4-2020

# Association Between Sulfur-Metabolizing Bacterial Communities in Stool and Risk of Distal Colorectal Cancer in Men

Long H. Nguyen

Wenjie Ma

Dong D. Wang

Yin Cao

Himel Mallick

*See next page for additional authors*

## Authors

Long H. Nguyen, Wenjie Ma, Dong D. Wang, Yin Cao, Himel Mallick, Teklu K. Gerbaba, Jason Lloyd-Price, Galeb Abu-Ali, A. Brantley Hall, Daniel Sikavi, David A. Drew, Raaj S. Mehta, Cesar Arze, Amit D. Joshi, Yan Yan, Tobyn Branck, Casey DuLong, Kerry L. Ivey, Shuji Ogino, Eric B. Rimm, Mingyang Song, Wendy S. Garrett, Jacques Izard, Cutis Huttenhower, and Andrew T. Chan

# Association Between Sulfur-Metabolizing Bacterial Communities in Stool and Risk of Distal Colorectal Cancer in Men

Long H. Nguyen,[1,2,3] Wenjie Ma,[1,2] Dong D. Wang,[3,4] Yin Cao,[2,5,6] Himel Mallick,[3,7] Teklu K. Gerbaba,[8] Jason Lloyd-Price,[3,7] Galeb Abu-Ali,[3] A. Brantley Hall,[7] Daniel Sikavi,[2] David A. Drew,[1,2,9] Raaj S. Mehta,[1,2] Cesar Arze,[3] Amit D. Joshi,[1,2] Yan Yan,[3] Tobyn Branck,[3] Casey DuLong,[3] Kerry L. Ivey,[4,10] Shuji Ogino,[7,9,11,12] Eric B. Rimm,[4] Mingyang Song,[1,2,4,9] Wendy S. Garrett,[7,13,15] Jacques Izard,[8,14] Curtis Huttenhower,[3,7] and Andrew T. Chan[1,2,7,15]

1 Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts
2 Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts
3 Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts
4 Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts
5 Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, St Louis, Missouri
6 Siteman Cancer Center, Washington University School of Medicine, St Louis, Missouri;
7 Broad Institute of MIT and Harvard, Cambridge, Massachusetts
8 Department of Food Science & Technology, University of Nebraska, Lincoln, Nebraska
9 Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts

Published in *Gastroenterology* 158:5 (April 2020), pp1313–1325.
doi 10.1053/j.gastro.2019.12.029.

Submitted September 30, 2019; accepted December 24, 2019; published January 20, 2020.

10 South Australian Health and Medical Research Institute, Microbiome & Host Health Programme, Precision Medicine Theme, South Australia, Australia
11 Cancer Immunology and Cancer Epidemiology Programs, Dana-Farber Harvard Cancer Center, Boston, Massachusetts
12 Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts
13 Department of Medicine, Dana-Farber Cancer Institute, Boston, Massachusetts
14 Fred and Pamela Buffett Cancer Center, University of Nebraska Medical Center, Omaha, Nebraska
15 Department of Immunology and Infectious Disease, Harvard T.H. Chan School of Public Health, Boston, Massachusetts

Curtis Huttenhower and Andrew T. Chan share co-senior authorship.

*Corresponding authors* — Andrew T. Chan, MD, MPH, Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, 55 Fruit Street, GRJ-825C, Boston, Massachusetts 02114; *email* achan@partners.org; or Curtis Huttenhower, PhD, Department of Biostatistics, Harvard T.H. Chan School of Public Health, Associate Member, Broad Institute of MIT and Harvard, 677 Huntington Avenue, Boston, Massachusetts 02115; *email* chuttenh@hsph.harvard.edu.

## Abstract

**Background & Aims:** Sulfur-metabolizing microbes, which convert dietary sources of sulfur into genotoxic hydrogen sulfide ($H_2S$), have been associated with development of colorectal cancer (CRC). We identified a dietary pattern associated with sulfur-metabolizing bacteria in stool and then investigated its association with risk of incident CRC using data from a large prospective study of men.

**Methods:** We collected data from 51,529 men enrolled in the Health Professionals Follow-up Study since 1986 to determine the association between sulfur-metabolizing bacteria in stool and risk of CRC over 26 years of follow-up. First, in a subcohort of 307 healthy men, we profiled serial stool metagenomes and metatranscriptomes and assessed diet using semiquantitative food frequency questionnaires to identify food groups associated with 43 bacterial species involved in sulfur metabolism. We used these data to develop a sulfur microbial dietary score. We then used Cox proportional hazards modeling to evaluate adherence to this pattern among eligible individuals (n = 48,246) from 1986 through 2012 with risk for incident CRC.

**Results:** Foods associated with higher sulfur microbial diet scores included increased consumption of processed meats and low-calorie drinks and lower consumption of vegetables and legumes. Increased sulfur microbial diet scores were associated with risk of distal colon and rectal cancers, after adjusting for other risk factors (multivariable relative risk, highest vs lowest quartile, 1.43; 95% confidence interval 1.14–1.81; P-trend = .002). In contrast, sulfur microbial diet scores were not associated with risk of proximal colon cancer (multivariable relative risk 0.86; 95% CI 0.65–1.14; P-trend = .31).

**Conclusions:** In an analysis of participants in the Health Professionals Follow-up Study, we found that long-term adherence to a dietary pattern associated with sulfur-metabolizing bacteria in stool was associated with an increased risk of distal CRC. Further studies are needed to determine how sulfur-metabolizing bacteria might contribute to CRC pathogenesis.

**Keywords:** Colorectal Carcinogenesis, Cancer Biogeography, Fecal Microbes, FFQ

**Abbreviations**: **BMI**, body mass index; **CRC**, colorectal cancer; **EC**, Enzyme Commission; **FFQ**, food frequency questionnaire; **GI**, gastrointestinal; $H_2S$, hydrogen sulfide; **HPFS**, Health Professionals Follow-up Study; **IBD**, inflammatory bowel disease; **MLVS**, Men's Lifestyle Validation Study.

## What you need to know

### Background and context

Sulfur-metabolizing microbes, which convert dietary sources of sulfur into genotoxic hydrogen sulfide, have been associated with development of colorectal cancer (CRC).

### New findings

In an analysis of participants in the Health Professionals Follow-up Study, we found that long-term adherence to a dietary pattern associated with sulfur-metabolizing bacteria in stool was associated with an increased risk of distal CRC.

### Limitations

This study analyzed diets and stool from male health professionals. Studies among different populations are needed, as well as investigations to identify the mechanisms through which sulfur-producing microbes might promote colorectal carcinogenesis.

### Impact

Foods contributing to high sulfur microbial diet scores, including increased intake of processed meats and low-calorie drinks and fewer vegetables and legumes, are associated with development of distal CRC.

In the United States, colorectal cancer (CRC) is the third most frequently occurring and third most lethal form of cancer,[1] with most new diagnoses occurring among those without a clear genetic predisposition.[2] Thus, disease prevention through changes in lifestyle, such as dietary intake, is a high priority.[3] Diet has been convincingly shown to be a determinant in the taxonomic makeup and metabolic activities of the human gut microbiome,[4,5] but whether the well-established diet-CRC relationship is mediated through alterations in specific microbes and patterns of community metabolism remains vastly underexplored.[6]

Substantial experimental evidence links the presence and activity of sulfur-metabolizing microbes with increased microbial production of hydrogen sulfide ($H_2S$). Sulfur-metabolizing microbes, most often bacteria, are a specialized group of phylogenetically diverse microbes with the capacity to metabolize organic compounds for energy, often while reducing dietary sulfur to $H_2S$.[7–9] These bacteria and their resultant metabolites have been associated with CRC,[10,11] whereas $H_2S$ itself may cause epithelial DNA damage,[8] and promote alterations in immune cell populations associated with inflammation[7,11] and CRC.[8,12,13] Further, gut-derived $H_2S$ may fragment the mucus bilayer of the gastrointestinal tract. This barrier is typically held together by disulfide bonds that may be broken by excess $H_2S$. This breach may precede tumorigenesis by exposing gut epithelium to immunogenic luminal bacteria.[14–16]

Distinct foods may serve as critical inputs upstream of this process. Diets rich in processed animal meats, often at the expense of fiber sources, such as fruits and vegetables, likely provide proinflammatory sulfur-containing amino acids more conducive to the proliferation of these harmful bacteria.[4] This pattern of intake has previously been associated with increased CRC risk.[17–21] Conversely, a diet enriched with legumes and other vegetables may be associated with decreased populations of sulfur-metabolizing bacteria.[22–24] Further, these foods are a rich source of glucosinolates, sulfur-containing compounds with anti-inflammatory and possibly cancer preventive properties, [25–27] and have previously been associated with both a reduction in risk of CRC and precancerous adenomas.[19,26,28]

Thus, we performed a novel 2-stage study to clarify the role of sulfur-metabolizing bacteria in CRC. First, in a developmental subcohort of 307 men from the Men's Lifestyle Validation Study (MLVS) nested

within the Health Professionals Follow-up Study (HPFS) with diet and longitudinal stool sampling with next-generation sequencing, we generated the sulfur microbial diet, a de novo pattern of dietary constituents associated with the enrichment of sulfur-metabolizing bacteria in stool. We then used the larger HPFS as a testing cohort to prospectively associate long-term adherence to this dietary pattern with risk of incident CRC among participants for whom we had collected detailed information on long-term diet and other relevant exposures.

