Association of *Fusobacterium nucleatum* with Specific T-cell Subsets in the Colorectal Carcinoma Microenvironment

Jennifer Borowsky[1,2,3,4], Koichiro Haruki[1,5], Mai Chan Lau[1,5], Andressa Dias Costa[6], Juha P. Väyrynen[1,5,6,7], Tomotaka Ugai[1,8], Kota Arima[1,5], Annacarolina da Silva[5], Kristen D. Felt[9], Melissa Zhao[5], Carino Gurjao[6], Tyler S. Twombly[1,5], Kenji Fujiyoshi[1,5], Sara A. Väyrynen[5,6], Tsuyoshi Hamada[1,5], Kosuke Mima[5], Susan Bullman[10], Tabitha A. Harrison[11], Amanda I. Phipps[11,12], Ulrike Peters[11,12], Kimmie Ng[6], Jeffrey A. Meyerhardt[6], Mingyang Song[13,14,15], Edward L. Giovannucci[8,13], Kana Wu[8,13,16], Xuehong Zhang[16], Gordon J. Freeman[6], Curtis Huttenhower[17,18], Wendy S. Garrett[6,18,19], Andrew T. Chan[14,15,16,19], Barbara A. Leggett[3,4,20], Vicki L.J. Whitehall[3,4,21], Neal Walker[4,22], Ian Brown[4,22], Mark Bettington[3,4,22], Reiko Nishihara[1,5,8,13,17], Charles S. Fuchs[23,24,25], Jochen K. Lennerz[2], Marios Giannakis[6,18,26], Jonathan A. Nowak[5], and Shuji Ogino[1,5,8,18,27]

ABSTRACT

Purpose: While evidence indicates that *Fusobacterium nucleatum* (*F. nucleatum*) may promote colorectal carcinogenesis through its suppressive effect on T-cell–mediated antitumor immunity, the specific T-cell subsets involved remain uncertain.

Experimental Design: We measured *F. nucleatum* DNA within tumor tissue by quantitative PCR on 933 cases (including 128 *F. nucleatum*–positive cases) among 4,465 incident colorectal carcinoma cases in two prospective cohorts. Multiplex immunofluorescence combined with digital image analysis and machine learning algorithms for CD3, CD4, CD8, CD45RO (PTPRC isoform), and FOXP3 measured various T-cell subsets. We leveraged data on *Bifidobacterium*, microsatellite instability (MSI), tumor whole-exome sequencing, and M1/M2-type tumor-associated macrophages [TAM; by CD68, CD86, IRF5, MAF, and MRC1 (CD206) multimarker assay]. Using the 4,465 cancer cases and inverse probability weighting method to control for selection bias due to tissue availability, multivariable-adjusted logistic regression analysis assessed the association between *F. nucleatum* and T-cell subsets.

Results: The amount of *F. nucleatum* was inversely associated with tumor stromal $CD3^+$ lymphocytes [multivariable OR, 0.47; 95% confidence interval (CI), 0.28–0.79, for *F. nucleatum*–high vs. -negative category; $P_{trend} = 0.0004$] and specifically stromal $CD3^+CD4^+CD45RO^+$ cells (corresponding multivariable OR, 0.52; 95% CI, 0.32–0.85; $P_{trend} = 0.003$). These relationships did not substantially differ by MSI status, neoantigen load, or exome-wide tumor mutational burden. *F. nucleatum* was not significantly associated with tumor intraepithelial T cells or with M1 or M2 TAMs.

Conclusions: The amount of tissue *F. nucleatum* is associated with lower density of stromal memory helper T cells. Our findings provide evidence for the interactive pathogenic roles of microbiota and specific immune cells.

[1]Department of Oncologic Pathology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts. [2]Department of Pathology, Center for Integrated Diagnostics, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts. [3]Conjoint Gastroenterology Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [4]School of Medicine, The University of Queensland, Brisbane, Queensland, Australia. [5]Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts. [6]Department of Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts. [7]Cancer and Translational Medicine Research Unit, Medical Research Center Oulu, Oulu University Hospital, and University of Oulu, Oulu, Finland. [8]Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, Massachusetts. [9]Center for Immuno-Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts. [10]Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington. [11]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington. [12]Department of Epidemiology, University of Washington, Seattle, Washington. [13]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. [14]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts. [15]Division of Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts. [16]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts. [17]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. [18]Broad Institute of MIT and Harvard, Cambridge, Massachusetts. [19]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. [20]The Royal Brisbane and Women's Hospital, Brisbane, Queensland, Australia. [21]Conjoint Internal Medicine Laboratory, Pathology Queensland, Queensland Health, Brisbane, Queensland, Australia. [22]Envoi Specialist Pathologists, Brisbane, Queensland, Australia. [23]Yale Cancer Center, New Haven, Connecticut. [24]Department of Medicine, Yale School of Medicine, New Haven, Connecticut. [25]Smilow Cancer Hospital, New Haven, Connecticut. [26]Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts. [27]Cancer Immunology and Cancer Epidemiology Programs, Dana-Farber Harvard Cancer Center, Boston, Massachusetts.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

J. Borowsky, K. Haruki, M.C. Lau, and A. Dias Costa contributed equally as co–first authors of this article.

J.K. Lennerz, M. Giannakis, J.A. Nowak, and S. Ogino contributed equally as co–last authors of this article.

Corresponding Authors: Shuji Ogino, Department of Pathology, Brigham and Women's Hospital, 221 Longwood Avenue, Boston, MA 02115. Phone: 617-525-8953; Fax: 617-264-5149; E-mail: sogino@bwh.harvard.edu; and Jonathan A. Nowak, Phone: 617-732-7641; E-mail: janowak@bwh.harvard.edu

Clin Cancer Res 2021;27:2816–26

doi: 10.1158/1078-0432.CCR-20-4009

Translational Relevance

Our multiplex immunofluorescence assay combined with digital image analyses and machine learning enabled us to robustly quantify T-cell subsets in 933 colorectal cancer cases within a large database of 4,465 incident colorectal cancers in two prospective U.S.-wide cohort studies. We found an inverse association of *Fusobacterium nucleatum* DNA amount with tumor stromal density of $CD3^+CD4^+CD45RO^+$ memory helper T cells. Fewer memory T cells may contribute to a lack of immune attack on developing tumors. Our unique human population–based evidence for microbial–immune interactions supports possible interventions targeting the microbiota and/or immunity for cancer prevention and therapy.

Introduction

Accumulating evidence indicates the complex role of the microbiota and host immunity in carcinogenesis (1–4). As T-cell–mediated adaptive immune response influences tumor development (5, 6), immunotherapies targeting immune checkpoints that regulate T-cell activity have emerged as promising treatment strategies for various types of malignancy, including colorectal cancer (7–9). In colorectal cancer, high densities of tumor-infiltrating immune cells, including $CD3^+$ cells, $CD8^+$ cells, and $CD45RO^+$ cells have been associated with better survival (10–12). Recent advances in multiplex immunofluorescence technologies have enabled detailed evaluation of cellular phenotype, as well as information on spatial distribution of each cell in relation to tumor epithelium and stroma, thereby providing deeper insights into the tumor immune microenvironment (13).

Fusobacterium nucleatum (*F. nucleatum*), a Gram-negative anaerobic bacterium, is nearly ubiquitous in the oral cavity, yet rarely found in other anatomic sites under healthy conditions (14, 15). In disease states, it has been identified in various extraoral sites, with studies demonstrating an enrichment of *F. nucleatum* in colorectal carcinoma tissue compared with adjacent normal tissue (16–18). Experimental studies have shown that *F. nucleatum* activates the WNT/CTNNB1 (β-catenin) signaling pathway in colorectal cancer cells, promotes tumor growth (19, 20), and inhibits T-cell–mediated immune responses against colorectal cancer (20, 21). Consistent with these findings, a higher amount of *F. nucleatum* DNA in colorectal cancer tissue has been associated with advanced disease stage, worse survival, and lower T-cell density in tumor (18, 22). However, the relationship between *F. nucleatum* and specific T-cell subsets in the colorectal tumor microenvironment has not been well defined.

Using a multiplex assay that allows detailed characterization of T cells and their spatial localization within tumor intraepithelial and stromal regions, this analysis represents one of the first human studies to test the hypothesis that *F. nucleatum* DNA amount in tumor tissue might be inversely associated with specific T-cell subset densities.

