Fusobacterium nucleatum Enrichment in Colorectal Cancer and its Relationship with Tumor Attributes

Matthew Goldberg

A thesis
Submitted in partial fulfillment of the
Requirements for the degree of

Master of Public Health

University of Washington
2021

Committee:
Amanda Phipps
Polly A. Newcomb
Meredith Hullar

Program Authorized of Offer Degree:
Epidemiology

University of Washington

Abstract:

Fusobacterium Nucleatum Enrichment in Colorectal Cancer and its Relationship with Tumor Attributes

Matthew Goldberg

Chair of Supervisory Committee:

Amanda Phipps, Ph.D., MPH
Epidemiology

Background*:* Aspects of the human gut microbiome have recently been linked to the etiology of colorectal cancer (CRC). Specifically, the commensal microbe *Fusobacterium nucleatum* (*F. nucleatum*) has been implicated in the development and progression of CRC, and has been found to be enriched in colorectal tumor tissue. Reflecting the fact that CRC is a heterogenous disease, *F. nucleatum* enrichment in CRC has been suggested to associate with various tumor molecular characteristics.
Methods*:* In this study, we used matched tumor and normal adjacent tissue collected from 403 participants in the Puget Sound Colorectal Cancer Cohort study to quantify the amount of *F. nucleatum* DNA in each respective tissue sample. Based on these generated data, we determined trends in tissue enrichment among tumors exhibiting different molecular and phenotypic characteristics of etiologic interest. Logistic regression was used to determine the odds of tumor tissue enrichment among tumors exhibiting different clinicopathologic and molecular characteristics, adjusting for age, sex, and BMI.
Results*: F. nucleatum* enrichment in tumor (relative to matched normal tissue) was found in 25% of samples analyzed. Samples from right sided tumors (OR: 2.67; 95% CI: 1.64, 4.74), MSI-high samples (OR: 2.81; 95% CI: 1.99, 3.93), and those with a mutation in the *BRAF* v600E gene (OR: 1.75; 95% CI: 1.09, 2.67) had significantly higher odds of *F. nucleatum* enrichment.
Conclusions*:* While CRC is a well-described disease, the relationship between its development and features of the gut microbiome (including *F. nucleatum*) are still becoming understood. The results of this study provide evidence linking *F. nucleatum* enrichment in tumor tissue to various molecular and phenotypic tumor attributes, and highlight several gaps in our understanding of the relationship between the development of CRC and the gut microbiome.

Introduction

The relationship between the human gastrointestinal (GI) tract and its associated microbiome is complex and multifaceted, and mediates beneficial and harmful health effects[7]. The balance of the microbial communities comprising the intestinal microbiome is in constant flux due to variations in diet and environment, and has been linked to numerous disorders, including inflammatory bowel disorders and colorectal cancer (CRC)[4]. Perturbations in the balance of beneficial and harmful bacteria in the gut creates opportunities for increased gut permeability and bacterial translocation as well as altered immune responses, which can lead to conditions favorable for tumor development[7].

The bacterium *Fusobacterium nucleatum* (*F. nucleatum*) has previously been implicated in the etiology and progression of CRC, and several epidemiological studies have found it to be highly enriched in colorectal tumor tissue[1, 6, 5]. This enrichment has been suggested to vary across subsets of CRC tissue samples and subtypes, suggesting a differential role of *F. nucleatum* across certain etiologic pathways[2,3]. In particular, previous epidemiological studies investigating *F. nucleatum* in the development of CRC have found that *F. nucleatum* is associated with tumors exhibiting mutations in *KRAS* and *BRAF* genes, as well as microsatellite instability[3,8,9,10]. These molecular features may associate with oncogenic signaling, and have been previously associated with etiologic pathways of functional and prognostic relevance to the development of CRC[3]. This study used a large, population-based cohort to investigate the relationship of *F. nucleatum* enrichment in CRC tumor tissue (as compared to matched normal tissue) with tumor site and stage at diagnosis, as well as relevant molecular characteristics (*KRAS* mutation status, *BRAF* v600E mutation status, and microsatellite instability).