## Methods

### *Overall Study Population*

The HPFS is an ongoing prospective cohort study of 51,529 US male podiatrists, dentists, physicians, veterinarians, pharmacists, and optometrists aged 40 to 75 years at enrollment in 1986. Participants have been followed since inception with detailed biennial questionnaires on medical, lifestyle, and other health-related information. Dietary intake was assessed every 4 years through a semiquantitative food frequency questionnaire (FFQ). Follow-up among eligible subjects exceeds 90% of available person-time.

Nested within the HPFS, the MLVS was established among generally healthy participants, specifically excluding those with a prior history of coronary heart disease, prior cerebrovascular events, nonmelanoma cancer, or major neurologic comorbidities. Both cohorts have previously been described extensively. [29,30] We recruited the 307 male individuals in the MLVS who provided longitudinal stool samples between July 2012 and July 2013 (**Figure 1**).

### *Developmental cohort (MLVS)*

**Sample and data collection.** MLVS participants were asked to provide stool samples from 2 consecutive bowel movements 24 to 72 hours apart, followed approximately 6 months later by collection of a second, similarly paired set of samples. The collection protocol used has previously been detailed and validated to impart minimal perturbative effect from at-home collection of gut metagenomes and metatranscriptomes compared with fresh-frozen sample collections.[31–33] Each

bowel movement was placed into a container with RNA later. Participants completed a questionnaire detailing the date and time of evacuation, Bristol stool scale, and other relevant metadata. Paired samples were stored at ambient temperature, at which point both were sent overnight to the Massachusetts General Hospital and the Harvard T.H. Chan School of Public Health and held in –80-C freezers until nucleic acid extraction for subsequent sequencing at the Broad Institute of MIT and Harvard. To obtain metagenomes and metatranscriptomes, we used the Illumina HiSeq paired-end (2 × 101 nucleotides) shotgun sequencing platform. RNA was extracted and sequenced from the subset of participants who provided stool during both sampling periods (initial and 6 months later) and did not report the use of antibiotics within the past year. One individual was excluded from our analytic cohort after study enrollment because of prior history of total colectomy.

**Sequence bioinformatics.** Taxonomic and functional profiles were generated using the bioBakery shotgun metagenome workflow v0.9.0.[34] Shotgun metagenomes yielded relative taxonomic profiles using MetaPhlAn2 v2.6.0.[35] Functional profiling was done using HUMAnN2 v0.11.0.[36] For reads that did not map at the nucleotide level, a subsequent translated search was performed against a UniRef90-based protein sequence catalog,[37] resulting in final gene family abundance tables for both metagenomic and metatranscriptomic profiles, stratified by species contribution.

These gene families were then assembled into higher-order groupings, such as Enzyme Commission (EC) numbers. Species transcriptional activity was also quantified by summing the total sum-normalized stratified abundance attributed to each organism with the expression ratio defined as the ratio between species-stratified metatranscriptomic-to-metagenomic functional profiles. Only samples with greater than 1 read per kilobase of transcript per million mapped reads were used in downstream analysis.

**Identification of sulfur-modifying enzymes and sulfur-metabolizing bacteria.** First, we excluded microbial species that did not surpass minimum prevalence (10% of samples) and abundance (0.1% relative abundance) thresholds in human stool. We then used 2 complementary approaches to catalog sulfur-metabolizing bacteria: (1) a

MetaCyc version 22 pathway search[38,39] for at least 2 reactions generating $H_2S$ in the bacterial pangenome, and (2) a comprehensive literature review. We derived relevant sulfur-metabolism enzymes and their EC numbers by searching BRENDA release 2018.1[39,40] and MetaCyc version 22.[38,39] "Hydrogen sulfide" with exact and synonymous matching was used to identify relevant reactions. In addition, to prevent the inclusion of promiscuous enzymes not intimately involved in sulfur metabolism, we further refined our enzyme list to those that actively participate in sulfur group modification, for example, of a thiol (-SH) to a sulfide (-S-), and required that most of the reactions they catalyze to be sulfur metabolism reactions. The curated EC lists from BRENDA and MetaCyc were then merged and used for agnostic analysis (Supplementary Table 1).

A given species was considered a sulfur-metabolizing bacterium if representative strains disclosed reactions that generated $H_2S$ or if there was prior, high-quality experimental evidence supporting that classification. Membrane, transport, unknown direction, and reversible $H_2$-generating reactions were excluded from the analysis. In some cases, microbial generation of $H_2S$ was deemed unlikely if robust experimental evidence of nonproduction, product uptake, or product use in biotransformation was observed. For some bacteria, $H_2S$ generation was extended from species to higher taxonomic levels if prior studies indicated this to be a core function among members. For example, because *Odoribacter splanchnicus* is known to produce $H_2S$,[41] this designation was extended to all members of the *Odoribacter* genus. *Veillonella atypica*, *Veillonella parvula*, and *Veillonella* unclassified were each categorized similarly.[42]

**Dietary information and the derivation of the sulfur microbial diet score.** The validity and reproducibility of the FFQ used in this cohort have been previously reported. [43] Briefly, the FFQ includes 131 food items with specified serving sizes using common portions (e.g., 1 orange or 2–3 celery sticks) or standard weight and volume measurements. For each item, participants indicated their average frequency of consumption over the past year with regard to serving size and frequency ranging from "almost never" to "≥6 times/day," which was then converted to servings/day. In the HPFS, the FFQ has been administered every 4 years since inception.

For the first stage, participants in the MLVS microbiome collection served as the developmental cohort to assess the dietary predictors of sulfur-metabolizing bacteria abundance. These participants were administered 2 additional FFQs (6 months apart). To dampen measurement errors from random interindividual variance, their responses across both FFQs were averaged and used for de novo dietary derivation. After collapsing food items into servings/day for 40 predefined food groups formed on the basis of culinary usage and nutrient profiles, consistent with prior methods,[44] we used reduced rank regression models and stepwise linear regression analyses to identify a dietary pattern of intake most predictive of the log-transformed abundance of our bacterial species of interest.

### *Testing Cohort (HPFS)*

Assessing long-term adherence to the sulfur microbial diet. In the second stage of our study, we calculated sulfur microbial diet scores for each participant in the much larger HPFS based on nearly 3 decades of diet data (1986–2010) by summing the intake of foods retained from the final stepwise linear regression analyses weighted by their regression coefficients. To represent long-term usual dietary habits, sulfur microbial diet scores were updated at each follow-up cycle using the cumulative average method, with each score averaged across all assessments before the current questionnaire. The food-based sulfur microbial diet score represents a data-driven prediction for how much sulfur-metabolizing bacteria an individual may harbor over the long-term.

**Colorectal cancer ascertainment.** The primary endpoint was incident CRC. Participants prospectively reported new CRC diagnoses on biennial questionnaire or were identified through reporting from family, postal authorities, or the National Death Index. Physicians blinded to risk factor status reviewed relevant medical records for case confirmation and to retrieve data on anatomic site, histologic features, and stage of presentation.

**Assessment of covariates.** Height and weight were reported at study inception, and weight was updated biennially. Body mass index (BMI) was calculated as weight in kilograms/ height in meters$^2$. Physical activity was self-reported using validated questionnaires every 2 to 4

years.[45] We also assessed and updated the age they started or stopped smoking, number of cigarettes smoked daily, family history of CRC among first-degree relatives, regular use of aspirin, prior health care engagement (visit to a care provider in the past 2 years), and prior history of lower gastrointestinal endoscopy.

### *Statistical Analysis*

At baseline, we excluded participants with CRC, inflammatory bowel disease (IBD), or with missing information on dietary intake. We also excluded individuals reporting implausible energy intake (<800 or >4200 kcal/day). In total, 48,246 subjects comprised our final study population. Follow-up time accrued from study enrollment until the date of CRC diagnosis, death from any cause, or the end of follow-up (January 31, 2012), whichever occurred first. We used Cox proportional hazards models to estimate age and multivariable-adjusted relative risks and their 95% confidence intervals (CIs). Schoenfield residual testing confirmed no violations of the proportional hazards assumption. Covariates were chosen a priori and updated on a time-varying basis among major CRC risk factors and confounders, including age (continuous), family history of CRC (yes/no), BMI (quartiles), physical activity (metabolic equivalent task hours/week, quartiles), smoking (categories, never, past, current: 1–14, 15–25, and >25 cigarettes per day), regular aspirin use (yes/no), total caloric intake (continuous), prior endoscopy within the past 2 years, and physical examination/health care engagement in the past 2 years (each yes/no). For missing data, we carried forward nonmissing covariate data from 1 previous data cycle. SAS version 9.4 (SAS, Inc., Cary, NC) and R 3.5.1 (Vienna, Austria) were used for all statistical analyses.