Materials and Methods

Study population

We utilized data from two prospective cohort studies in the United States: the Nurses' Health Study (NHS; 121,701 women ages 30–55 years followed since 1976) and the Health Professionals Follow-up Study (HPFS; 51,529 men aged 40–75 years followed since 1986; refs. 22–24). Study participants have been sent questionnaires biennially to update information on lifestyle factors and newly diagnosed diseases, including colorectal cancer. The follow-up rate has been more than 90% for each follow-up questionnaire cycle in both cohort studies. In both studies, the National Death Index was used to ascertain deaths of study participants and identify unreported lethal colorectal cancer cases. Study physicians, who were blinded to exposure data, reviewed medical records of identified colorectal cancer cases to confirm the disease diagnosis and to collect data on tumor size, tumor anatomic location, and disease stage based on the American Joint Committee on Cancer tumor–node–metastasis classification. We included both colon and rectal carcinomas based on the colorectal continuum model (25). In the cohort studies, 4,465 incident colorectal carcinoma cases had been documented up to 2012. We attempted to collect formalin-fixed, paraffin-embedded (FFPE) tumor tissue blocks from hospitals throughout the United States (where patients with colorectal cancer had undergone surgical resection), resulting in 933 cases with sufficient tissue and T-cell data. A single pathologist (S. Ogino), blinded to other data, reviewed hematoxylin and eosin–stained tissue sections and recorded pathologic features. Tumor differentiation was categorized into well/moderate versus poor (>50% vs. ≤50% gland formation, respectively).

The study was conducted in accordance with the U.S. Common Rule. All participants gave written informed consent for the study. This study was approved by the Institutional Review Boards of Brigham and Women's Hospital (Boston, MA) and Harvard T.H. Chan School of Public Health (Boston, MA), and those of participating registries as required.

Multiplex immunofluorescence analyses for T cells and macrophages in tumor

We constructed tissue microarrays of colorectal cancer cases with sufficient tissue materials, including up to four tumor cores from each case in a tissue microarray block (26). As described previously (27, 28), 4 μm sections from tissue microarray blocks were sequentially stained using the following antibody and fluorophore combinations, in order: anti-CD3 antibody (clone F7.2.38, Dako; Agilent Technologies)/Opal-520, anti-FOXP3 (clone 206D, BioLegend)/Opal-540, anti-CD45RO (one of PTPRC isoforms, clone UCHL1, Dako)/Opal-650, anti-CD8 (clone C8/144B, Dako)/Opal-570, anti-CD4 (clone 4B12, Dako)/Opal-690, and anti-KRT (keratin, pan-cytokeratins; clone AE1/AE3, Dako and clone C11, Cell Signaling Technology)/Opal-620 (Supplementary Fig. S1; with standardized protein nomenclature recommended by a panel of experts; ref. 29).

Digital images of all tissue microarray cores were acquired at 200× magnification using the Vectra Multispectral Imaging Platform (Vectra 3.0, Akoya Biosciences; Hopkinton, MA). Images of each core underwent tissue segmentation to characterize regions of tumor epithelium and peritumoral stroma based on KRT (keratin) expression using supervised machine learning algorithms within Inform 2.4.1 (Akoya Biosciences). Following tissue segmentation, cell enumeration and segmentation were performed using the DAPI signal to aid in identification of nuclei (Supplementary Fig. S1). We evaluated comparison of T-cell subset densities between the initial and reprocessed data, demonstrating a high degree of concordance (Pearson $r = 0.48$–0.99; Spearman $r = 0.77$–0.99) between the density measurements. Each cell was further segmented into nuclear, cytoplasmic, and membranous compartments. A separate supervised machine learning algorithm was used to identify T cells based upon a combination of cytomorphology and subcellular T-cell marker expression patterns. We evaluated self-reproducibility and the reliability of tissue segmentation and T-cell identification using supervised machine learning and

confirmed a high degree of concordance (J. Borowsky). We also evaluated pathologist-to-pathologist (J. Borowsky and A. Dias Costa) concordance through their independent analyses starting from raw multispectral image data and confirmed moderate to high concordance. These single-cell data were then used to calculate T-cell subpopulation densities within separate regions. Aggregate tumor-level densities were then determined by calculating the average density for each subset across all regions from each tumor. T-cell densities were initially classified into quartile categories (C1–C4). If more than 25% of all cases had zero density for a specific cell type, these zero-value cases were grouped together (C1 category), and the remaining (nonzero-value) cases were divided into tertile categories according to density (C2–C4).

In exploratory analyses, we evaluated tumor-associated macrophage (TAM) densities and polarization using a separate multiplex immunofluorescence panel that included a pan-macrophage marker (CD68), two markers generally expressed in M1 TAMs (CD86 and IRF5), two markers generally expressed in M2 TAMs [MAF and MRC1 (CD206)], a tumor cell marker (KRT), and DAPI, as described previously (30). We calculated an M1:M2 polarization index using the formula "(CD86 × IRF5)/(MRC1 × MAF)" based on the expression level of each protein. The TAMs within the highest 30% of the index were regarded as M1-like TAMs, while the macrophages within the lowest 30% were regarded as M2-like TAMs for these analyses, according to our previous study (30). We calculated each cell density (cell count/mm^2) in tumor intraepithelial and stromal regions separately, and macrophage densities were classified into quartile categories (C1–C4).

DNA analyses for *F. nucleatum* and *Bifidobacterium* genus in tumor

Genomic DNA was extracted from archival FFPE tissue sections of colorectal carcinoma using the QIAamp DNA FFPE Tissue Kit (Qiagen). As described previously (18, 31), we performed quantitative PCR assays to measure the amounts of *F. nucleatum* and *Bifidobacterium* genus DNA in tumor tissue, using *SLCO2A1* (for *F. nucleatum*) or a universal 16S primer set (for *Bifidobacterium* genus) as reference genes. Cases with any detectable *F. nucleatum* DNA (or *Bifidobacterium* genus) were categorized as low versus high on the basis of the median level of *F. nucleatum* (or *Bifidobacterium* genus), while cases without detectable *F. nucleatum* (or *Bifidobacterium* genus) were categorized as negative.

Evaluation of tumor molecular characteristics

Tumor microsatellite instability (MSI) status was analyzed using PCR for 10 microsatellite markers (D2S123, D5S346, D17S250, BAT25, BAT26, BAT40, D18S55, D18S56, D18S67, and D18S487), and MSI-high was defined as presence of instability in ≥30% of the markers, as described previously (24, 28, 32). Using bisulfite-treated DNA, methylation status of eight CpG island methylator phenotype (CIMP)-specific promoters (*CACNA1G, CDKN2A, CRABP1, IGF2, MLH1, NEUROG1, RUNX3*, and *SOCS1*) and long interspersed nucleotide element-1 (LINE-1) was determined as described previously (24, 28, 32). CIMP-high was defined as ≥6 methylated promoters of eight promoters, and CIMP-low/negative as 0–5 methylated promoters, as described previously (24, 28, 32). PCR and pyrosequencing were performed for *KRAS* (codons 12, 13, 61, and 146), *BRAF* (codon 600), and *PIK3CA* (exons 9 and 20), as described previously (23, 24, 28, 33). Whole-exome sequencing was performed using DNA from tumor and matched normal tissue pairs, as described previously (34). Exome-wide tumor mutational burden was defined as the number of nonsynonymous somatic mutations identified per megabase in the sequenced exome. Using a neoantigen prediction pipeline for somatic mutations, the neoantigen load (i.e., the number of proteins that likely give rise to immunogenic peptides in the tumor microenvironment) were estimated by counting peptides that bind to personal human leukocyte antigen molecules with high affinity (<500 nmol/L). Using NetMHCpan (version 2.4; ref. 35), we predicted the binding affinities of all possible 9- and 10-mer mutant peptides to the corresponding human leukocyte antigen alleles inferred by the POLYSOLVER algorithm, as described previously (28, 34). On the basis of all colorectal cancers with available whole-exome sequencing data, neoantigen load and exome-wide tumor mutational burden were divided into quartiles (Q1–Q4).

Statistical analysis

All statistical analyses were conducted using SAS Software (version 9.4, SAS Institute), and all *P* values were two-sided. We used the two-sided α level of 0.005 as recommended previously (36). The primary hypothesis of this study was an assessment of a statistical trend of association of the amount of *F. nucleatum* DNA (negative, low, and high; as an ordinal predictor variable) with each T-cell subset density (an ordinal outcome variable). All other analyses represented secondary analyses. In secondary analyses, to assess the association between *F. nucleatum* DNA amount and each of categorical clinicopathologic variables, the χ^2 test was performed. To compare continuous variables, an ANOVA assuming equal variances or Spearman correlation test was performed.

For our primary hypothesis testing, we conducted initial analyses to select T-cell subset variables using Spearman correlation tests that assessed the correlations of *F. nucleatum* DNA amount with densities of T-cell subsets. To control for selection bias due to tissue data availability, we used the inverse probability weighting (IPW) method (24, 37), which used 4,465 incident colorectal cancer cases including the 933 cases with tissue data. First, we constructed a multivariable logistic regression model that had covariates as predictors and tissue data availability as an outcome variable. On the basis of the fitted regression model using all of the 4,465 cases, we calculated the probability of tissue availability in each case with a set of covariates. Then, each of 933 tissue data available cases was weighted by the inverse of the probability. Weights greater than the 95th percentile were truncated and set to the value of the 95th percentile to reduce outlier effects (37). We confirmed that results without weight truncation did not differ substantially from those with weight truncation (data not shown). The logistic regression analyses without IPW yielded similar results (Supplementary Table S1) to the IPW-adjusted model.