METHODS

Study Population

This study used samples and data from the Puget Sound Colorectal Cancer Cohort (PSCCC). The PSCCC is comprised of study participants who were diagnosed with incident CRC between 1998-2007, who were 20-74 years old and residing in western Washington State at the time of CRC diagnosis. Subjects were identified as CRC patients and potential PSCCC participants through the use of the Surveillance, Epidemiology, and End Results (SEER) cancer registry serving the western Washington State region. Further description of the PSCCC population has been published previously[12]. All study participants completed a baseline questionnaire, from which information on epidemiologic and demographic factors was obtained.

Tumor and normal adjacent colorectal tissue biopsy samples for were obtained from diagnostic materials for PSCCC study participants from treating hospitals. The present analysis includes a subset of PSCCC participants for whom tumor and matched normal adjacent tissue specimens were available for analysis. All samples were available as formalin-fixed paraffin-embedded (FFPE) tissue.

All study participants included in the present analysis provided informed consent for the collection and analysis of their study materials. This project was approved by the Institutional Review Board of the Fred Hutchinson Cancer Research Center.

Tumor clinicopathologic and molecular characteristics

Information regarding tumor site and stage at diagnosis was obtained from SEER records. Tumor site was categorized as "right sided," "left sided," or "rectal" based on ICD-O-3 codes:[9] tumors located in the cecum, ascending colon, hepatic flexure and transverse colon were grouped together as right-sided tumors [ICD-O-3 codes were C18.0, C18.2, C18.3, and C18.4 respectively]; tumors located in the splenic fixture, descending colon, and sigmoid colon were grouped together as left-sided tumors [ICD-O-3 codes were C18.5, C18.6, and C18.7 respectively]; tumors located in the rectosigmoid junction and rectum were grouped together as rectal tumors [ICD-O-3 codes were C19.9 and C20.9 respectively]. Stage at diagnosis was classified as "localized," "regional," or "distant" according SEER summary stage definitions.

Information regarding tumor molecular characteristics was previously collected as part of the PSCCC. In particular, information regarding MSI status (dichotomized into "microsatellite stable or MSI-low" and "MSI-high"), *BRAF* v 600E somatic mutation status, and *KRAS* somatic mutation status was obtained as described elsewhere [8,9,10,11].

DNA extraction

In order to measure bacterial DNA in FFPE tissue, a modified DNA extraction protocol was used[15]. This protocol incorporates both mechanical and enzymatic lysis to optimize the extraction of bacterial genomic DNA. These lysis techniques are necessary to detect prokaryotic DNA through degradation of bacterial cell walls. Additional analysis techniques were then adapted to ensure that bacterial DNA signals were detectable in the presence of human eukaryotic genomes. In order to extract bacterial DNA for analysis, slides of FFPE tissue were sectioned; for tumor slides documented to be comprised of <80% lesional tissue, macrodissection was used to ensure that tissue was taken from tumor with a high degree

of specificity. The bacterial DNA was then extracted from the tissue for analysis with a microbiome-specific DNA extraction procedure to maximize bacterial DNA coverage (Qiagen QIAamp DNA FFPE Tissue Kit and QIAamp DNA Mini Kit).

Enrichment status for F. nucleatum

Bacterial DNA extracted from tumor and normal colorectal tissue were analyzed with digital droplet PCR (ddPCR) to determine the presence and extent of *F. nucleatum* enrichment in the tissue samples. Reactions used primers for the *nusG* gene of *F. nucleatum* to quantify the presence of these bacteria, as well as primers for the *SLCO2A1* eukaryotic housekeeping gene to enable normalization of observed *F. nucleatum* counts in the context of eukaryotic DNA levels.

The outcome of interest for this study was the enrichment of tumor tissue with *F. nucleatum*. This was done two ways based upon the presence or absence of *F. nucleatum* DNA in the matched normal adjacent tissue. In participants for whom no *F. nucleatum* DNA was detected in normal tissue, matched tumor tissue was considered "enriched" if the level of *F. nucleatum* DNA detected via ddPCR analysis exceeded the lower 25^{th} quartile, or a normalized value of 0.0003. For participants with detectable *F. nucleatum* DNA in both tumor and normal tissue samples, the tumor tissue was considered "enriched" if the detected level of *F. nucleatum* DNA in tumor tissue was greater than or equal to 2.5 times that observed in the matched normal tissue sample.