### *Regulatory Compliance and Data Availability*

Participant recruitment and study-related protocols were approved by the Harvard T. H. Chan School of Public Health Institutional Review Board #HSPH 22067-102. Informed consent was implied through voluntary return of study questionnaires and bio specimens. Sequencing data have been deposited in the Sequence Read Archive under BioProject ID: PRJNA354235. Non–sequencing-based cohort data may be obtained with written request.

## Results

### *Developmental Cohort (MLVS)*

In our nested cohort of 307 men, mean age was 70.6 ± 4.3 years at the time of first stool collection. They generally did not smoke, consumed alcohol in moderation, and stool was typically of normal consistency by Bristol Stool Score (Supplementary Table 2).[46] In the 12 months before sampling, 26.6% of participants had been exposed to antibiotics, and 5.2% underwent bowel preparation within the prior 2 months. We found no clinically important differences between those who elected to participate in the MLVS stool sample collection and those who did not on the basis of age (mean age 71 years), ethnicity (each 97% white), BMI (25.7 vs 25.6 kg/m$^2$), physical activity (46 vs 50 metabolic equivalent of tasks hours/wk), and smoking status (1.5 vs 1% current smokers).

DNA was extracted from 925 stool samples. RNA was reverse transcribed to complementary DNA for the 340 samples from participants who provided stool at both sampling periods and did not report the use of antibiotics within the past year (**Figure 1**). Before and after computational quality control, sequencing depth was 3.8 ± 1.6

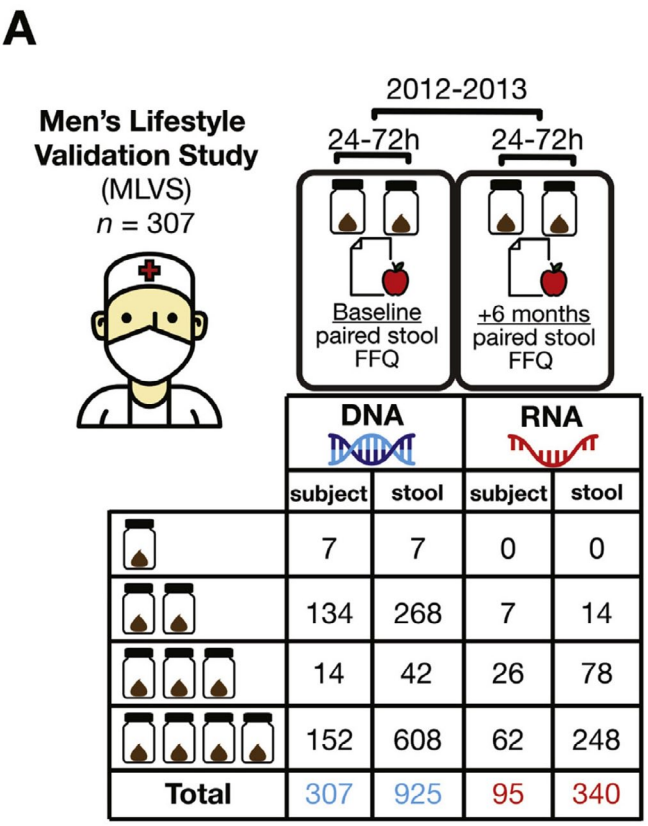
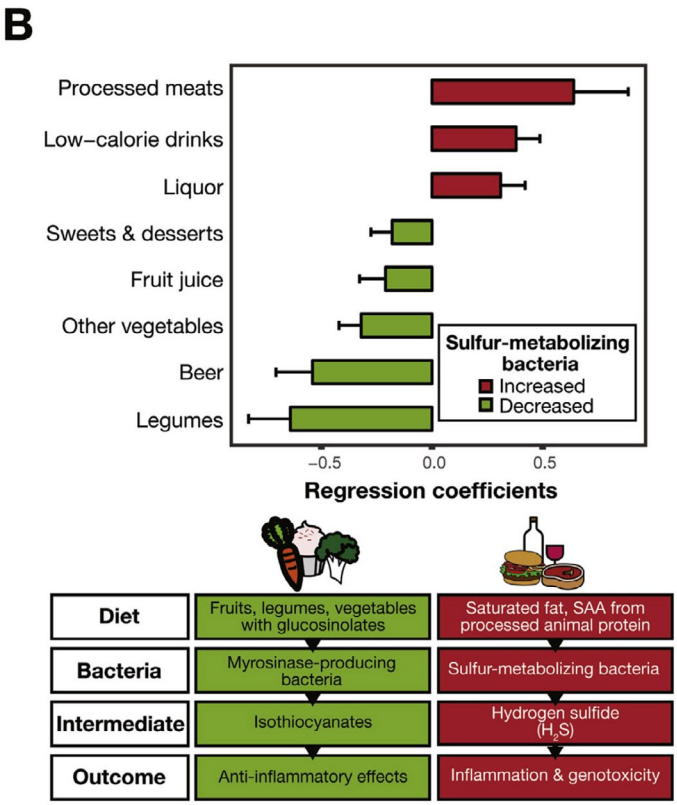
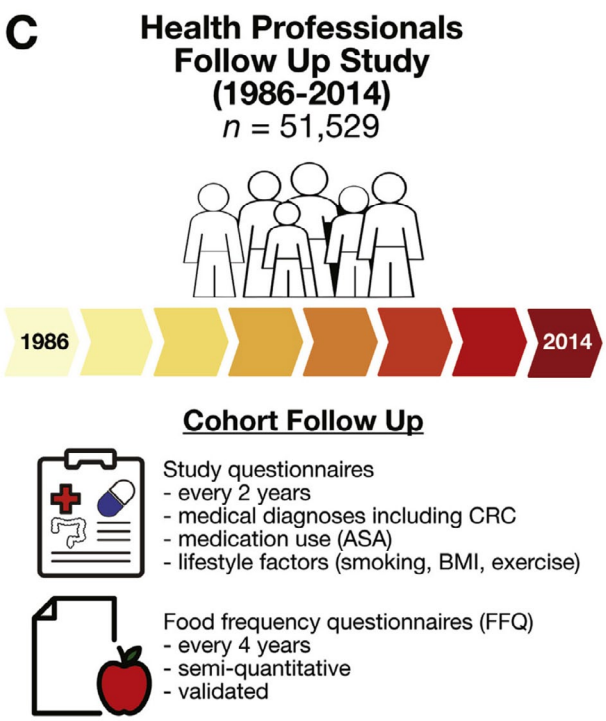



**Figure 1.** Experimental design.
**(A)** Study population and sampling details. Participants in the MLVS provided up to 4 stool samples over a 6-month study period with measurement of dietary intake via FFQ, identical to the FFQ given to participants in the HPFS. Stool underwent metagenomic and metatranscriptomic sequencing.




**(B)** Creation of the sulfur microbial diet. In the first stage, among MLVS participants with longitudinal stool metagenomes and FFQs, we used supervised clustering and regression techniques to determine the foods most commonly associated with increased abundance of sulfur-metabolizing bacteria to generate the sulfur microbial diet score.




**(C)** Predicted microbial carriage and risk of CRC. In the second stage, leveraging access to the much larger HPFS cohort with diet assessed every 4 years since inception, as well as other factors that may confound the relationship between diet and CRC, we calculated sulfur microbial diet scores in all 48,246 eligible participants of the HPFS, with higher scores reflecting closer adherence to a diet predicted to enrich for sulfur-metabolizing bacteria. ASA, aspirin or acetylsalicylic acid; SAA, sulfur-containing amino acids.

giganucleotides (Gnt) and 1.8 ± 0.7 Gnt for DNA and 2.8 ± 2.4 Gnt and 1.2 ± 1.0 Gnt for RNA, respectively. We identified and retained 139 bacterial species after minimum prevalence and abundance filtering, 43 of which were deemed sulfur-metabolizing microbes on the basis of both prior supportive experimental evidence plus the presence of at least 2 sulfur-modifying enzymes in their respective pangenomes (Supplementary Table 3).

We found modest concordance between the presence of sulfur-metabolizing enzymes found in stool metagenomes (DNA) with their downstream expression (RNA from metatranscriptomes; **Figure 2**). This suggests that for reactions catalyzed by the subset of enzymes used to define sulfur-metabolizing microbes, the presence of species that encode for these gene products is generally reflective of and a reasonable proxy for the underlying sulfur-metabolizing activity of the human gut in stool.

Using the FFQs collected most proximate to biospecimen collection, the sulfur microbial diet score was calculated using the weighted sum of the 7 food groups retained after stepwise linear regression and deemed most predictive of the log-transformed abundance of our 43 sulfur-metabolizing species (Supplementary Table 4). The component food groups were processed meat, liquor, and low-calorie drinks (each positively associated with the relative abundance of sulfur-metabolizing bacteria), as well as beer, fruit juice, legumes, mixed (other) vegetables, and sweets/desserts (each negatively associated).