To control for confounding, multivariable IPW-adjusted logistic regression analyses initially included sex (female vs. male), age at diagnosis (continuous), year of diagnosis (continuous), family history of colorectal cancer in any first-degree relative (present vs. absent), tumor location (proximal colon vs. distal colon vs. rectum), MSI status (MSI-high vs. non–MSI-high), CIMP status (high vs. low/negative), LINE-1 methylation level (continuous), *KRAS* status (mutant vs. wild-type), *BRAF* status (mutant vs. wild-type), and *PIK3CA* status (mutant vs. wild-type). A backward elimination was conducted with a threshold *P* of 0.20 to select variables for the final models. Cases with missing data [family history of colorectal cancer in a first-degree relative (0.9%), tumor location (0.4%), MSI (2.9%), CIMP (7.0%), *KRAS* (2.9%), *BRAF* (2.0%), and *PIK3CA* mutation (8.5%)] were included in the majority category of a given categorical covariate to limit the degrees of freedom of the models. For the cases with missing data on LINE-1 methylation (2.7%), we assigned a

Table 1. Clinical, pathologic, and molecular characteristics of colorectal cancer cases according to *F. nucleatum* DNA amount in tumor tissue.

Characteristic[a]	All cases ($N = 933$)	*F. nucleatum* DNA in tumor tissue			P^b
		Negative ($n = 805$)	Low ($n = 64$)	High ($n = 64$)	
Sex					0.61
Female (NHS)	513 (55%)	439 (55%)	35 (55%)	39 (61%)	
Male (HPFS)	420 (45%)	366 (45%)	29 (45%)	25 (39%)	
Mean age ± SD (years)	69.1 ± 8.8	69.0 ± 8.9	70.6 ± 8.6	68.9 ± 8.0	0.37
Year of diagnosis					0.13
1995 or before	300 (32%)	268 (33%)	13 (20%)	19 (30%)	
1996–2000	307 (33%)	264 (33%)	20 (31%)	23 (36%)	
2001–2012	326 (35%)	273 (34%)	31 (48%)	22 (34%)	
Family history of colorectal cancer in first-degree relative(s)					0.69
Absent	730 (79%)	627 (78%)	51 (81%)	52 (83%)	
Present	195 (21%)	172 (22%)	12 (19%)	11 (17%)	
Tumor location					0.084
Cecum	166 (18%)	135 (18%)	16 (25%)	15 (23%)	
Ascending to transverse colon	300 (32%)	253 (32%)	23 (36%)	24 (38%)	
Descending to sigmoid colon	276 (30%)	251 (30%)	10 (16%)	15 (23%)	
Rectum	187 (20%)	162 (20%)	15 (23%)	10 (16%)	
Tumor differentiation					<0.0001
Well to moderate	846 (91%)	745 (93%)	52 (83%)	49 (77%)	
Poor	86 (9.2%)	60 (7.5%)	11 (17%)	15 (23%)	
AJCC disease stage					0.090
I	201 (23%)	184 (25%)	8 (14%)	9 (15%)	
II	284 (33%)	235 (32%)	22 (38%)	27 (44%)	
III	251 (29%)	214 (29%)	22 (38%)	15 (25%)	
IV	127 (15%)	111 (15%)	6 (10%)	10 (16%)	
MSI status					<0.0001
Non–MSI-high	750 (83%)	669 (86%)	42 (69%)	39 (61%)	
MSI-high	156 (17%)	112 (14%)	19 (31%)	25 (39%)	
CIMP status					<0.0001
Low/negative	708 (82%)	627 (84%)	46 (78%)	35 (58%)	
High	160 (18%)	122 (16%)	13 (22%)	25 (42%)	
Mean LINE-1 methylation level ± SD (%)	62.5 ± 9.6	62.2 ± 9.6	63.0 ± 9.3	64.8 ± 9.9	0.099
KRAS mutation					0.33
Wild-type	536 (59%)	466 (59%)	30 (51%)	40 (63%)	
Mutant	370 (41%)	318 (41%)	29 (49%)	23 (37%)	
BRAF mutation					0.0009
Wild-type	776 (85%)	680 (86%)	52 (85%)	44 (69%)	
Mutant	138 (15%)	109 (14%)	9 (15%)	20 (31%)	
PIK3CA mutation					0.95
Wild-type	715 (84%)	618 (84%)	46 (82%)	51 (84%)	
Mutant	139 (16%)	119 (16%)	10 (18%)	10 (16%)	
Neoantigen load					0.0002
Q1 (lowest)	107 (25%)	94 (26%)	6 (19%)	7 (23%)	
Q2	105 (25%)	91 (25%)	10 (31%)	4 (13%)	
Q3	106 (25%)	100 (28%)	4 (13%)	2 (6.7%)	
Q4 (highest)	106 (25%)	77 (21%)	12 (37%)	17 (57%)	
Exome-wide tumor mutation burden					0.001
Q1 (lowest)	107 (25%)	95 (26%)	6 (19%)	6 (20%)	
Q2	106 (25%)	95 (26%)	7 (22%)	4 (13%)	
Q3	105 (25%)	95 (26%)	7 (22%)	3 (10%)	
Q4 (highest)	106 (25%)	77 (21%)	12 (38%)	17 (57%)	
Median stromal $CD3^+$ cell density (IQR; cells/mm^2)	145 (22–493)	162 (26–538)	94 (6.8–302)	71 (6.9–238)	0.0002[c]
Median stromal $CD3^+CD4^+$ cell density (IQR; cells/mm^2)	76 (3.8–361)	84 (4.9–392)	44 (0–207)	37 (0–164)	0.0007[c]
Median stromal $CD3^+CD8^+$ cell density (IQR; cells/mm^2)	10 (0–54)	10 (0–56)	11 (0–56)	9.3 (0–34)	0.69[c]
Median stromal $CD3^+CD4^+CD45RO^+$ cell density (IQR; cells/mm^2)	57 (0–287)	67 (2.9–329)	33 (0–162)	24 (0–134)	0.0006[c]
Median stromal $CD3^+CD4^+CD45RO^-$ cell density (IQR; cells/mm^2)	7.2 (0–48)	7.7 (0–51)	0 (0–38)	4.2 (0–26)	0.034[c]
Median stromal overall macrophage density (IQR; cells/mm^2)	871 (489–1,428)	859 (479–1,399)	966 (579–1,667)	894 (609–1,477)	0.33[c]
Median stromal M1-like macrophage density (IQR; cells/mm^2)	175 (65–398)	170 (62–379)	202 (68–548)	192 (85–452)	0.14[c]
Median stromal M2-like macrophage density (IQR; cells/mm^2)	192 (69–429)	192 (69–435)	191 (69–435)	196 (63–431)	0.33[c]

(Continued on the following page)

Table 1. Clinical, pathologic, and molecular characteristics of colorectal cancer cases according to *F. nucleatum* DNA amount in tumor tissue. (Cont'd)

Characteristic[a]	All cases ($N = 933$)	*F. nucleatum* DNA in tumor tissue: Negative ($n = 805$)	Low ($n = 64$)	High ($n = 64$)	P[b]
Bifidobacterium genus DNA in tumor tissue					
Negative	631 (71%)	556 (72%)	38 (64%)	37 (62%)	0.094
Low	128 (14%)	104 (14%)	9 (15%)	15 (25%)	
High	128 (14%)	108 (14%)	12 (20%)	8 (13%)	

Abbreviations: AJCC, American Joint Committee on Cancer; IQR, interquartile range.
[a]Percentage indicates the proportion of patients with a specific clinical, pathologic, or molecular characteristic among all patients or in strata of *F. nucleatum* DNA amount.
[b]To compare categorical data between subgroups classified by *F. nucleatum* DNA amount, the χ^2 test was performed, unless otherwise noted. To compare continuous variables, an ANOVA was performed.
[c]To assess associations between *F. nucleatum* DNA amount (continuous) and densities of T-cell subsets and macrophage (continuous), the Spearman correlation test was performed.

separate indicator variable. We confirmed that excluding the cases with missing information in any of the covariates did not substantially alter results (data not shown). The proportional odds assumption was assessed using the ordinal logistic regression model using the ordinal categories of T-cell subset density (negative and tertile or quartile; as an ordinal outcome variable). We observed evidence of violation of this assumption in $CD3^+CD4^+$ cells, while this assumption for the other subsets was generally satisfied ($P > 0.07$). Therefore, for $CD3^+CD4^+$ cells, we used a binary variable dichotomized at the median value as the outcome variable in logistic regression analysis.

