033-1036 (2007).

179. Calavia R, Annanouch FE, Correig X, Yanes O: Nanostructure Initiator Mass Spectrometry for tissue imaging in metabolomics: future prospects and perspectives. *J. Proteomics* 75(16), 5061-5068 (2012).

180. Dakubo GD, Jakupciak JP, Birch-Machin MA, Parr RL: Clinical implications and utility of field cancerization. *Cancer. Cell. Int.* 7, 2 (2007).

181. Campilongo R, Di Martino ML, Marcocci L *et al.*: Molecular and functional profiling of the polyamine content in enteroinvasive E. coli : looking into the gap between commensal E. coli and harmful Shigella. *PLoS One* 9(9), e106589 (2014).

182. Soda K: The mechanisms by which polyamines accelerate tumor spread. *J. Exp. Clin. Cancer Res.* 30, 95-9966-30-95 (2011).

183. Patel CN, Wortham BW, Lines JL, Fetherston JD, Perry RD, Oliveira MA: Polyamines are essential for the formation of plague biofilm. *J. Bacteriol.* 188(7), 2355-2363 (2006).

184. Louis P, Hold GL, Flint HJ: The gut microbiota, bacterial metabolites and colorectal cancer. *Nat. Rev. Microbiol.* 12(10), 661-672 (2014).

185. Hiramatsu K, Takahashi K, Yamaguchi T *et al.*: N(1),N(12)-Diacetylspermine as a sensitive and specific novel marker for early- and late-stage colorectal and breast cancers. *Clin. Cancer Res.* 11(8), 2986-2990 (2005).

186. Kuwata G, Hiramatsu K, Samejima K *et al.*: Increase of N1, N12-diacetylspermine in tissues from colorectal cancer and its liver metastasis. *J. Cancer Res. Clin. Oncol.* 139(6), 925-932 (2013).

187. Linsalata M, Cavallini A, Messa C, Orlando A, Refolo MG, Russo F: Lactobacillus rhamnosus GG influences polyamine metabolism in HGC-27 gastric cancer cell line: a strategy toward nutritional approach to chemoprevention of gastric cance. *Curr. Pharm. Des.* 16(7), 847-853 (2010).

188. Bussiere FI, Chaturvedi R, Cheng Y *et al.*: Spermine causes loss of innate immune response to Helicobacter pylori by inhibition of inducible nitric-oxide synthase translation. *J. Biol. Chem.* 280(4), 2409-2412 (2005).

189. Iacopetta B: Are there two sides to colorectal cancer? *Int. J. Cancer* 101(5), 403-408 (2002).

190. Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G: XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal. Chem.* 84(11), 5035-5039 (2012).

191. Patti GJ, Tautenhahn R, Rinehart D *et al.*: A view from above: cloud plots to visualize global metabolomic data. *Anal. Chem.* 85(2), 798-804 (2013).

192. Wu S, Rhee KJ, Albesiano E *et al.*: A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* 15(9), 1016-1022 (2009).

193. Prorok-Hamon M, Friswell MK, Alswied A *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(5), 761-770 (2014).

194. Boleij A, Hechenbleikner EM, Goodwin AC *et al.*: The Bacteroides fragilis Toxin Gene Is Prevalent in the Colon Mucosa of Colorectal Cancer Patients. *Clin. Infect. Dis.* (2014).

195. Sears CL, Pardoll DM: Perspective: alpha-bugs, their microbial partners, and the link to colon cancer. *J. Infect. Dis.* 203(3), 306-311 (2011).

196. Tjalsma H, Boleij A, Marchesi JR, Dutilh BE: A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nat. Rev. Microbiol.* 10(8), 575-582 (2012).

197. Uemura N, Okamoto S, Yamamoto S *et al.*: Helicobacter pylori infection and the development of gastric cancer. *N. Engl. J. Med.* 345(11), 784-789 (2001).

198. Hand TW, Dos Santos LM, Bouladoux N *et al.*: Acute gastrointestinal infection induces long-lived microbiota-specific T cell responses. *Science* 337(6101), 1553-1556 (2012).