We found notable relationships between sulfur microbial diet scores and the abundance of 2 sulfur-metabolizing bacteria previously identified as important microbes in the CRC microbiome. Specifically, sulfur microbial diet scores were associated with relative enrichments for *Erysipelotrichaceae bacterium 21_3* and *Bilophila wadsworthia* (**Figure 3**). *E bacterium 21_3*, of the Firmicutes phylum, has previously been shown to be increased in mice fed high-fat diets,[47–50] as well as in patients with CRC.[47] Several *Bilophila* spp are overrepresented in CRC and precursor colonic adenoma tissue, particularly from African American patients.[51,52]

Some sulfur-metabolizing enzymes were carried and expressed only by a single species, such as *B wadsworthia* and EC 1.8.99.3 (hydrogen sulfite reductase), a major contributor to dissimilatory sulfite reduction and an important proxy for microbially driven $H_2S$ production

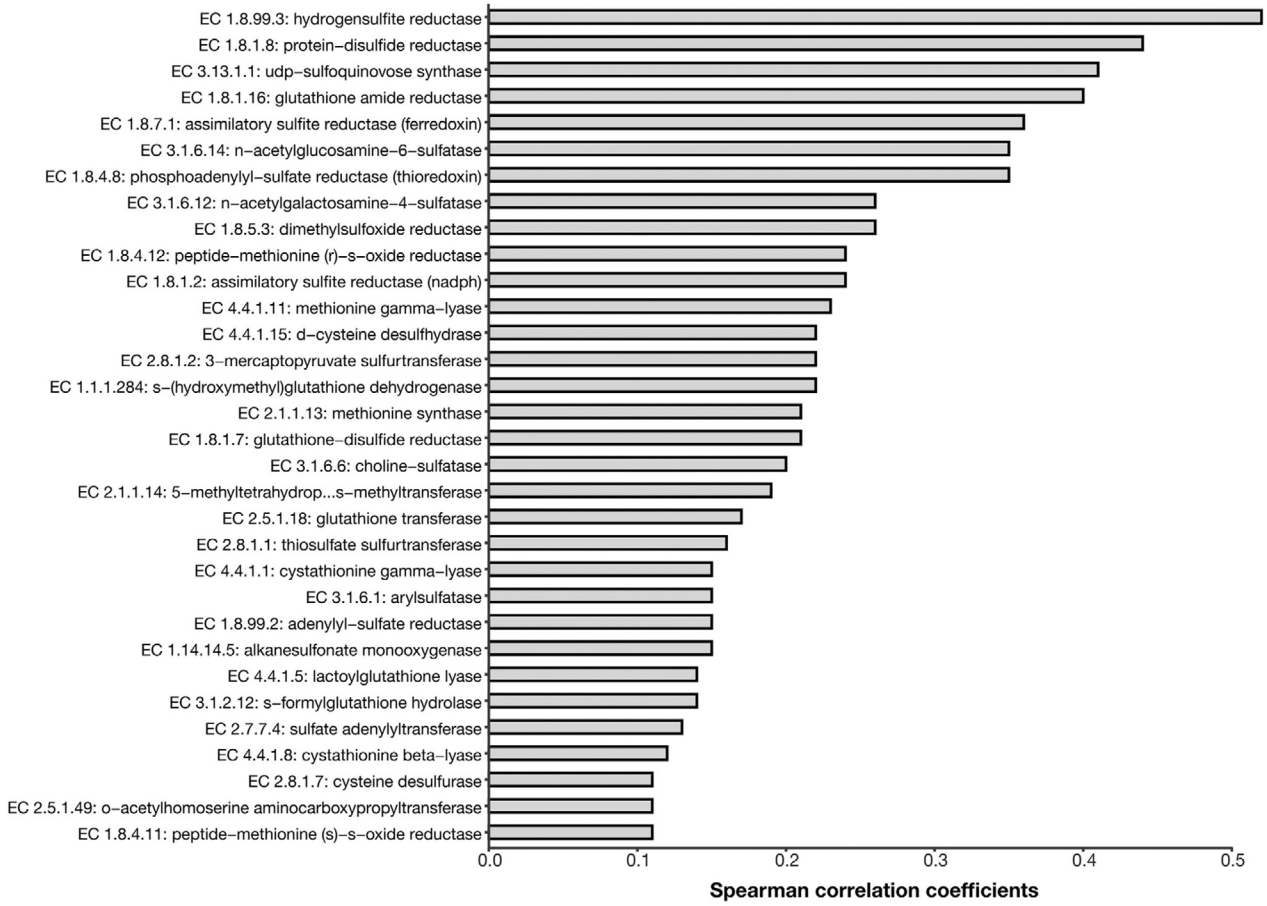



**Figure 2.** Bar plot of correlation between sulfur-related enzyme functional potential (DNA) and functional activity (RNA). Modest concordance was observed between observed metagenomes and their downstream metatranscriptomes among enzymes intimately involved in sulfur modification and sulfur metabolism. This suggests that for reactions catalyzed by our subset of enzymes, the sulfur microbial diet trained on the species that encode for these gene products is reflective of the underlying sulfur-metabolizing activity of the human gut in stool. Only enzymes represented in both DNA and RNA with $P < .05$ shown.

(**Figure 4**).[53] Conversely, enzymes such as EC 2.7.7.4 (sulfate adenylyltransferase), which catalyzes reactions to metabolize dietary sulfur and purine, were more widely shared among phylogenetically diverse microbes, including those from the *Parabacteroides* genera, *Ruminococcus* spp, and *Desulfovibrio desulfuricans*. This may indicate a community-level response to substrate availability through enrichment of microbes capable of metabolizing the sulfur- and purine-rich

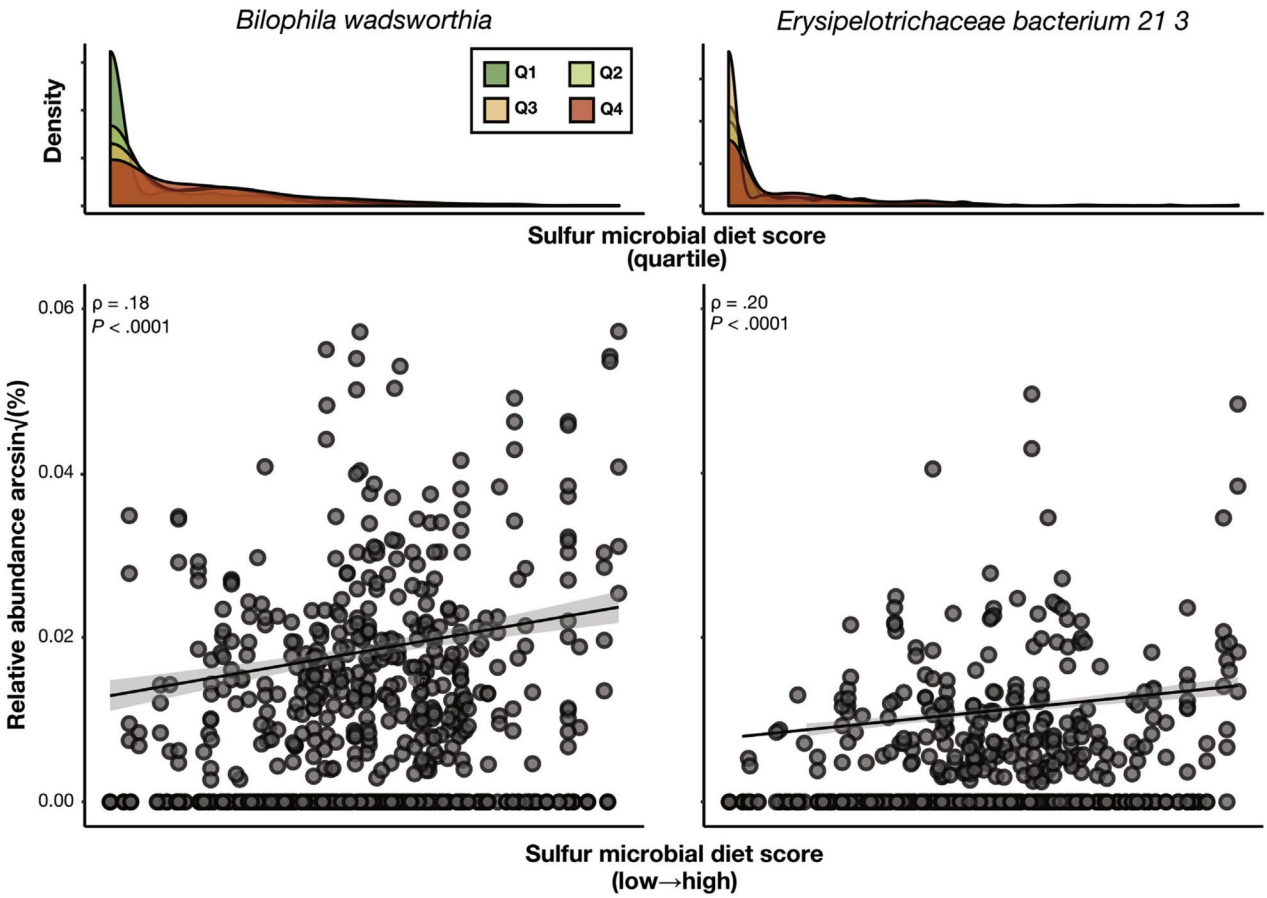



**Figure 3.** The association between the sulfur microbial diet and 2 representative sulfur-metabolizing bacteria. Higher sulfur microbial diet scores were associated with a relative enrichment of 2 sulfur-metabolizing microbes previously implicated in dysbiosis-associated CRC, *Bilophila wadsworthia* and *Erysipelotrichaeceae bacterium*. Trend line fit to nonzero data.

processed meats that characterize the sulfur microbial diet. In general, for sulfur-metabolizing enzymes, microbial transcriptional activity was highly correlated with their metagenomic functional capacities, that is, the greater number of species encoding for a given sulfur-metabolizing enzyme at the DNA level, the greater the number of species that will contribute to that enzyme's RNA level. This observation indicates the biological importance of these enzymes: when present, they are expressed. This supports the need to evaluate the entire community of microbes involved in sulfur economy, rather than several in comparative isolation.