In secondary analyses, we assessed the statistical interaction between *F. nucleatum* status in colorectal cancer tissue (negative, low, and high) and MSI status (high vs. non-high), neoantigen load [high (Q3–Q4) vs. low (Q1–Q2)], or exome-wide tumor mutational burden [high (Q3–Q4) vs. low (Q1–Q2)] in relation to T-cell subset densities. We used the Wald test for the cross-product in multivariable-adjusted logistic regression models. We estimated the OR for a unit increase in the three ordinal categories of the amount of *F. nucleatum* in strata of MSI status, neoantigen load, or exome-wide tumor mutational burden using reparameterization of the interaction term in a single regression model (32). For interaction analyses, the proportional odds assumption in an ordinal logistic regression model was violated in subpopulations defined by MSI status, neoantigen loads, and exome-wide tumor mutational burden, therefore, we used binary outcome variables for all of the T-cell measurements.

As exploratory analyses, we performed Spearman correlation tests that assessed the correlation between *F. nucleatum* and TAMs densities. We also assessed the association of *Bifidobacterium* genus with T-cell subset densities.

Results

During the longitudinal follow-up of the two prospective cohort studies, we documented 4,465 incident colorectal cancer cases, including 933 cases with available data on *F. nucleatum* and T cells in tumor tissue. We used covariate data of the 4,465 cases to adjust for selection bias in the 933 cases in multivariable analyses to conduct our primary hypothesis testing. *F. nucleatum* DNA was detected using a quantitative PCR assay in 128 (14%) of the 933 cases. **Table 1** shows clinical, pathologic, and molecular features of colorectal cancer cases according to the amount of *F. nucleatum* DNA. Greater amounts of *F. nucleatum* DNA were associated with poor tumor differentiation, MSI-high status, CIMP-high status, BRAF mutation, higher tumor neoantigen load, and higher exome-wide tumor mutational burden ($P < 0.005$; with the α level of 0.005), but not with *KRAS* or *PIK3CA* mutation status.

Initial analyses using Spearman correlation test on each of the T-cell subset densities in tumor intraepithelial and stromal regions revealed that the densities of $CD3^+$ cells, $CD3^+CD4^+$ cells, and $CD3^+CD4^+CD45RO^+$ cells in tumor stromal areas were inversely correlated with the amount of *F. nucleatum* ($P < 0.005$ with the α level of 0.005; **Table 1** and **Fig. 1A**). In contrast, the amount of tissue *F. nucleatum* was not associated with intraepithelial densities of T-cell subsets. Representative multiplex immunofluorescence and cell phenotype images of *F. nucleatum*–negative, -low, and -high cases are shown in **Fig. 1B**.

In our primary hypothesis testing, we used a logistic regression analysis to assess the association of the amount of *F. nucleatum* DNA with the densities of T-cell subsets, which were selected by the initial analyses (**Table 2**; Supplementary Tables S1 and S2). In the multivariable analyses, the amount of *F. nucleatum* DNA in colorectal cancer tissue was inversely associated with density of stromal $CD3^+$ cells ($P_{trend} = 0.0004$) and $CD3^+CD4^+CD45RO^+$ cells ($P_{trend} = 0.003$) with the α level of 0.005. For a unit increase in four ordinal categories of stromal $CD3^+$ cell density, multivariable ORs were 0.51 [95% confidence interval (CI), 0.30–0.87] for *F. nucleatum*–low cases and 0.47 (95% CI, 0.28–0.79) for *F. nucleatum*–high cases, compared with *F. nucleatum*–negative cases. For a unit increase in four ordinal categories of stromal $CD3^+CD4^+CD45RO^+$ cell density, multivariable ORs were 0.67 (95% CI, 0.41–1.10) for *F. nucleatum*–low cases and 0.52 (95% CI, 0.32–0.85) for *F. nucleatum*–high cases, compared with *F. nucleatum*–negative cases. The amount of *F. nucleatum* in colorectal cancer tissue was not significantly associated with the density of $CD3^+CD4^+$ cells or $CD3^+CD4^+CD45RO^-$ cells (with the α level of 0.005).

As secondary analyses, we examined the statistical interaction between *F. nucleatum* and MSI status in relation to densities of stromal $CD3^+$ cells and $CD3^+CD4^+CD45RO^+$ cells. We did not observe a significant interaction between *F. nucleatum* and MSI status (**Table 3**).

Additional secondary analyses were performed using a subset of cases with available neoantigen load or exome-wide tumor mutational burden data as measured by whole-exome sequencing. Evaluation of the statistical interaction between *F. nucleatum* and neoantigen load (or exome-wide tumor mutational burden) in relation to densities of stromal $CD3^+$ cells and $CD3^+CD4^+CD45RO^+$ cells did not identify a

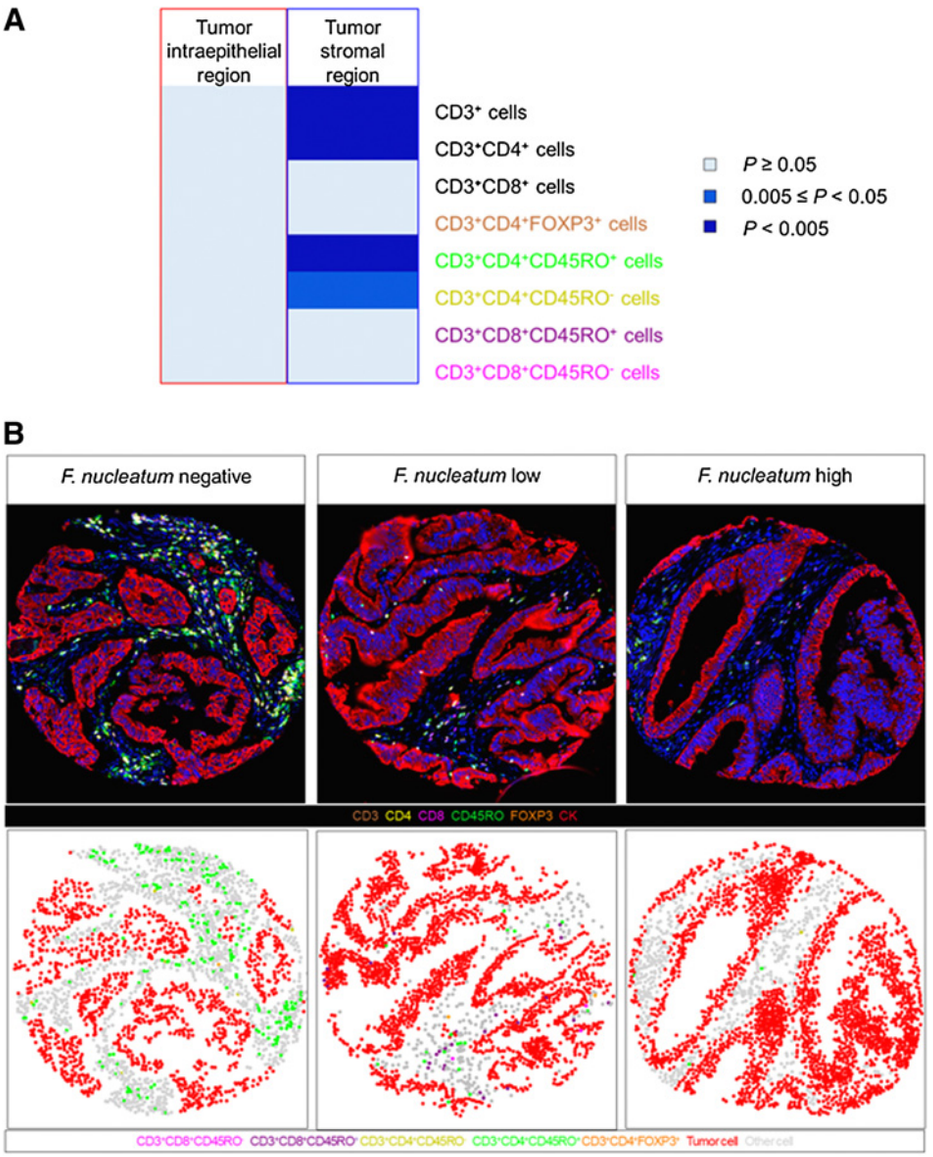

Figure 1.
Relationships between *F. nucleatum* and T-cell subsets in the colorectal cancer microenvironment. **A,** Correlation between *F. nucleatum* DNA amount in tumor tissue and density of T-cell subsets in tumor intraepithelial and stromal regions. **B,** Multiplex immunofluorescence and cell phenotype images for representative *F. nucleatum*–negative, -low, and -high case.

significant interaction between *F. nucleatum* and neoantigen load (**Table 4**) or exome-wide tumor mutational burden (**Table 5**).

In exploratory analyses, we did not observe a statistically significant association between *F. nucleatum* DNA amount and any TAM subset in tumor intraepithelial or stromal regions ($P > 0.01$; with the α level of 0.005; **Table 1**; Supplementary Table S3).