Covariates

We considered covariates known to be associated with the etiology of CRC for inclusion in our analysis as potential confounders. These included BMI (dichotomized into groups of participants with BMI $>25 kg/m^2$, or those with BMI $\leq 25 kg/m^2$), age (categorized by those aged <40, 40-49, 50-59, 60-69, and 70-74 years at the time of tissue sampling), sex, family history of CRC (dichotomized into those with and without a self-reported first degree family history), and cigarette use (dichotomized as "ever" and "never" smokers).

Statistical Analysis

Descriptive analyses of our study sample were conducted, including determination of the frequency of all confounders based on the previously defined *F. nucleatum* enrichment status of individual tissue sample pairs. Chi-squared analyses were conducted to determine the presence of significant patterns in the distribution of covariates along strata defined by the outcome variables. Logistic regression analyses

were carried out to examine associations of tumor clinicopathologic factors (i.e., tumor site and stage at diagnosis) and tumor markers (i.e., *BRAF* mutation status, *KRAS* mutation status, and MSI status) with *F. nucleatum* enrichment status. Age, sex, and BMI were included as covariates in all analyses. For analyses of clinicopathologic factors, we also assessed the impact of adjusting for MSI status. In analyses of tumor markers, we examined associations with markers individually and jointly, and assessed the impact of further adjustment by tumor site. All statistical analyses were conducted using R software.

Results

Details regarding the study population and distribution of *F. nucleatum* enrichment status are provided in Table 1. Overall, 25% (N=103) of study participants had tumors exhibiting enrichment for *F. nucleatum* relative to normal tissue. Compared to cases with tumors not exhibiting *F. nucleatum* enrichment, those with tumors enriched for *F. nucleatum* were more likely to have tumors located in the right side of the colon (65.0% vs. 42.7%; p<0.01), to have a somatic *BRAF* v.600E mutation (19.4% vs. 10.7%; p= 0.03), and to have MSI-high CRC (34.9% vs. 12.0%; p<0.01) (Table 1).

In regression analyses investigating the association of tumor site with *F. nucleatum* enrichment status, the odds of *F. nucleatum* enrichment was 2.67 (95% CI: 1.64, 4.74) times higher among those with right-sided tumors than among those with rectal cancer (Table 2). This association with tumor site persisted even when adjusting for MSI status (POR: 1.89, 95% CI: 1.07, 3.58), and when further accounting for stage at diagnosis (POR: 1.77, 95% CI: 1.07, 3.58). While the differences in odds of enrichment for left-sided tumors vs. rectal tumors were non-significant, there was a significant trend of increasing prevalence of *F. nucleatum* enrichment along a gradient of 3 distinct tumor sites in the colon and rectum, with the prevalence of enrichment increasing linearly from tumors located in the rectum, to those located on the left side of the colon, to those on the right side (p <0.01). Stage at diagnosis was not significantly associated with *F. nucleatum* enrichment status, with the exception of a suggestive association indicating elevated enrichment in distant stage CRC in models adjusting for MSI status and baseline covariates (POR=1.65, 95% CI: 1.06, 2.87).

In the regression analyses assessing the association of MSI status with *F. nucleatum* enrichment status, MSI-high tumors were significantly more likely to exhibit *F. nucleatum* enrichment (POR: 2.81, 95% CI: 1.99, 3.83), even after adjusting for tumor site (POR: 2.26, 95% CI: 1.55, 3.31) (Table 3). In regressions analyzing the role of *BRAF* v600E mutations in tumor tissue, *BRAF*-mutated CRC tumors were found to have significantly higher odds of *F. nucleatum* enrichment as compared to those with wildtype *BRAF* (POR: 1.75, 95% CI: 1.09, 2.67); however, this association was no longer significant after

adjustment for MSI status. When investigating the effect of *KRAS* mutations in tumor tissue on the odds of *F. nucleatum* enrichment, no significant relationship was found across all evaluated models. In analyses investigating the joint effect of MSI status, *BRAF* v600E mutation status, and *KRAS* mutation status, as well as tumor site, on the odds of enrichment, MSI status was again found to be associated with an increased prevalence of *F. nucleatum* enrichment (POR: 2.35, 95% CI: 1.58, 3.38); no associations were observed with *BRAF* and *KRAS* mutations.