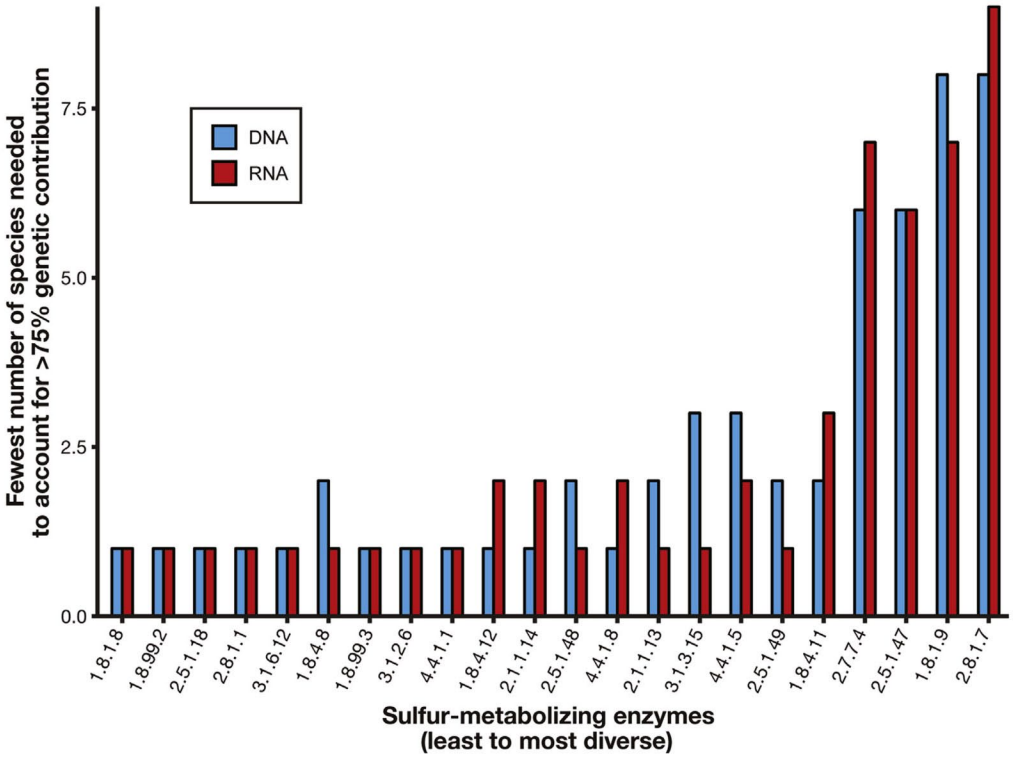



**Figure 4.** Contributional genetic diversity among sulfur-metabolizing enzymes by sulfur-metabolizing bacteria. Enzymes are arranged along the x-axis from least to greatest number of attributable species encoding for it at the DNA level (EC 1.8.1.8 = 1 species, EC 2.8.1.7 = 36 species). In general, the greater number of species encoding for a given enzyme (DNA), the greater the number of species that will contribute to that enzyme's functional activity or transcript level. Only ECs with 1 or more assignable taxon for both DNA and RNA are shown.

### *Testing Cohort (HPFS)*

After developing our de novo dietary pattern among individuals with dietary inventories during longitudinal stool sampling, we calculated sulfur microbial diet scores for each participant in the much larger HPFS based on their FFQs collected serially from 1986 to 2010. We found that participants more closely adhering to the sulfur microbial diet tended to have a slightly higher BMI, more frequently smoked (currently and in the past), and were more likely to be regular users of aspirin (**Table 1**). We sought to ensure the sulfur microbial diet was not serving as a surrogate measure for a Western-style diet, a diet characterized by high-fat intake and deficient in fiber, and a pattern

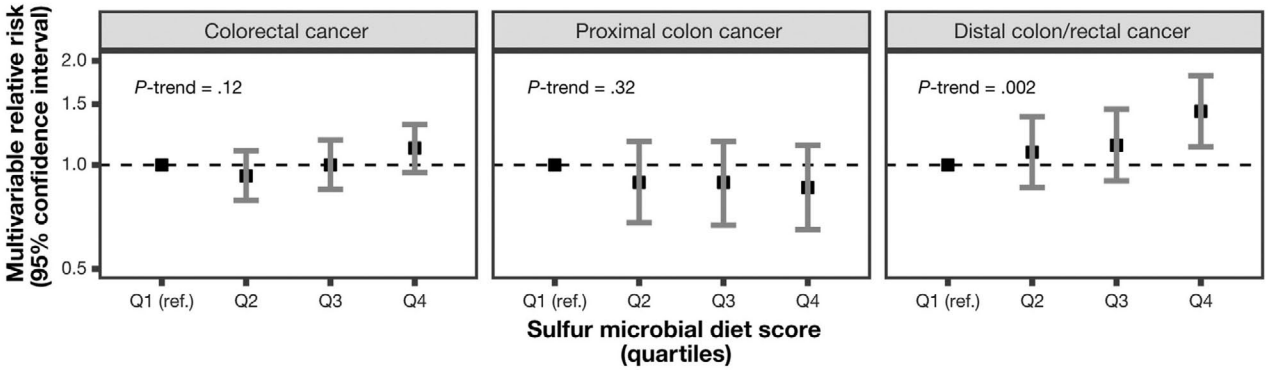
**Table 1.** Baseline Age-Standardized Characteristics by Sulfur Microbial Diet Score (HPFS, 1986)

| | Sulfur microbial diet score | | | |
|---|---|---|---|---|
| | Quartile 1 (n = 12,035) | Quartile 2 (n = 12,240) | Quartile 3 (n = 12,004) | Quartile 4 (n = 11,967) |
| Age, y | 54.2 (10.0) | 54.2 (9.9) | 54.3 (9.8) | 54.0 (9.6) |
| BMI, kg/m$^2$ | 24.9 (3.1) | 25.2 (3.1) | 25.6 (3.2) | 26.3 (3.5) |
| Alcohol intake, g/d | 13.0 (17.3) | 8.3 (10.8) | 8.4 (11.3) | 16.0 (19.2) |
| Physical activity, MET-h/wk | 21.7 (28.3) | 18.7 (24.8) | 17.7 (26.6) | 17.0 (24.1) |
| Past smokers, % | 39 40 42 47 | | | |
| Current smokers, % | 8 | 7 | 9 | 13 |
| Regular aspirin use, % | 28 | 29 | 28 | 31 |
| Family history of colorectal cancer, % | 15 | 15 | 14 | 14 |
| White race, % | 96 | 96 | 95 | 96 |
| Screening lower endoscopy within past 2 y, % | 27 | 28 | 27 | 26 |
| Calories, kcal/d | 2339 (629) | 1964 (557) | 1798 (550) | 1846 (588) |
| Dietary intake (servings/wk) | | | | |
| Processed meats | 1.9 (2.1) | 2.0 (2.1) | 2.4 (2.3) | 4.0 (4.3) |
| Liquor | 1.1 (2.5) | 1.3 (2.7) | 1.8 (3.6) | 5.6 (8.3) |
| Low-calorie drinks | 1.3 (2.8) | 1.7 (3.1) | 2.6 (3.9) | 8.4 (10.5) |
| Beer | 4.3 (7.5) | 1.5 (2.6) | 1.0 (2.0) | 1.0 (1.9) |
| Fruit juice | 8.5 (8.7) | 5.8 (4.8) | 4.3 (3.9) | 3.7 (3.9) |
| Legumes | 4.9 (3.4) | 3.1 (1.9) | 2.4 (1.6) | 2.2 (1.6) |
| Other vegetables | 5.4 (4.3) | 3.7 (2.6) | 2.9 (2.1) | 2.7 (2.1) |
| Sweets & desserts | 11.9 (12.1) | 7.1 (6.3) | 5.3 (5.0) | 4.8 (5.2) |

All values other than age have been directly standardized to age distribution (in 5-year age group) of all participants.
Mean (standard deviation) is presented for continuous variables.
MET, metabolic equivalent of tasks.

of intake previously associated with CRC risk.[19,54] Despite sharing several food groups, sulfur microbial diet scores were not associated with Western dietary scores (Spearman $\rho$ = –0.009 at study baseline), suggesting sulfur microbial diet scores capture a novel, independent signal within the well-established diet-CRC relationship.