We further examined the association of *Bifidobacterium* genus with T-cell subset densities and did not observe a significant association between *Bifidobacterium* genus and any T-cell subset in intraepithelial or stromal regions ($P > 0.2$; Supplementary Table S4).

Discussion

As colorectal cancer is a group of heterogenous tumors influenced by the microbiota and immune system, we conducted this study utilizing a molecular pathologic epidemiology database (38) based on the two prospective cohort studies, to examine the relations between *F. nucleatum* and T-cell infiltrates while controlling for confounders and selection bias. By employing a quantitative, multiplexed immunofluorescence assay, we found an inverse association of *F. nucleatum* with tumor stromal $CD3^+$ T cells, particularly stromal $CD3^+CD4^+CD45RO^+$ memory helper T cells, that was independent of MSI status, neoantigen load, and exome-wide tumor mutational burden. To our knowledge, our analysis is the first human population study to show the relationship between *F. nucleatum* amount and memory helper T cells in tumor stroma. Our findings not only aid the understanding of *F. nucleatum*–associated immunosuppression, but also underscore the importance of *in situ* localization of specific T-cell subsets in the colorectal cancer microenvironment.

The tumor immune microenvironment comprises neoplastic cells, infiltrating immune cells, stromal cells, and extracellular matrix. Recent advances in digital pathology have revealed the importance of both the characterization and localization of immune cells (10). Multiplex immunofluorescence allows detailed phenotyping of immune cells, which improves understanding of tumor–immune

Table 2. IPW-adjusted logistic regression analysis to assess the association of *F. nucleatum* (predictor) with T-cell density (outcome).

	Univariable OR (95% CI)[a]	Multivariable OR (95% CI)[a,b]
Model for stromal $CD3^+$ cell density (as an ordinal outcome variable)		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.54 (0.33–0.91)	0.51 (0.30–0.87)
High	0.51 (0.30–0.84)	0.47 (0.28–0.79)
P_{trend}[c]	0.001	0.0004
Model for stromal $CD3^+CD4^+$ cell density (as a binary outcome variable)		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.86 (0.49–1.48)	0.80 (0.46–1.39)
High	0.52 (0.29–0.91)	0.50 (0.28–0.89)
P_{trend}[d]	0.023	0.015
Model for stromal $CD3^+CD4^+CD45RO^+$ cell density (as an ordinal outcome variable)		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.66 (0.40–1.09)	0.67 (0.41–1.10)
High	0.53 (0.33–0.85)	0.52 (0.32–0.85)
P_{trend}[c]	0.003	0.003
Model for stromal $CD3^+CD4^+CD45RO^-$ cell density (as an ordinal outcome variable)		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.73 (0.43–1.25)	0.74 (0.43–1.28)
High	0.75 (0.47–1.17)	0.75 (0.47–1.17)
P_{trend}[c]	0.12	0.13

[a]IPW was applied to reduce bias due to the availability of tumor tissue after cancer diagnosis (see "statistical analysis" subsection for details).
[b]The multivariable ordinal logistic regression model initially included age, sex, year of diagnosis, family history of colorectal cancer, tumor location, MSI, CIMP, LINE-1 methylation level, and *KRAS, BRAF*, and *PIK3CA* mutation status. A backward elimination with a threshold *P* of 0.20 was used to select variables for the final model. The variables which remained in the final models are shown in Supplementary Table S2.
[c]P_{trend} was calculated as a linear trend across the ordinal categories of *F. nucleatum* DNA amount (negative, low, and high; as an ordinal predictor variable) in the IPW-adjusted ordinal logistic regression model for T-cell density (four ordinal categories, as an ordinal outcome variable).
[d]To avoid violation of the proportional odds assumption, the density of $CD3^+CD4^+$ cells was dichotomized at the median value of subset. P_{trend} was calculated as a linear trend across the ordinal categories of *F. nucleatum* DNA amount (negative, low, and high; as an ordinal predictor variable) in the IPW-adjusted logistic regression model for the density of $CD3^+CD4^+$ cells (binary categories, as an outcome variable).

Table 3. IPW-adjusted logistic regression analysis to assess the association of *F. nucleatum* (predictor) with T-cell density (outcome) in strata of MSI status.

N = 906	Univariable OR (95% CI)[a]	Multivariable OR (95% CI)[a,b]
Model for stromal $CD3^+$ cell density[c] (as a binary outcome variable)		
Non–MSI-high		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.55 (0.27–1.11)	0.56 (0.28–1.15)
High	0.36 (0.16–0.79)	0.36 (0.16–0.80)
MSI-high		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.60 (0.21–1.68)	0.58 (0.20–1.66)
High	0.52 (0.21–1.31)	0.58 (0.23–1.46)
$P_{interaction}$[d]	0.53	0.47
Model for stromal $CD3^+CD4^+CD45RO^+$ cell density[c] (as a binary outcome variable)		
Non–MSI-high		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.56 (0.28–1.12)	0.55 (0.28–1.11)
High	0.43 (0.20–0.91)	0.42 (0.20–0.89)
MSI-high		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.76 (0.27–2.14)	0.72 (0.25–2.08)
High	0.57 (0.23–1.45)	0.62 (0.25–1.54)
$P_{interaction}$[d]	0.52	0.46

[a]IPW was applied to reduce bias due to the availability of tumor tissue after cancer diagnosis (see "statistical analysis" subsection for details).
[b]The multivariable logistic regression model initially included age, sex, year of diagnosis, family history of colorectal cancer, tumor location, CIMP, LINE-1 methylation level, and *KRAS, BRAF*, and *PIK3CA* mutation status. A backward elimination with a threshold *P* of 0.20 was used to select variables for the final model.
[c]To avoid violation of the proportional odds assumption, T-cell densities were dichotomized at the median value of each subset.
[d]$P_{interaction}$ (two-sided) was calculated using the Wald test for the cross-product of *F. nucleatum* DNA amount (negative, low, and high; as an ordinal predictor variable) and MSI status (high vs. non-high) in the IPW-adjusted logistic regression model.

interactions (13). Mature T cells, expressing CD3, are largely comprised of $CD8^+$ T cells and $CD4^+$ helper T cells. Within these two major classes, naïve and memory T cells can be distinguished on the basis of CD45RO expression, and regulatory T cells can be identified by expression of the FOXP3 transcription factor (39). While T cells play a major role in the adaptive immune response against cancer, specific T-cell subsets have divergent functions. Tumor-infiltrating lymphocytes, specifically $CD8^+$ T cells, represent the cytotoxic arm of the adaptive immune response and have been associated with better survival, while $CD4^+$ T cells (mostly in the helper T cell lineage) appear to enhance the antitumor activity of cytotoxic T cells (40, 41). Tumor stromal lymphocytes, which are generally more abundant than intraepithelial ones, have also been associated with favorable prognosis in colorectal cancer (42); however, different cell types, such as naïve, memory, and regulatory T cells, may have different implications. Our intriguing findings suggest the potential interplay between *F. nucleatum* and memory helper T cells in tumor stroma.

There is a growing body of evidence on the influential role of microbiota on cancer immunosurveillance and the modulation of immunotherapy responsiveness (43–46). *F. nucleatum* has emerged as a potentially oncogenic microorganism, which may drive inflammation-related carcinogenesis and recruit myeloid-derived suppressor cells (20). Experimental studies have also shown that *F. nucleatum* may inhibit the T-cell- and natural killer–cell–mediated immune

Table 4. IPW-adjusted logistic regression analysis to assess the association of *F. nucleatum* (predictor) with T-cell density (outcome) in strata of tumor neoantigen load.

N = 424	Univariable OR (95% CI)[a]	Multivariable OR (95% CI)[a,b]
Model for stromal $CD3^+$ cell density[c] (as an outcome variable)		
Neoantigen low		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.86 (0.26–2.80)	0.94 (0.28–3.13)
High	0.50 (0.12–2.18)	0.47 (0.10–2.23)
Neoantigen high		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.49 (0.16–1.49)	0.51 (0.17–1.55)
High	0.71 (0.27–1.90)	0.80 (0.29–2.25)
$P_{interaction}$[d]	0.71	0.85
Model for stromal $CD3^+CD4^+CD45RO^+$ cell density[c] (as an outcome variable)		
Neoantigen low		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.84 (0.26–2.73)	0.90 (0.27–2.98)
High	0.49 (0.11–2.12)	0.49 (0.11–2.13)
Neoantigen high		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.51 (0.17–1.58)	0.53 (0.18–1.60)
High	0.74 (0.28–1.98)	0.89 (0.32–2.50)
$P_{interaction}$[d]	0.88	0.99

[a]IPW was applied to reduce bias due to the availability of tumor tissue after cancer diagnosis (see "statistical analysis" subsection for details).
[b]The multivariable logistic regression model initially included age, sex, year of diagnosis, family history of colorectal cancer, tumor location, MSI, CIMP, LINE-1 methylation level, and *KRAS, BRAF*, and *PIK3CA* mutation status. A backward elimination with a threshold *P* of 0.20 was used to select variables for the final model.
[c]To avoid violation of the proportional odds assumption, T-cell densities were dichotomized at the median value of each subset.
[d]$P_{interaction}$ (two-sided) was calculated using the Wald test for the cross-product of *F. nucleatum* DNA amount (negative, low, and high; as an ordinal predictor variable) and neoantigen load [high (Q3–Q4) vs. low (Q1–Q2)] in the IPW-adjusted logistic regression model.