199. Evans AS: Causation and disease: the Henle-Koch postulates revisited. *Yale J. Biol. Med.* 49(2), 175-195 (1976).

CHRISTINE M. DEJEA CRAIG
CURRICULUM VITAE

Johns Hopkins School of Public Health	226 S. Madeira St.
Department of Molecular Microbiology and Immunology	Baltimore, MD 21231
615 N. Wolfe St., Suite E5132	(443) 955-2065
Baltimore, MD 21205	cdejea@jhsph.edu
(410) 955-9686	

EDUCATION:

Johns Hopkins Bloomberg School of Public Health • Baltimore, MD
> *Doctor of Philosophy*
> January 2015
> W. Harry Feinstone Department of Molecular Microbiology & Immunology
> GPA: 3.5

University of California, Davis • Davis, CA
> *Bachelor of Science (Summa Cum Laude)*
> May 2006
> Department of Biological Sciences
> Emphasis: Medical Microbiology (Departmental Honors)
> GPA: 3.9

RESEARCH EXPERIENCE:

Doctoral Thesis Research:
Johns Hopkins Bloomberg School of Public Health • Baltimore, MD • 2006-2013
> *Thesis Title:*
> > Biofilms of Colorectal Cancer

Undergraduate Thesis Research (Highest Honors):
University of California, Davis • Davis, CA • 2005-2006
> *Thesis Title:*
> > Evolution of bacteriophage virulence

TEACHING EXPERIENCE:

Instructor:
Johns Hopkins School of Nursing •Baltimore, MD • 2014-current
     NR.110.203.8101 Microbiology with Lab

PUBLICATIONS:

Dejea CM, Wick EM, Heckenbleikner EM, White JR, Mark-Welch JL, Rossetti BJ, Peterson SN, Snesrud EC, Borisy G, Lazarev M, Stein E, Vadivelu J, Roslani A, Malik A, Wanyiri J, Goh K, Iyadorai T, Fu K, Wan F, Llosa N, Housseae F, Romans K, Wu X, McAllister F, Wu S, Vogelstein B, Kinzler K, Pardoll DM, Sears CL. Microbiota organization rather than composition is an underlying feature of many colorectal cancers. PNAS submitted 2014.

Johnson C.*, **Dejea CM***, Santidrian AF, Felding-Habermann BH, Hoang LT, Wick EC, Heckenbleikner EM, Pardoll D, Sears CL, Siuzdak G. Metabolomics profiling correlates polyamine biosynthesis with colon cancer. PNAS, submitted 2015.

Dejea C, Wick EM, Sears CL. Bacterial oncogenesis in the colon. Future Microbiol, 2013

Lepore A.C., Tolmie C, O'Donell J, Wright M, **Dejea C**, Rauck B, Hoke A, Ignagni A, Onders R, Maragakis NJ. Peripheral Hyperstimulation Alters Site of Disease Onset and Course in SOD1 G93a Rats. Neurobiol Dis., 2010.☐

Lepore A.C., Rauck B, **Dejea C**, Pardo A.C. Roa M.S., Rothstein J.D., Maragakis N.J. Focal transplatation-based astrocyte replacement is neuroprotective in a model of motor neuron disease. Nature Neuroscience, 2008.

Lepore A.C., **Dejea C**, Carmen J, Rauck B, Kerr D.A., Sofroniew M.V., Maragakis N.J. Selective alation of proliferating astrocytes does not affect disease outcome in either acute or chronic models of motor neuron degeneration. Experimental Neurology, 2008.

PRESENTATIONS

     Dejea CM (Oral Presentation)
     Identification and Characterization of Biofilms on Sporadic Colorectal Cancer
     June 2012• Anaerobe Society of the Americas

COURSES AND CERIFICATES

2010 • Vaccine Certificate, Bloomberg School of Public Health
2011 • Advanced Anaerobic Bacteriology Workshop, Anaerobe systems
2012 • Institute for Genome Sciences Sequencing Workshop

OTHER HONORS & AWARDS:

2005 • Scholar Award, University of California, Davis
2006 • Phi Kappa Phi Honor Society
2006 • Phi Sigma Honor Society
2006 • Deans Citation Award (Contribution to Research), University California, Davis
2009 • Martin Frobisher Fellowship (Graduate Research Funding)
2010 • Martin Frobisher Fellowship (Graduate Research Funding)

RELATED WORK EXPERIENCE:

La Clinica Tepati Intern:
University of California, Davis, School of Medicine • Davis, CA • Summers 2005-2006

Research Assistant:
University of California, Davis • Davis, CA • 2005-2006
Department of Evolution and Ecology, Laboratory of Dr. Sergey Nuzhdin, PhD

Research Assistant
Johns Hopkins University, School of Medicine • Baltimore, MD • 2006-2008
Department of Neurology, Laboratory of Nicholas Maragakis, MD

LABORATORY SKILLS:

Animal (mouse) and infectious agents handling, DNA/RNA/protein isolation and purification, standard and real-time PCR, western blotting, ELISAs, immunohistochemistry and immunofluorescence microscopy, electron microscopy, recombinant DNA cloning, protein production and purification, cell culture (primary and cell lines), tissue preparation for histological processing.