We documented 1264 cases of incident CRC over 26 years of follow-up encompassing 1,077,325 person-years. Greater adherence to the sulfur microbial diet was associated with an increased risk of distal colon and rectal cancer (**Figure 5**). Compared with having a sulfur microbial diet score in the first quartile, men in the highest quartile had a multivariable relative risk of 1.43 (95% confidence interval 1.13–1.81; *P*-trend = .002), after adjusting for putative CRC risk factors. In contrast, no clear association was observed for proximal colorectal cancer (*P*-trend = 0.31).




| | Q1 | Q2 | Q3 | Q4 | *P*-trend |
|---|---|---|---|---|---|
| Colorectal cancer | | | | | |
| Cases | 305 | 294 | 317 | 348 | |
| Person-Years | 266,389 | 272,651 | 269,256 | 269,029 | |
| Age-adjusted RR (95% CI) | 1 [ref.] | 0.94 (0.80, 1.11) | 1.04 (0.89, 1.22) | 1.21 (1.04, 1.41) | **.008** |
| Multivariable RR (95% CI) | 1 [ref.] | 0.93 (0.79, 1.10) | 1.00 (0.85, 1.18) | 1.12 (0.95, 1.31) | .12 |
| Proximal colon cancer | | | | | |
| Cases | 116 | 107 | 107 | 101 | |
| Age-adjusted RR (95% CI) | 1 [ref.] | 0.91 (0.70, 1.18) | 0.92 (0.71, 1.20) | 0.93 (0.71, 1.22) | .63 |
| Multivariable RR (95% CI) | 1 [ref.] | 0.89 (0.68, 1.17) | 0.89 (0.67, 1.17) | 0.86 (0.65, 1.14) | .31 |
| Distal colon and rectal cancer | | | | | |
| Cases | 134 | 151 | 157 | 195 | |
| Age-adjusted RR (95% CI) | 1 [ref.] | 1.10 (0.87, 1.39) | 1.18 (0.94, 1.49) | 1.53 (1.23, 1.91) | **<.0001** |
| Multivariable RR (95% CI) | 1 [ref.] | 1.09 (0.86, 1.38) | 1.14 (0.90, 1.45) | 1.43 (1.13, 1.81) | **.002** |

**Figure 5.** Sulfur microbial diet and risk of CRC. Multivariable modeling demonstrating an association between increased adherence to the sulfur microbial diet and risk of distal colon and rectal cancer. Models adjusted for age, family history of CRC, BMI, physical activity, smoking, aspirin use, total caloric intake, prior endoscopy, and recent physical examination. Tests for trend were conducted using the median value of each quartile category as a continuous variable. RR, relative risk.

Further analyses by subgroups were notable for several reasons. When comparing extreme quartiles of sulfur microbial diet scores, the association was stronger among subjects without a family history of CRC (*P*-interaction = .02; **Figure 6**). Similarly, those with a lower BMI and no prior history of smoking were at greater risk for distal colon and rectal cancer than their referent counterparts. Taken together, this could suggest that adherence to the sulfur microbial diet may have an outsized influence on disease risk among those with few or no prior CRC risk factors.

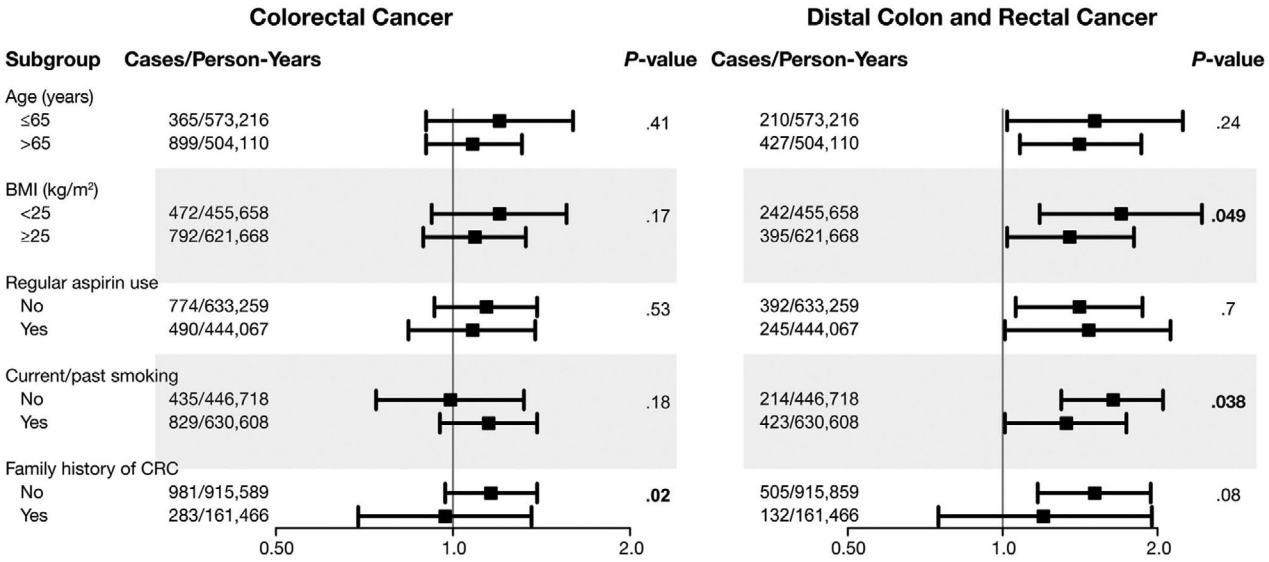



**Figure 6.** Sulfur microbial diet and risk of overall CRC and distal colon and rectal cancer by various subgroups. Multivariable relative risks comparing extreme quartiles of sulfur microbial diet scores were adjusted for age, family history of CRC, BMI, physical activity, smoking, aspirin use, total caloric intake, prior endoscopy, and recent physical examination with the exception of a given stratification variable.

## Conclusions

Our study, composed of the MLVS subcohort nested within the much larger, prospective HPFS, offered a unique opportunity to not only explore the dietary determinants of sulfur-metabolizing bacteria, but how human gut communities enriched for these microbes may correspond with disease risk. We linked predicted long-term enrichment for these bacteria using a de novo dietary-based index, the sulfur microbial diet score, to increased risk for distal CRC. Interestingly, the sulfur microbial diet was more closely aligned with elevated risk for CRC among individuals with fewer traditional CRC risk factors. Taken together, we offer first-of-its-kind evidence linking targeted microbiome discovery with diet-driven differences in disease risk at an epidemiologic scale, implicating gut microbial communities as a potential intermediary in the well-established diet-to- CRC relationship. These results may provide a rationale for considering dietary modification as a means to modulate long-term ecological states implicated in an array of GI diseases, including CRC and the IBDs.

Our primary findings were motivated by a collection of long-standing mechanistic factors that relate colonic nutrients to microbial biochemistry. A mucus bilayer in the colon partitions epithelial cells from the microbial contents of the gut. This typically sterile fortification separates primed epithelial tissue from highly immunogenic microbial antigens. Consequently, disruption of this interface is considered an important event in the development of CRC and IBD.[55,56] The mucins that comprise this barrier are joined by disulfide bonds, which are fragmented in the presence of excess $H_2S$, a harmful byproduct of sulfur metabolism.[14,16,57] This creates a breach in a crucial protective barrier. Other efforts have convincingly demonstrated how diet may modulate overall gut microbial composition.[4,5,58] Prior cross-sectional evidence found that high-fat diets enriched for meat-based proteins may promote distinct and less diverse populations of sulfur-metabolizing bacteria in humans.[24] Processed meats, a key component of the sulfur microbial diet, are viewed as particularly problematic because of their high sulfur content from both sulfur-containing amino acids and the inorganic sulfurs found in preservatives.[9] In contrast, plant-based sulfur sources, such as those found in legumes and vegetables, 2 food groups associated with the relative depletion of sulfur-metabolizing bacteria, are distinct from animal-based sources, and may include compounds called glucosinolates. Non–sulfur-metabolizing bacteria may produce myrosinases that hydrolyze glucosinolates to isothiocyanates, which have been extensively studied for their cancer preventive properties.[59–61] Thus, the source of sulfur, rather than quantifiable sulfur content of foods, may be more strongly predictive for sulfur-metabolizing bacterial abundance, a plausible explanation for prior inconsistent findings relating overall sulfur content with CRC risk.