Discussion

The relationship between resident gut microbiota and colorectal cancer (CRC) development has been found to be increasingly significant in recent years, with much of the scrutiny resting on candidate bacteria shown to be involved in the etiology of cancerous tissue[7,14]. Animal and *in vitro* studies have shown multiple mechanisms through which these resident microbes might induce carcinogenesis, including the induction of pro-inflammatory signaling, oxidative stress, and bacterially-derived toxins[5,7,14]. Our present study focused on the impact of *F. nucleatum* on CRC through investigation of patterns in *F. nucleatum* enrichment in tumors according to tumor site, stage at diagnosis, and different molecular characteristics. In regression analyses investigating the impact of tumor site at diagnosis, we observed that the odds of *F. nucleatum* enrichment were significantly increased among individuals with right sided tumors as compared to those with rectal tumors, even after accounting for MSI status and stage at diagnosis. In analyses investigating the role of various molecular characteristics on the odds of *F. nucleatum* enrichment, we observed significantly higher prevalence of enrichment in MSI-high and *BRAF* v600E mutated tumors. When MSI status, *BRAF* v600E mutation status, and *KRAS* mutation status were investigated simultaneously, MSI-high tumors were again found to be significantly associated with increased odds of enrichment in tumor tissue.

Previous work investigating the role of the gut microbiome has found significant relationships between aspects of the gut microbial community and the development of various types of CRC. Dysbiosis in the gut microbial communities has been linked to the development colorectal inflammation and CRC, as well as a wide variety of associated diseases [2,5,7,13,19]. In studies more specifically evaluating the impact of increased prevalence and enrichment of *F. nucleatum* in the colon, it has been shown that *F. nucleatum* infiltration of colonic tissue is associated with negative impacts on the innate immune system, as well as substantially increased odds of development of CRC, type II diabetes, and obesity (even beyond the risks associated with generalized dysbiosis of the gut)[4,5,6,7,14,21]. The mechanism for the *F. nucleatum*-associated development of these diseases is still under investigation, but some previous

studies have found that the infiltration of *F. nucleatum* into intestinal epithelial tissue enhances pro-inflammatory signaling and gene expression, and may induce oxidative stress-associated tissue damage potentially linked to the onset of these disorders[4,7].

Several previous studies have found significant associations in line with the present investigation's findings, in which they describe significant associations between *F. nucleatum* enrichment and tumor tissue exhibiting *BRAF* mutations and MSI [9,10,11]. These previous findings indicate that tumor tissue enriched with *F. nucleatum* tend to be MSI-high or *BRAF* mutated, and are associated with significantly worse patient-specific and overall survival[21,22,23]. Other recent studies have also observed trends in the prevalence of *F. nucleatum* enrichment along a colonic site gradient, where there is increasing prevalence of enrichment from the rectum to the cecum theorized to reflect differences in the impact of pathogenic exposure of the gut microbiome on neoplastic and immune cells in different regions of the colon[12]. However, there has been mixed evidence of an association of enrichment with tumor stage at diagnosis, with many studies failing to find a statistically significant association.

Many studies have investigated the connection between *F. nucleatum* infiltration of intestinal epithelial tissue and CRC and have found several etiologic pathways might explain the development of CRC in *F. nucleatum* enriched tissues. The presence of *F. nucleatum* in colonic tissue is associated with increased numbers of infiltrating immune cells, which creates elevated levels of inflammatory signaling markers such as TNFα, IL-8, and other pro-inflammatory signaling molecules [7,13]. The presence of *F. nucleatum* also inhibits the T-cell mediated response against colorectal tumors, and is inversely associated with $CD3^+$ T-cell density in persons with CRC [7,13]. Studies investigating the role of *F. nucleatum* in CRC have also highlighted the butyrogenic nature of the microbe—butyrogenic species have been found to be associated with a particularly rapid increase in the proliferation of colorectal cells, leading to a potential increase in the development of precancerous lesions and growths in the colon [7,13]. *F. nucleatum* also expresses the virulence factor FadA on its surface, which previous studies have shown activates WNT signaling in colorectal cells, thus promoting cell division and enhanced potentiate tumor growth[12]. In the context of CRC, a high abundance of *F. nucleatum* in colonic tissue is associated with microsatellite instability and BRAF mutations, as well as a wide array of other phenotypic changes associated with an increased risk of the development and potentiation of CRC [9,10,11].