Table 5. IPW-adjusted logistic regression analysis to assess the association of *F. nucleatum* (predictor) with T-cell density (outcome) in strata of exome-wide tumor mutational burden.

N = 424	Univariable OR (95% CI)[a]	Multivariable OR (95% CI)[a,b]
Model for stromal $CD3^+$ cell density[c] (as an outcome variable)		
Exome-wide tumor mutational burden low		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	1.08 (0.32–3.64)	1.07 (0.30–3.76)
High	0.48 (0.11–2.07)	0.43 (0.09–2.09)
Exome-wide tumor mutational burden high		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.42 (0.15–1.18)	0.47 (0.17–1.33)
High	0.68 (0.26–1.80)	0.75 (0.28–2.06)
$P_{interaction}$[d]	0.68	0.77
Model for stromal $CD3^+CD4^+CD45RO^+$ cell density[c] (as an outcome variable)		
Exome-wide tumor mutational burden low		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	1.15 (0.34–3.87)	1.15 (0.34–3.97)
High	0.51 (0.12–2.20)	0.50 (0.12–2.19)
Exome-wide tumor mutational burden high		
Amount of *F. nucleatum* DNA		
Negative	1 (reference)	1 (reference)
Low	0.41 (0.14–1.14)	0.43 (0.15–1.22)
High	0.65 (0.25–1.71)	0.78 (0.29–2.15)
$P_{interaction}$[d]	0.71	0.82

[a]IPW was applied to reduce bias due to the availability of tumor tissue after cancer diagnosis (see "statistical analysis" subsection for details).
[b]The multivariable logistic regression model initially included age, sex, year of diagnosis, family history of colorectal cancer, tumor location, MSI, CIMP, LINE-1 methylation level, and *KRAS, BRAF*, and *PIK3CA* mutation status. A backward elimination with a threshold *P* of 0.20 was used to select variables for the final model.
[c]To avoid violation of the proportional odds assumption, T-cell densities were dichotomized at the median value of each subset.
[d]$P_{interaction}$ (two-sided) was calculated using the Wald test for the cross-product of *F. nucleatum* DNA amount (negative, low, and high; as an ordinal predictor variable) and exome-wide tumor mutational burden [high (Q3–Q4) vs. low (Q1–Q2)] in the IPW-adjusted logistic regression model.

response against colorectal cancer through the immune cell receptor TIGIT (21). In accordance with these experimental lines of evidence, our findings support the immunosuppressive role of *F. nucleatum* in the human colorectal cancer microenvironment. Interestingly, despite the ability of *F. nucleatum* to adhere and invade into tumor epithelial cells (19, 47), the densities of T cells within tumor epithelial regions were not correlated with *F. nucleatum* levels. Given that invasive *F. nucleatum* distribution appears highly heterogenous and focal in colorectal cancer in terms of gross tumor center versus invasive margin, as well as microscopic tumor epithelia versus stroma (48), *F. nucleatum* may interact with stromal components, including myeloid-derived suppressor cells, which inhibit T-cell proliferation and induce T-cell apoptosis (20). The amount of *F. nucleatum* in colorectal cancer has been associated with proximal tumor location, high-level MSI status, lower T-cell infiltrates, and worse prognosis (18, 49–51). Evidence also suggests that *F. nucleatum* may exert differential immunosuppressive or modulatory effects according to tumor MSI status (52). A better understanding of the interaction between *F. nucleatum* and specific immune cell phenotypes could have considerable implications on the development of therapeutic strategies that help overcome scarcity of T-cell infiltration and increase the fraction of patients responding to immunotherapy.

We acknowledge limitations of this study. First, considering the cross-sectional nature of our study, we cannot exclude the possibility of

reverse causation. Although it is possible that T cells may contribute to the elimination of *F. nucleatum*, our specific hypothesis was based on several lines of experimental evidence indicating that *F. nucleatum* suppresses the adaptive immune responses against colorectal cancer (21). Second, because our study was driven by a specific hypothesis, we focused on *F. nucleatum* in relation to T cells. Accumulating evidence suggests that various microbial species are involved in tumor development and the antitumor immune response (1, 2). Although we determined that there was no significant association between *Bifidobacterium* genus and T-cell densities as an exploratory analysis, more comprehensive bacterial analyses, such as metagenomic sequencing, would help further characterize the relationship between tumor microbiota and the antitumor immune response. Third, we used quantitative PCR assay for *F. nucleatum* in FFPE tissue specimens. Histopathology procedures and storage conditions may have influenced the detection rates and quantification. Nonetheless, our previous validation study using the quantitative PCR assay showed both a good concordance in detection of *F. nucleatum* in paired FFPE and frozen tissue specimens, as well as high linearity and reproducibility of *F. nucleatum* measurements in FFPE tissue specimens (18).

This study has notable strengths, including the use of a molecular pathologic epidemiology database derived from two U.S.-based large prospective cohort studies. Because no experimental model can recapitulate the complexity of human tumor immune microenvironment, which can be modified by various factors, including genetic and epigenetic alterations, lifestyle and environmental exposures, the microbiota, and host factors (53–55), the importance of the integrated data analyses on microbial features, tumor molecular characteristics, clinicopathologic findings, and immunologic profiling cannot be overemphasized. Second, our analyses within the prospective cohort studies enabled us to use the IPW method and covariate data of all 4,465 incident colorectal cancer cases to control for selection bias due to tissue data availability. Third, we utilized multiplexed immunofluorescence assays to identify and quantify specific subsets of T cells and macrophages in archival tumor tissue. In contrast to commonly used IHC, our assays enabled us to deeply subclassify immune cells and discover the link between *F. nucleatum* and memory helper T cells. Fourth, unlike common studies based on cases drawn from one or several hospitals, our study subjects were derived from hundreds of hospitals located throughout the United States, increasing the generalizability of our findings. Nevertheless, our findings need to be validated in independent studies.

In conclusion, this cross-sectional study utilizing two U.S.-wide prospective cohort studies has shown an inverse association of *F. nucleatum* DNA in colorectal carcinoma tissue with tumor stromal densities of $CD3^+$ cells and $CD3^+CD4^+CD45RO^+$ cells. These findings provide a compelling rationale for further investigations into the interplay of the microbiota and T lymphocytes in colorectal carcinoma, potentially leading to novel strategies for cancer prevention and therapy.

Authors' Disclosures

K. Fujiyoshi reports grants from Uehara Memorial Foundation and Grant of the Clinical Research Promotion Foundation during the conduct of the study. S.A. Väyrynen reports grants from Finnish Cultural Foundation and Orion Research Foundation sr during the conduct of the study. S. Bullman reports personal fees from BiomX outside the submitted work. K. Ng reports grants from NCI, Colorectal Cancer Alliance, Cancer Research UK, and Department of Defense during the conduct of the study. K. Ng also reports grants from Janssen, Revolution Medicines, Genentech, and Gilead Sciences; nonfinancial support from Pharmavite and Evergrande Group; and personal fees from Seattle Genetics, Array Biopharma, BiomX, and X-Biotix Therapeutics outside the submitted work. J.A. Meyerhardt reports personal fees from COTA Healthcare and Taiho Pharmaceutical outside the submitted work. G.J. Freeman reports grants from NCI during the conduct of the study, as well as personal fees from Roche, Bristol-Myers Squibb, Xios, OrigiMed, Triursus, iTeos, Nextpoint, IgM, Jubilant, GV20, and Trillium outside the submitted work. G.J. Freeman also reports a patent for CD274 (PD-L1)/PDCD1 (PD-1) pathway issued, licensed, and with royalties paid from Roche; a patent for CD274 (PD-L1)/PDCD1 (PD-1) pathway issued, licensed, and with royalties paid from Merck MSD; a patent for CD274 (PD-L1)/PDCD1 (PD-1) pathway issued, licensed, and with royalties paid from Bristol Myers Squibb; a patent for CD274 (PD-L1)/PDCD1 (PD-1) pathway issued, licensed, and with royalties paid from Merck KGA; a patent for CD274 (PD-L1)/PDCD1 (PD-1) pathway issued, licensed, and with royalties paid from AstraZeneca; a patent for CD274 (PD-L1)/PDCD1 (PD-1) pathway issued, licensed, and with royalties paid from Dako; a patent for CD274 (PD-L1)/PDCD1 (PD-1) pathway issued, licensed, and with royalties paid from Mayo Clinic; and a patent for CD274 (PD-L1)/PDCD1 (PD-1) pathway issued, licensed, and with royalties paid from Novartis. G.J. Freeman also reports equity in Nextpoint, Triursus, Xios, iTeos, IgM, GV20, and Geode. C. Huttenhower reports personal fees from Seres Therapeutics and Empress Therapeutics, as well as other from ZOE Nutrition outside the submitted work. A.T. Chan reports personal fees from Pfizer Inc. and Boehringer Ingelheim, as well as grants and personal fees from Bayer Pharma AG outside the submitted work. R. Nishihara reports other support from Pfizer outside the submitted work. C.S. Fuchs reports personal fees from Agios, Amylin Pharmaceuticals, AstraZeneca, Bain Capital, CytomX Therapeutics, Daiichi Sankyo, Eli Lilly, Entrinsic Health, EvolveImmune Therapeutics, Genentech, Merck, Taiho, and Unum Therapeutics outside the submitted work; C.S. Fuchs serves as a director of CytomX Therapeutics and owns unexercised stock options for CytomX Therapeutics and Entrinsic Health, is a cofounder for EvolveImmune Therapeutics and owns equity in this private company, and has provided expert testimony for Amylin Pharmaceuticals and Eli Lilly. M. Giannakis reports grants from Cancer Research UK during the conduct of the study, as well as grants from Bristol Myers Squibb, Merck, and Servier outside the submitted work. S. Ogino reports grants from NIH during the conduct of the study. No disclosures were reported by the other authors.