In addition, our finding of elevated risk in the distal colorectum fits in the context of prior work in humans demonstrating CRC to be a molecular heterogeneous disease by anatomic site, suggesting there may be etiological differences among cancers arising from the proximal vs the distal large bowel.[62] Prior animal studies have demonstrated differential expression of toll-like receptor-2 and toll-like receptor-4 along the colorectum in specific-pathogen free, but not germ-free mice, suggesting that GI microbes may alter regional expression of these markers of innate immunity.[63] Other investigations have also reaffirmed that patterns of host gene function may be altered by

microbial signals dependent on favorable anatomic conditions.[64–66] Amino acid fermentation, including sulfur-containing cysteine and methionine, may be greater in the distal, rather than the proximal colon.[67] Taken together, multiple host factors critical to maintaining immune and microbial homeostasis decrease caudally and could promote the emergence of disease-promoting microbes, rendering the distal colorectum more susceptible to injury. The presented work adds a new dimension beyond traditional or well-established CRC risk factors with a proportion of attributable risk perhaps driven by biogeographical differences in microbial ecology (i.e., differences in the regional and anatomical diversity of the gut microbiome).[68]

This study expands on prior research in CRC and the gut microbiome, which have largely focused on either diet or microbiology in isolation. However, a collection of smaller scale or time-limited studies in humans,[4,22,24] corroborating experimental evidence,[9,69–71] and large-scale investigations of the CRC microbiome with little information on dietary intake,[23,72] have each substantiated the importance of sulfur-metabolizing bacteria, which we identified using a comprehensive literature review complemented by agnostic bioinformatics approaches. In one such study, participants fed a high-meat, high-fat diet for several days had an increase in bile-resistant sulfur-metabolizing bacteria and microbial DNA and RNA encoding for $H_2S$ -producing enzymes. [4] In a large cohort of patients with precancerous polyps and CRCs, there was relative enrichment of sulfur-metabolizing microbes, including *Bilophila* and *Desulfovibrio* spp, as well as sulfur-related pathways, all along the adenoma-carcinoma sequence.[72]

Our study has several strengths. First, we leveraged a large, deeply characterized study population with more than 26 years of follow-up, including a developmental cohort with contemporaneous dietary inventories and serial stool sampling. This allowed us to more comprehensively and longitudinally associate the intake of certain foods with the enrichment and depletion of sulfur-metabolizing bacteria, and subsequently link these putative foods and pattern of intake to risk for GI malignancy. Although $H_2S$ has long been believed to contribute to inflammation and carcinogenesis, it exists in gaseous, dissolved, and anionic forms in the colonic environment, making it challenging to measure in vivo. These limitations are circumvented by observing the alterations in bacterial counts and their enzymatic activity related to sulfur economy. Jointly incorporating metagenomic

taxonomic assignment with metatranscriptomic functional activity currently serve as the best available proxy for the genotoxic effects of $H_2S$ and allows us to more confidently assign the observed association between the sulfur microbial diet and distal CRC to the microbial metabolism of sulfur, rather than other functions also possessed by this class of bacteria. Second, information on dietary intake and cancer outcomes were regularly updated and collected prospectively with high follow-up rates, limiting recall, ascertainment, and selection bias. Third, we also collected details on several known risk factors for CRC that may confound the relationship between the sulfur microbial diet and CRC risk, and their inclusion in our multivariate models did not significantly alter our estimates. Finally, we were able to demonstrate that the presence of sulfur-metabolizing bacteria was coupled to their functional activity and transcript levels, an attempt to further link correlation to underlying biology.

We acknowledge several limitations. Our study is composed of older male health professionals, minimizing heterogeneity and potentially confounding by socioeconomic status. This is particularly noteworthy given prior evidence of anatomic heterogeneity in CRC risk by socioeconomic status.[73] However, concern for generalizability is minimal because our observations address a possible underlying mechanism relating diet to health likely to be substantially present among different populations. Despite these challenges, we found compelling evidence that certain foods may modulate the presence and activity of sulfur-metabolizing bacteria, likely related to differences in energy and substrate availability. Given the observational nature of our study, we cannot exclude the possibility of residual confounding; however, we carefully adjusted for multiple potential confounders. Although strain-level differences in pathogenic potential and microbial transcription have been well-established,[74–77] our compiled list of sulfur-metabolizing bacteria was limited to species-level taxonomy. This was a necessary constraint given our intention to strictly and confidently categorize microbes using both the cross-classification of bacterial pangenomes to known, but sparse sulfur enzymes, as well as requiring prior, robust experimental evidence of their involvement in sulfur metabolism. Finally, we restricted our analysis to microbes of sufficient prevalence and abundance in the healthy human gut. *Fusobacterium* spp were present in only 6 stool samples (mean relative abundance: $2.4 \times 10^{-3}$). This low prevalence is consistent with other

large-scale investigations of healthy humans.[74] Thus, our analysis did not specifically examine *Fusobacterium* spp. In contrast, *Fusobacterium* is more abundant among individuals with late-stage colorectal neoplasia,[72,78–80] especially CRC.[79,81]

Dietary modulation of the gut microbiome is of significant appeal as a strategy for risk minimization in chronic disease. Further confirmation of our findings will benefit from multidisciplinary approaches to elucidate the mechanisms that underlie how these phylogenetically diverse microbes influence intestinal inflammation and tumorigenesis. Epidemiologic investigations on whether the sulfur microbial diet influences risk of precursor adenomatous lesions, as well as trials in humans testing avoidance of implicated foods, are needed to identify at-risk populations and promising targets to ameliorate potential harms, respectively. Our findings may help unify the previously observed relationships linking sulfur-metabolizing microbes and CRC and offer a plausible mechanism in support of the well-characterized link between diet and CRC risk.

**Supplementary Material** (Supplementary Tables 1–4) is attached to this archive record.

## References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394–424.
2. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med 2000;343:78–85.
3. Song M, Chan AT. Environmental factors, gut microbiota, and colorectal cancer prevention. Clin Gastroenterol Hepatol 2019;17:275–289.
4. David LA, Maurice CF, Carmody RN, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature 2014;505:559–563.
5. Wu GD, Chen J, Hoffmann C, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science 2011;334:105–108.
6. Brennan CA, Garrett WS. Gut microbiota, inflammation, and colorectal cancer. Annu Rev Microbiol 2016;70:395–411.
7. Ramasamy S, Singh S, Taniere P, et al. Sulfidedetoxifying enzymes in the human colon are decreased in cancer and upregulated in differentiation. Am J Physiol Gastrointest Liver Physiol 2006;291:G288–G296.

8. Attene-Ramos MS, Wagner ED, Gaskins HR, et al. Hydrogen sulfide induces direct radical-associated DNA damage. Mol Cancer Res 2007;5:455–459.
9. Magee EA, Richardson CJ, Hughes R, et al. Contribution of dietary protein to sulfide production in the large intestine: an in vitro and a controlled feeding study in humans. Am J Clin Nutr 2000;72:1488–1494.
10. Scanlan PD, Shanahan F, Marchesi JR. Culture-independent analysis of desulfovibrios in the human distal colon of healthy, colorectal cancer and polypectomized individuals. FEMS Microbiol Ecol 2009;69:213–221.
11. Cai WJ, Wang MJ, Ju LH, et al. Hydrogen sulfide induces human colon cancer cell proliferation: role of Akt, ERK and p21. Cell Biol Int 2010;34:565–572.
12. Roediger WE, Duncan A, Kapaniris O, et al. Reducing sulfur compounds of the colon impair colonocyte nutrition: implications for ulcerative colitis. Gastroenterology 1993;104:802–809.
13. Zeng H, Combs GF Jr. Selenium as an anticancer nutrient: roles in cell proliferation and tumor cell invasion. J Nutr Biochem 2008;19:1–7.
14. Ijssennagger N, van der Meer R, van Mil SW. Sulfide as a mucus barrier-breaker in inflammatory bowel disease? Trends Mol Med 2016;22:190–199.
15. Schroeder BO, Birchenough GMH, Stahlman M, et al. Bifidobacteria or fiber protects against diet-induced microbiota-mediated colonic mucus deterioration. Cell Host Microbe 2018;23:27–40.e7.
16. Ijssennagger N, Belzer C, Hooiveld GJ, et al. Gut microbiota facilitates dietary heme-induced epithelial hyperproliferation by opening the mucus barrier in colon. Proc Natl Acad Sci U S A 2015;112:10038–10043.
17. Magalhaes B, Peleteiro B, Lunet N. Dietary patterns and colorectal cancer: systematic review and meta-analysis. Eur J Cancer Prev 2012;21:15–23.
18. Kesse E, Clavel-Chapelon F, Boutron-Ruault MC. Dietary patterns and risk of colorectal tumors: a cohort of French women of the National Education System (E3N). Am J Epidemiol 2006;164:1085–1093.
19. Fung T, Hu FB, Fuchs C, et al. Major dietary patterns and the risk of colorectal cancer in women. Arch Intern Med 2003;163:309–314.
20. Hou JK, Abraham B, El-Serag H. Dietary intake and risk of developing inflammatory bowel disease: a systematic review of the literature. Am J Gastroenterol 2011; 106:563–573.
21. Kaplan GG, Ng SC. Understanding and preventing the global increase of inflammatory bowel disease. Gastroenterology 2017;152:313–321.e2.
22. Rey FE, Gonzalez MD, Cheng J, et al. Metabolic niche of a prominent sulfate-reducing human gut bacterium. Proc Natl Acad Sci U S A 2013;110:13582–13587.
23. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. Nat Rev Microbiol 2014;12:661–672.
24. Nava GM, Carbonero F, Ou J, et al. Hydrogenotrophic microbiota distinguish native Africans from African and European Americans. Environ Microbiol Rep 2012; 4:307–315.