There are several limitations to this study that should be considered when reviewing our findings. One such limitation comes from the relatively limited set of samples available for this study, with ddPCR data from 405 subjects available for analysis. As tissue samples from the PSCCC continue to be analyzed, this dataset will increase in size; however, at this early stage of the study, there was a

limited set of matched tumor and normal tissue data available for analysis. While this prevented us from looking at several more detailed classifications related to previously defined etiologic pathways of CRC, there was still sufficient data to ensure that our analyses maintained adequate power to detect meaningful and significant associations. Additional limitations arose from the historical storage of formalin-fixed, paraffin-embedded (FFPE) tissue. The DNA of FFPE tissue has been shown to degrade over time, and tissues stored in paraffin blocks have been shown to be subject to external DNA contamination as well [15]. These concerns were addressed through protocols designed to detect nucleic acid degradation in preliminary analyses, as well as protocols which allowed for the detection and normalization of ddPCR data to protect from potential contamination acquired in storage. Additionally, as these tissues are collected and analyzed cross-sectionally, we are unable to determine whether *F. nucleatum* infiltration of colonic tissue preceded or followed tumor formation and tumor marker acquisition. However, previous studies have suggested that *F. nucleatum* has an etiologic role in cancer formation[1,2,5,6,14,16,20].

There are also several notable strengths to our study. A major strength is the fact that we were able to define *F. nucleatum* enrichment through the use of both tumor and matched normal adjacent tissue. Several previous studies have linked the presence of *F. nucleatum* in CRC tissue to the development of CRC, but most have been limited by their access to either tumor tissue alone or non-matched normal tissue[20]. In the absence of matched samples, some studies have defined *F. nucleatum* enrichment in tumor tissue through visualization techniques such as fluorescent in-situ hybridization imaging[20]. Others defined *F. nucleatum* status in CRC based on whether the quantity of *F.* nucleatum DNA found in metagenomic and PCR screening exceeded objective thresholds established *a priori*, or through comparison to normal tissue collected from patients without CRC[20]. By defining *F. nucleatum* enrichment in tumor tissue based on a comparison of matched adjacent normal tissue samples, this study was able to develop a more robust and normalized definition of enrichment for analysis. This study was also able to investigate the associations with of multiple molecular markers on the odds of enrichment simultaneously. In doing so, this study was able to align our findings to distinct development pathways relevant to our understanding of the etiology, prognosis, and diagnostic modalities of CRC[20]. Finally, this study was able to use ddPCR analysis to increase the sensitivity of our nucleic acid analysis. ddPCR functions by partitioning nucleic acid samples into thousands of droplets within which PCR amplification is carried out, allowing for greater sensitivity and reduction in background signal[24]. The use of ddPCR allowed for more efficient analysis, and optimized use of the tissue samples available.

The results of the analyses conducted in this study indicate clear associations of MSI status and right sided tumor location with odds of tumor tissue being enriched with *F. nucleatum* as compared to matched normal adjacent tissue. The association between *F. nucleatum* enrichment in tumor tissue on the development of CRC have been well characterized in recent years, with links discovered between tissue damage and virulence factors associated with *F. nucleatum* in the intestine and the development of CRC along several previously described etiologic pathways. While CRC is a well-described disease, the relationship between its development and progression with features of the gut microbiome (including *F. nucleatum*) are still becoming understood. There is enormous potential to learn more about the microbiome-associated factors linked to the onset of CRC, and the impact of various resident microbial species in isolation, as well as in terms of their role in the overall composition of the gut microbial community. The results of this study provide evidence linking *F. nucleatum* enrichment in tumor tissue to various molecular and phenotypic tumor attributes, and highlights several gaps in our understanding of the relationship between the development of CRC and the gut microbiome. Further investigation is needed to clarify the ways that *F. nucleatum* contributes to the development of CRC along different etiologic pathways; however, this study's findings provide further evidence of its association to traits linked to distinct CRC developmental pathways.