Disclaimer

The authors assume full responsibility for analyses and interpretation of these data. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the article.

Authors' Contributions

J. Borowsky: Conceptualization, data curation, formal analysis, investigation, methodology, writing–original draft, writing–review and editing. **K. Haruki:** Data curation, formal analysis, investigation, writing–original draft, writing–review and editing. **M.C. Lau:** Formal analysis, writing–original draft, writing–review and editing. **A. Dias Costa:** Data curation, methodology, writing–review and editing. **J.P. Väyrynen:** Data curation, formal analysis, investigation, methodology, writing–original draft, writing–review and editing. **T. Ugai:** Formal analysis, writing–review and editing. **K. Arima:** Data curation, investigation, writing–review and editing. **A. da Silva:** Data curation, investigation, writing–review and editing. **K.D. Felt:** Investigation, methodology, writing–review and editing. **M. Zhao:** Writing–review and editing. **C. Gurjao:** Data curation, formal analysis, writing–review and editing. **T.S. Twombly:** Writing–review and editing. **K. Fujiyoshi:** Writing–review and editing. **S.A. Väyrynen:** Writing–review and editing. **T. Hamada:** Writing–review and editing. **K. Mima:** Investigation, methodology, writing–original draft. **S. Bullman:** Investigation, methodology, writing–review and editing. **T.A. Harrison:** Funding acquisition, writing–review and editing. **A.I. Phipps:** Funding acquisition, writing–review and editing. **U. Peters:** Funding acquisition, writing–review and editing. **K. Ng:** Funding acquisition, writing–review and editing. **J.A. Meyerhardt:** Data curation, formal analysis, writing–review and editing. **M. Song:** Data curation, formal analysis, writing–review and editing. **E.L. Giovannucci:** Writing–review and editing. **K. Wu:** Writing–review and editing. **X. Zhang:** Writing–review and editing. **G.J. Freeman:** Writing–review and editing. **C. Huttenhower:** Funding acquisition, project administration, writing–review and editing. **W.S. Garrett:** Funding acquisition, project administration, writing–review and editing. **A.T. Chan:** Data curation, project administration, writing–review and editing. **B.A. Leggett:** Writing–review and editing. **V.L.J. Whitehall:** Writing–review and editing. **N. Walker:** Writing–review and editing. **I. Brown:** Writing–review and editing. **M. Bettington:** Writing–review and editing. **R. Nishihara:** Data curation, formal analysis, project administration,

writing–review and editing. **C.S. Fuchs:** Data curation, formal analysis, project administration, writing–review and editing. **J.K. Lennerz:** Conceptualization, supervision, writing–review and editing. **M. Giannakis:** Conceptualization, supervision, funding acquisition, writing–review and editing. **J.A. Nowak:** Conceptualization, data curation, supervision, investigation, methodology, writing–review and editing. **S. Ogino:** Conceptualization, data curation, formal analysis, supervision, funding acquisition, writing–original draft, project administration, writing–review and editing.

Acknowledgments

We would like to thank the participants and staff of the Nurses' Health Study and the Health Professionals Follow-up Study for their valuable contributions, as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, and WY. This work was supported by NIH grants (P01 CA87969 to M.J. Stampfer; UM1 CA186107 to M.J. Stampfer; P01 CA55075 to W.C. Willett; UM1 CA167552 to W.C. Willett; U01 CA167552 to W.C. Willett and L.A. Mucci; P50 CA127003 to C.S. Fuchs; R01 CA118553 to C.S. Fuchs; R01 CA169141 to C.S. Fuchs; R01 CA137178 to A.T. Chan; K24 DK098311 to A.T. Chan; R35 CA197735 to S. Ogino; R01 CA151993 to S. Ogino; K07 CA190673 to R. Nishihara; K07 CA188126 to X. Zhang; R37 CA225655 to J.K. Lennerz; R01 CA248857 to S. Ogino, U. Peters, and A.I. Phipps; and P50 CA101942 to G.J. Freeman), Cancer Research UK's Grand Challenge Award (UK C10674/A27140 to K. Ng, W.S. Garrett, M. Giannakis, C. Huttenhower, and S. Ogino), Nodal Award (2016-02) from the Dana-Farber Harvard Cancer Center (to S. Ogino and G.J. Freeman), a Stand Up to Cancer Colorectal Cancer Dream Team Translational Research grant (SU2C-AACR-DT22-17 to C.S. Fuchs and M. Giannakis), and grants from the Project P Fund, The Friends of the Dana-Farber Cancer Institute, Bennett Family Fund, and the Entertainment Industry Foundation through National Colorectal Cancer Research Alliance and SU2C. Stand Up to Cancer is a division of the Entertainment Industry Foundation. The indicated SU2C research grant is administered by the American Association for Cancer Research, the scientific partner of SU2C. J. Borowsky was supported by a grant from the Australia Awards-Endeavour Scholarships and Fellowships Program. K. Haruki was supported by fellowship grants from the Uehara Memorial Foundation and the Mitsukoshi Health and Welfare Foundation. K. Arima was supported by a grant from Overseas Research Fellowship (JP2018-60083) from Japan Society for the Promotion of Science. K. Fujiyoshi was supported by a fellowship grant from the Uehara Memorial Foundation. S.A. Väyrynen was supported by Finnish Cultural Foundation and Orion Research Foundation. J.A. Meyerhardt research was supported by the Douglas Gray Woodruff Chair fund, the Guo Shu Shi Fund, Anonymous Family Fund for Innovations in Colorectal Cancer, P fund, and the George Stone Family Foundation. M. Giannakis was supported by an ASCO Conquer Cancer Foundation Career Development Award. A.T. Chan is a Stuart and Suzanne Steele MGH Research Scholar.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received November 5, 2020; revised January 9, 2021; accepted February 19, 2021; published first February 25, 2021.