Tables

Table 1: Descriptive Statistics of Puget Sound Colorectal Cancer Cohort (PSCCC) Participants by intra-tissue Fusobacterium nucleatum enrichment status (n = 403)

		Tumor Tissue Enriched in *F. Nucleatum* (N, %)	**Tumor Tissue non-Enriched with *F. Nucleatum* (N, %)**
		N = 103	N = 300
Age at diagnosis[1]			
	<40	15 (14.5)	34 (11.3)
	40-49	36 (34.9)	107 (35.7)
	50-59	15 (14.5)	46 (15.3)
	60-69	24 (23.3)	72 (24)
	70-74	13 (12.6)	41 (13.7)
BMI[1]			
	<25.0kg/m^2	38 (36.9)	98 (32.7)
	≥25.0kg/m^2	65 (63.1)	202 (67.3)
Sex[1]			
	Male	35 (33.9)	129 (43.0)
	Female	68 (66.1)	171 (57.0)
Tumor Site[1,2]			
	Right Sided	67 (65.0)	128 (42.7)
	Left Sided	21 (20.3)	80 (26.7)
	Rectal	14 (13.7)	88 (29.3)
Tumor Stage[1]			
	Local	29 (28.2)	84 (28.0)
	Regional	57 (55.3)	174 (58.0)
	Distant	16 (15.5)	39 (14.0)
BRAF V600E Mutation[2]			
	Wildtype	75 (72.8)	249 (83)
	Mutated	20 (19.4)	32 (10.7)
	Missing	*19 (18.4)*	*8 (6.3)*
KRAS mutation			
	Wildtype	69 (66.9)	183 (61.0)
	Mutated	25 (24.2)	87 (29.0)
	Missing	*9 (8.7)*	*30 (10.0)*
Microsatellite Instability (MSI) Status[2,3]			
	MSS or MSI-L	53 (53.4)	237 (79)
	MSI-H	36 (34.9)	36 (12)
	Missing	*27 (26.3)*	*14 (9)*

1: Missing for less than 1% of sample

2: This variable was found to be significantly associated with the probability of *F. nucleatum* enrichment via a X^2 test of independence
3: MSS: Microsatellite Stable; MSI-L: low Microsatellite Instability; MSI-H: high Microsatellite Instability

Table 2: Prevalence Ratios of tissue enrichment with Fusobacterium nucleatum by clinicopathologic characteristics

	N	*Prevalence Ratio in Model 1 (95% CI)[1]*	*Prevalence Ratio in Model 2 (95% CI)[2]*	*Prevalence Ratio in Model 3 (95% CI)[3]*
Site:				
Right colon	197	2.67 (1.64, 4.74) ***	1.89 (1.07, 3.58) *	1.77 (1.07, 3.57) *
Left colon	101	1.56 (0.86, 2.98)	1.37 (0.72, 2.70)	1.32 (0.72, 2.69)
Rectum	102	1.0 (ref)	1.0 (ref)	1.0 (ref)
Stage at diagnosis:				
Localized	113	1.0 (ref)	1.0 (ref)	1.0 (ref)
Regional	233	0.93 (0.64, 1.39)	0.98 (0.62, 1.46)	0.98 (0.66, 1.39)
Distant	55	1.12 (0.65, 1.86)	1.65 (1.06, 2.87) *	1.48 (0.82, 2.91)

1: Model covariates include age, sex, BMI
2: Model covariates include age, sex, BMI, MSI Status
3: Model covariates includes age, sex, BMI, MSI, tumor site, and tumor stage
*: $p < 0.05$
**: $p < 0.001$
***: $P < 0.0001$

Table 3: Prevalence Ratios of tissue enrichment with Fusobacterium nucleatum by molecular characteristics

	N	*Prevalence Ratio in Model 1 (95% CI)*[1]	*Prevalence Ratio in Model 2 (95% CI)*[2]	*Prevalence Ratio in Model 3 (95% CI)*[3]
MSI status[4]				
MSI-high	74	2.81 (1.99, 3.93)***	2.26 (1.55, 3.31)***	2.35 (1.58, 3.38) ***
MSS/MSI-low	290	1.0 (ref)	1.0 (ref)	1.0(ref)
KRAS mutation:	253	0.82 (0.54, 1.21)	0.85 (0.56, 1.24)	0.93 (0.57, 1.25)
Mutated	113	1.0 (ref)	1.0 (ref)	1.0 (ref)
Wildtype				
BRAF mutation	53	1.75 (1.09, 2.67)**	1.29 (0.72, 1.45)	0.70 (0.41, 1.36)
Mutated	325	1.0 (ref)	1.0 (ref)	1.0 (ref)
Wildtype				

1: Model covariates include age, sex, BMI
2: Model covariates include age, sex, BMI, Tumor site
3: Model covariates includes age, sex, BMI, tumor site, MSI, BRAF, KRAS
4: MSS: Microsatellite Stable; MSI-L: low Microsatellite Instability; MSI-H: high Microsatellite Instability
*: $p < 0.05$
**: $p < 0.001$
***: $p < 0.0001$

References:

1. Castellarin, Mauro, et al. “Fusobacterium Nucleatum Infection Is Prevalent in Human Colorectal Carcinoma.” *Genome Research*, vol. 22, no. 2, Feb. 2012, pp. 299–306. *genome.cshlp.org*, doi:10.1101/gr.126516.111.
2. Flanagan, L., et al. “Fusobacterium Nucleatum Associates with Stages of Colorectal Neoplasia Development, Colorectal Cancer and Disease Outcome.” *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 33, no. 8, Aug. 2014, pp. 1381–90. *Springer Link*, doi:10.1007/s10096-014-2081-3.
3. Jass, Jeremy R. “Molecular Heterogeneity of Colorectal Cancer: Implications for Cancer Control.” *Surgical Oncology*, vol. 16, Dec. 2007, pp. 7–9. *ScienceDirect*, doi:10.1016/j.suronc.2007.10.039.
4. Kau, Andrew L., et al. “Human Nutrition, the Gut Microbiome and the Immune System.” *Nature*, vol. 474, no. 7351, 7351, Nature Publishing Group, June 2011, pp. 327–36. *www.nature.com*, doi:10.1038/nature10213.
5. Kostic, Aleksandar D., et al. “Genomic Analysis Identifies Association of Fusobacterium with Colorectal Carcinoma.” *Genome Research*, vol. 22, no. 2, Feb. 2012, pp. 292–98. *genome.cshlp.org*, doi:10.1101/gr.126573.111.
6. McCoy, Amber N., et al. “Fusobacterium Is Associated with Colorectal Adenomas.” *PLOS ONE*, vol. 8, no. 1, Public Library of Science, Jan. 2013, p. e53653. *PLoS Journals*, doi:10.1371/journal.pone.0053653.
7. Peterson, C. T., et al. “Immune Homeostasis, Dysbiosis and Therapeutic Modulation of the Gut Microbiota.” *Clinical & Experimental Immunology*, vol. 179, no. 3, 2015, pp. 363–77. *Wiley Online Library*, doi:10.1111/cei.12474.
8. Phipps, Amanda I., et al. "KRAS-mutation status in relation to colorectal cancer survival: the joint impact of correlated tumour markers." *British journal of cancer* 108.8 (2013): 1757-1764.
9. Phipps, Amanda I., et al. "Association between molecular subtypes of colorectal cancer and patient survival." *Gastroenterology* 148.1 (2015): 77-87.
10. Phipps, Amanda I., et al. "BRAF mutation status and survival after colorectal cancer diagnosis according to patient and tumor characteristics." *Cancer Epidemiology and Prevention Biomarkers* 21.10 (2012): 1792-1798.
11. Newcomb PA, Zheng Y, Chia VM, Morimoto LM, Doria-Rose VP, Templeton A, Thibodeau SN, Potter JD. Estrogen plus progestin use, microsatellite instability, and the risk of colorectal cancer in women. Cancer Res. 2007 Aug 1;67(15):7534-9. doi: 10.1158/0008-5472.CAN-06-4275. PMID: 17671225.
12. Mima, Kosuke et al. “Fusobacterium nucleatum in Colorectal Carcinoma Tissue According to Tumor Location.” *Clinical and translational gastroenterology* vol. 7,11 e200. 3 Nov. 2016, doi:10.1038/ctg.2016.53
13. Purcell, Rachel V., et al. “Distinct Gut Microbiome Patterns Associate with Consensus Molecular Subtypes of Colorectal Cancer.” *Scientific Reports*, vol. 7, no. 1, 1, Nature Publishing Group, Sept. 2017, p. 11590. *www.nature.com*, doi:10.1038/s41598-017-11237-6.
14. Kostic, Aleksandar D., et al. “Fusobacterium Nucleatum Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment.” *Cell Host & Microbe*, vol. 14, no. 2, Elsevier, Aug. 2013, pp. 207–15. *www.cell.com*, doi:10.1016/j.chom.2013.07.007.