References

1. Rajpoot M, Sharma AK, Sharma A, Gupta GK. Understanding the microbiome: emerging biomarkers for exploiting the microbiota for personalized medicine against cancer. Semin Cancer Biol 2018;52:1–8.
2. Inamura K. Gut microbiota contributes towards immunomodulation against cancer: new frontiers in precision cancer therapeutics. Semin Cancer Biol 2020 Jun 21 [Epub ahead of print].
3. Chen B, Du G, Guo J, Zhang Y. Bugs, drugs, and cancer: can the microbiome be a potential therapeutic target for cancer management? Drug Discov Today 2019; 24:1000–9.
4. El Bairi K, Jabi R, Trapani D, Boutallaka H, Ouled Amar Bencheikh B, Bouziane M, et al. Can the microbiota predict response to systemic cancer therapy, surgical outcomes, and survival? The answer is in the gut. Expert Rev Clin Pharmacol 2020;13:403–21.
5. Kamal Y, Schmit SL, Frost HR, Amos CI. The tumor microenvironment of colorectal cancer metastases: opportunities in cancer immunotherapy. Immunotherapy 2020;12:1083–100.
6. Kather JN, Halama N. Harnessing the innate immune system and local immunological microenvironment to treat colorectal cancer. Br J Cancer 2019;120: 871–82.
7. Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. Science 2018;359:1350–5.
8. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 2015;372: 2509–20.
9. Ciardiello D, Vitiello PP, Cardone C, Martini G, Troiani T, Martinelli E, et al. Immunotherapy of colorectal cancer: challenges for therapeutic efficacy. Cancer Treat Rev 2019;76:22–32.
10. Pages F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet 2018;391:2128–39.
11. Nosho K, Baba Y, Tanaka N, Shima K, Hayashi M, Meyerhardt JA, et al. Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: cohort study and literature review. J Pathol 2010;222:350–66.
12. Marisa L, Svrcek M, Collura A, Becht E, Cervera P, Wanherdrick K, et al. The balance between cytotoxic T-cell lymphocytes and immune checkpoint expression in the prognosis of colon tumors. J Natl Cancer Inst 2018;110:68–77.
13. Gartrell RD, Marks DK, Hart TD, Li G, Davari DR, Wu A, et al. Quantitative analysis of immune infiltrates in primary melanoma. Cancer Immunol Res 2018; 6:481–93.
14. Brennan CA, Garrett WS. *Fusobacterium nucleatum* - symbiont, opportunist and oncobacterium. Nat Rev Microbiol 2019;17:156–66.
15. Luo K, Zhang Y, Xv C, Ji J, Lou G, Guo X, et al. *Fusobacterium nucleatum*, the communication with colorectal cancer. Biomed Pharmacother 2019;116:108988.
16. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. Genome Res 2012;22:292–8.
17. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. Genome Res 2012;22:299–306.
18. Mima K, Sukawa Y, Nishihara R, Qian ZR, Yamauchi M, Inamura K, et al. *Fusobacterium nucleatum* and T cells in colorectal carcinoma. JAMA Oncol 2015;1:653–61.
19. 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 2013;14:195–206.
20. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. Cell Host Microbe 2013;14:207–15.
21. Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, et al. Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack. Immunity 2015;42:344–55.
22. Mehta RS, Nishihara R, Cao Y, Song M, Mima K, Qian ZR, et al. Association of dietary patterns with risk of colorectal cancer subtypes classified by *Fusobacterium nucleatum* in tumor tissue. JAMA Oncol 2017;3:921–7.
23. Liao X, Lochhead P, Nishihara R, Morikawa T, Kuchiba A, Yamauchi M, et al. Aspirin use, tumor PIK3CA mutation, and colorectal-cancer survival. N Engl J Med 2012;367:1596–606.
24. Haruki K, Kosumi K, Hamada T, Twombly TS, Väyrynen JP, Kim SA, et al. Association of autophagy status with amount of *Fusobacterium nucleatum* in colorectal cancer. J Pathol 2020;250:397–408.
25. Yamauchi M, Morikawa T, Kuchiba A, Imamura Y, Qian ZR, Nishihara R, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. Gut 2012;61:847–54.
26. Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. N Engl J Med 2007;356:2131–42.
27. Fujiyoshi K, Väyrynen JP, Borowsky J, Papke DJ Jr, Arima K, Haruki K, et al. Tumour budding, poorly differentiated clusters, and T-cell response in colorectal cancer. EBioMedicine 2020;57:102860.

28. Lau MC, Borowsky J, Väyrynen JP, Haruki K, Zhao M, Costa AD, et al. Tumor-immune partitioning and clustering (TIPC) algorithm reveals distinct signatures of tumor-immune cell interactions within the tumor microenvironment. bioRxiv 2020.
29. Fujiyoshi K, Bruford EA, Mroz P, Sims CL, O'Leary TJ, Lo AWI, et al. Opinion: standardizing gene product nomenclature-a call to action. Proc Natl Acad Sci U S A 2021;118:e2025207118.
30. Väyrynen JP, Haruki K, Lau MC, Väyrynen SA, Zhong R, Dias Costa A, et al. The prognostic role of macrophage polarization in the colorectal cancer microenvironment. Cancer Immunol Res 2021;9:8–19.
31. Kosumi K, Hamada T, Koh H, Borowsky J, Bullman S, Twombly TS, et al. The amount of *Bifidobacterium* genus in colorectal carcinoma tissue in relation to tumor characteristics and clinical outcome. Am J Pathol 2018; 188:2839–52.
32. Nosho K, Irahara N, Shima K, Kure S, Kirkner GJ, Schernhammer ES, et al. Comprehensive biostatistical analysis of CpG island methylator phenotype in colorectal cancer using a large population-based sample. PLoS One 2008;3: e3698.
33. Imamura Y, Lochhead P, Yamauchi M, Kuchiba A, Qian ZR, Liao X, et al. Analyses of clinicopathological, molecular, and prognostic associations of KRAS codon 61 and codon 146 mutations in colorectal cancer: cohort study and literature review. Mol Cancer 2014;13:135.
34. Giannakis M, Mu XJ, Shukla SA, Qian ZR, Cohen O, Nishihara R, et al. Genomic correlates of immune-cell infiltrates in colorectal carcinoma. Cell Rep 2016;15: 857–65.
35. Nielsen M, Andreatta M. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. Genome Med 2016;8:33.
36. Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers EJ, Berk R, et al. Redefine statistical significance. Nat Hum Behav 2018;2:6–10.
37. Liu L, Nevo D, Nishihara R, Cao Y, Song M, Twombly TS, et al. Utility of inverse probability weighting in molecular pathological epidemiology. Eur J Epidemiol 2018;33:381–92.
38. Akimoto N, Ugai T, Zhong R, Hamada T, Fujiyoshi K, Giannakis M, et al. Rising incidence of early-onset colorectal cancer - a call to action. Nat Rev Clin Oncol 2020 Nov 20 [Epub ahead of print].
39. Golubovskaya V, Wu L. Different subsets of T cells, memory, effector functions, and car-T immunotherapy. Cancers (Basel) 2016;8:36.
40. Bou Nasser Eddine F, Ramia E, Tosi G, Forlani G, Accolla RS. Tumor immunology meets. . .immunology: modified cancer cells as professional APC for priming naive tumor-specific CD4+ T cells. Oncoimmunology 2017;6:e1356149.
41. Aspeslagh S, Morel D, Soria JC, Postel-Vinay S. Epigenetic modifiers as new immunomodulatory therapies in solid tumours. Ann Oncol 2018;29:812–24.
42. Haruki K, Kosumi K, Li P, Arima K, Vayrynen JP, Lau MC, et al. An integrated analysis of lymphocytic reaction, tumour molecular characteristics and patient survival in colorectal cancer. Br J Cancer 2020;122:1367–77.
43. Sears CL, Pardoll DM. The intestinal microbiome influences checkpoint blockade. Nat Med 2018;24:254–5.
44. Gopalakrishnan V, Helmink BA, Spencer CN, Reuben A, Wargo JA. The influence of the gut microbiome on cancer, immunity, and cancer immunotherapy. Cancer Cell 2018;33:570–80.
45. Zitvogel L, Pietrocola F, Kroemer G. Nutrition, inflammation and cancer. Nat Immunol 2017;18:843–50.
46. Murphy CL, O'Toole PW, Shanahan F. The gut microbiota in causation, detection, and treatment of cancer. Am J Gastroenterol 2019;114:1036–42.
47. Abed J, Emgard JE, Zamir G, Faroja M, Almogy G, Grenov A, et al. Fap2 mediates fusobacterium nucleatum colorectal adenocarcinoma enrichment by binding to tumor-expressed gal-GalNAc. Cell Host Microbe 2016;20:215–25.
48. Bullman S, Pedamallu CS, Sicinska E, Clancy TE, Zhang X, Cai D, et al. Analysis of *Fusobacterium* persistence and antibiotic response in colorectal cancer. Science 2017;358:1443–8.
49. Mima K, Nishihara R, Qian ZR, Cao Y, Sukawa Y, Nowak JA, et al. *Fusobacterium nucleatum* in colorectal carcinoma tissue and patient prognosis. Gut 2016;65: 1973–80.
50. Mima K, Cao Y, Chan AT, Qian ZR, Nowak JA, Masugi Y, et al. *Fusobacterium nucleatum* in colorectal carcinoma tissue according to tumor location. Clin Transl Gastroenterol 2016;7:e200.
51. Nosho K, Sukawa Y, Adachi Y, Ito M, Mitsuhashi K, Kurihara H, et al. Association of *Fusobacterium nucleatum* with immunity and molecular alterations in colorectal cancer. World J Gastroenterol 2016;22:557–66.
52. Hamada T, Zhang X, Mima K, Bullman S, Sukawa Y, Nowak JA, et al. *Fusobacterium nucleatum* in colorectal cancer relates to immune response differentially by tumor microsatellite instability status. Cancer Immunol Res 2018;6:1327–36.
53. Ogino S, Chan AT, Fuchs CS, Giovannucci E. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut 2011;60:397–411.
54. Ogino S, Nowak JA, Hamada T, Milner DA Jr, Nishihara R. Insights into pathogenic interactions among environment, host, and tumor at the crossroads of molecular pathology and epidemiology. Annu Rev Pathol 2019;14:83–103.
55. Wang ST, Cui WQ, Pan D, Jiang M, Chang B, Sang LX. Tea polyphenols and their chemopreventive and therapeutic effects on colorectal cancer. World J Gastroenterol 2020;26:562–97.