15. Flores Bueso, Yensi, et al. “Characterization of FFPE-Induced Bacterial DNA Damage and Development of a Repair Method.” *Biology Methods and Protocols*, vol. 5, no. bpaa015, Jan. 2020. *Silverchair*, doi:10.1093/biomethods/bpaa015.
16. McCoy, Amber N., et al. “Fusobacterium Is Associated with Colorectal Adenomas.” *PLOS ONE*, vol. 8, no. 1, Public Library of Science, Jan. 2013, p. e53653. *PLoS Journals*, doi:10.1371/journal.pone.0053653.
17. Jass, Jeremy R. “Molecular Heterogeneity of Colorectal Cancer: Implications for Cancer Control.” *Surgical Oncology*, vol. 16, Dec. 2007, pp. 7–9. *ScienceDirect*, doi:10.1016/j.suronc.2007.10.039.
18. Turnbaugh, Peter J., et al. “A Core Gut Microbiome in Obese and Lean Twins.” *Nature*, vol. 457, no. 7228, 7228, Nature Publishing Group, Jan. 2009, pp. 480–84. *www.nature.com*, doi:10.1038/nature07540.
19. Vatanen, Tommi, et al. “The Human Gut Microbiome in Early-Onset Type 1 Diabetes from the TEDDY Study.” *Nature*, vol. 562, no. 7728, 7728, Nature Publishing Group, Oct. 2018, pp. 589–94. *www.nature.com*, doi:10.1038/s41586-018-0620-2.
20. Shang, Fu-Mei, and Hong-Li Liu. “Fusobacterium Nucleatum and Colorectal Cancer: A Review.” *World Journal of Gastrointestinal Oncology*, vol. 10, no. 3, Mar. 2018, pp. 71–81. *PubMed Central*, doi:10.4251/wjgo.v10.i3.71.
21. Tahara, Tomomitsu, et al. “Fusobacterium in Colonic Flora and Molecular Features of Colorectal Carcinoma.” *Cancer Research*, vol. 74, no. 5, American Association for Cancer Research, Mar. 2014, pp. 1311–18. *cancerres.aacrjournals.org*, doi:10.1158/0008-5472.CAN-13-1865.
22. de Carvalho, Ana Carolina, et al. “Microbiota Profile and Impact of Fusobacterium Nucleatum in Colorectal Cancer Patients of Barretos Cancer Hospital.” *Frontiers in Oncology*, vol. 9, Frontiers, 2019. *Frontiers*, doi:10.3389/fonc.2019.00813.
23. Yamamoto, Shumpei, et al. “Heterogeneous Distribution of Fusobacterium Nucleatum in the Progression of Colorectal Cancer.” *Journal of Gastroenterology and Hepatology*, vol. n/a, no. n/a. *Wiley Online Library*, doi:https://doi.org/10.1111/jgh.15361. Accessed 25 May 2021.
24. Yang, Rongchang, et al. "Comparison of next-generation droplet digital PCR (ddPCR) with quantitative PCR (qPCR) for enumeration of Cryptosporidium oocysts in faecal samples." *International journal for parasitology* 44.14 (2014): 1105-1113.