25. Wirfalt E, Midthune D, Reedy J, et al. Associations between food patterns defined by cluster analysis and colorectal cancer incidence in the NIH-AARP diet and health study. Eur J Clin Nutr 2009;63:707–717.
26. Terry P, Hu FB, Hansen H, et al. Prospective study of major dietary patterns and colorectal cancer risk in women. Am J Epidemiol 2001;154:1143–1149.
27. Ananthakrishnan AN, Khalili H, Konijeti GG, et al. A prospective study of long-term intake of dietary fiber and risk of Crohn's disease and ulcerative colitis. Gastroenterology 2013;145:970–977.
28. Makambi KH, Agurs-Collins T, Bright-Gbebry M, et al. Dietary patterns and the risk of colorectal adenomas: the Black Women's Health Study. Cancer Epidemiol Biomarkers Prev 2011;20:818–825.
29. Mehta RS, Abu-Ali GS, Drew DA, et al. Stability of the human faecal microbiome in a cohort of adult men. Nat Microbiol 2018;3:347–355.
30. Abu-Ali GS, Mehta RS, Lloyd-Price J, et al. Metatranscriptome of human faecal microbial communities in a cohort of adult men. Nat Microbiol 2018;3:356–366.
31. Franzosa EA, Morgan XC, Segata N, et al. Relating the metatranscriptome and metagenome of the human gut. Proc Natl Acad Sci U S A 2014;111:E2329–E2338.
32. Song SJ, Amir A, Metcalf JL, et al. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. mSystems 2016;1(3).
33. Voigt AY, Costea PI, Kultima JR, et al. Temporal and technical variability of human gut metagenomes. Genome Biol 2015;16:73.
34. McIver LJ, Abu-Ali G, Franzosa EA, et al. bioBakery: a meta'omic analysis environment. Bioinformatics 2018; 34:1235–1237.
35. Truong DT, Franzosa EA, Tickle TL, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods 2015;12:902–903.
36. Abubucker S, Segata N, Goll J, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. PLoS Comput Biol 2012;8: e1002358.
37. Suzek BE, Wang Y, Huang H, et al. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. Bioinformatics 2015; 31:926–932.
38. Caspi R, Billington R, Fulcher CA, et al. The MetaCyc database of metabolic pathways and enzymes. Nucleic Acids Res 2018;46:D633–D639.
39. MetaCyc Metabolic Pathway Database. Available at: https://metacyc.org/. Accessed August 4, 2018.
40. Placzek S, Schomburg I, Chang A, et al. BRENDA in 2017: new perspectives and new tools in BRENDA. Nucleic Acids Res 2017;45:D380–D388.
41. Werner H, Rintelen G, Kunstek-Santos H. [A new butyric acid-producing bacteroides species: B. splanchnicus n. sp. (author's transl)]. Zentralbl Bakteriol Orig A 1975;
42. Washio J, Shimada Y, Yamada M, et al. Effects of pH and lactate on hydrogen sulfide production by oral Veillonella spp. Appl Environ Microbiol 2014;80:4184–4188.

43. Rimm EB, Giovannucci EL, Stampfer MJ, et al. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am J Epidemiol 1992; 135:1114–1126; discussion 1127–1136.
44. Tabung FK, Smith-Warner SA, Chavarro JE, et al. Development and validation of an empirical dietary inflammatory index. J Nutr 2016;146:1560–1570.
45. Wolf AM, Hunter DJ, Colditz GA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. Int J Epidemiol 1994;23:991–999.
46. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. Scand J Gastroenterol 1997; 32:920–924.
47. Chen W, Liu F, Ling Z, et al. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. PLoS One 2012;7:e39743.
48. Turnbaugh PJ, Ridaura VK, Faith JJ, et al. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Sci Transl Med 2009;1:6ra14.
49. Hildebrandt MA, Hoffmann C, Sherrill-Mix SA, et al. High-fat diet determines the composition of the murine gut microbiome independently of obesity. Gastroenterology 2009;137:1716–1724; e1–2.
50. Turnbaugh PJ, Backhed F, Fulton L, et al. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell Host Microbe 2008;3:213–223.
51. Yazici C, Wolf PG, Kim H, et al. Race-dependent association of sulfidogenic bacteria with colorectal cancer. Gut 2017;66:1983–1994.
52. Feng Q, Liang S, Jia H, et al. Gut microbiome development along the colorectal adenoma-carcinoma sequence. Nat Commun 2015;6:6528.
53. Carbonero F, Benefiel AC, Alizadeh-Ghamsari AH, et al. Microbial pathways in colonic sulfur metabolism and links with health and disease. Front Physiol 2012; 3:448.
54. Mehta RS, Song M, Nishihara R, et al. Dietary patterns and risk of colorectal cancer: analysis by tumor location and molecular subtypes. Gastroenterology 2017; 152:1944–1953.e1.
55. Johansson ME, Phillipson M, Petersson J, et al. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. Proc Natl Acad Sci U S A 2008;105:15064–15069.
56. Podolsky DK. Inflammatory bowel disease. N Engl J Med 2002;347:417–429.
57. Velcich A, Yang W, Heyer J, et al. Colorectal cancer in mice genetically deficient in the mucin Muc2. Science 2002;295:1726–1729.
58. Muegge BD, Kuczynski J, Knights D, et al. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 2011;332:970–974.
59. Song M, Garrett WS, Chan AT. Nutrients, foods, and colorectal cancer prevention. Gastroenterology 2015; 148:1244–1260.e16.
60. Bianchini F, Vainio H. Isothiocyanates in cancer prevention. Drug Metab Rev 2004;36:655–667.

61. Hashem FA, Motawea H, El-Shabrawy AE, et al. Myrosinase hydrolysates of *Brassica oleraceae* L. var. *italica* reduce the risk of colon cancer. Phytother Res 2012; 26:743–747.
62. Yamauchi M, Morikawa T, Kuchiba A, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. Gut 2012;61:847–854.
63. Wang Y, Devkota S, Musch MW, et al. Regional mucosa-associated microbiota determine physiological expression of TLR2 and TLR4 in murine colon. PLoS One 2010; 5:e13607.
64. Kong J, Zhang Z, Musch MW, et al. Novel role of the vitamin D receptor in maintaining the integrity of the intestinal mucosal barrier. Am J Physiol Gastrointest Liver Physiol 2008;294:G208–G216.
65. Su W, Bush CR, Necela BM, et al. Differential expression, distribution, and function of PPAR-gamma in the proximal and distal colon. Physiol Genomics 2007;30:342–353.
66. Sodir NM, Chen X, Park R, et al. Smad3 deficiency promotes tumorigenesis in the distal colon of ApcMin/þ mice. Cancer Res 2006;66:8430–8438.
67. Macfarlane GT, Gibson GR, Cummings JH. Comparison of fermentation reactions in different regions of the human colon. J Appl Bacteriol 1992;72:57–64.
68. Martinez-Guryn K, Leone V, Chang EB. Regional diversity of the gastrointestinal microbiome. Cell Host Microbe 2019;26:314–324.
69. Devkota S, Wang Y, Musch MW, et al. Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in Il10_/_ mice. Nature 2012;487:104–108.
70. Ridlon JM, Kang DJ, Hylemon PB. Bile salt biotransformations by human intestinal bacteria. J Lipid Res 2006;47:241–259.
71. Shen W, Wolf PG, Carbonero F, et al. Intestinal and systemic inflammatory responses are positively associated with sulfidogenic bacteria abundance in high-fat-fed male C57BL/6J mice. J Nutr 2014;144:1181–1187.
72. Yachida S, Mizutani S, Shiroma H, et al. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. Nat Med 2019;25:968–976.
73. Doubeni CA, Laiyemo AO, Major JM, et al. Socioeconomic status and the risk of colorectal cancer: an analysis of more than a half million adults in the National Institutes of Health-AARP Diet and Health Study. Cancer 2012;118:3636–3644.
74. Lloyd-Price J, Mahurkar A, Rahnavard G, et al. Strains, functions and dynamics in the expanded Human Microbiome Project. Nature 2017;550:61–66.
75. Needham BD, Carroll SM, Giles DK, et al. Modulating the innate immune response by combinatorial engineering of endotoxin. Proc Natl Acad Sci U S A 2013;110:1464–1469.
76. Bielaszewska M, Mellmann A, Zhang W, et al. Characterisation of the Escherichia coli strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. Lancet Infect Dis 2011; 11:671–676.

77. Bron PA, van Baarlen P, Kleerebezem M. Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. Nat Rev Microbiol 2011;10:66–78.
78. Flanagan L, Schmid J, Ebert M, et al. Fusobacterium nucleatum associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. Eur J Clin Microbiol Infect Dis 2014;33:1381–1390.
79. Wirbel J, Pyl PT, Kartal E, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. Nat Med 2019;25:679–689.
80. Bullman S, Pedamallu CS, Sicinska E, et al. Analysis of Fusobacterium persistence and antibiotic response in colorectal cancer. Science 2017;358:1443–1448.
81. Thomas AM, Manghi P, Asnicar F, et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. Nat Med 2019;25:667–678.

**Acknowledgments —** We thank the participants who graciously participated in this research, 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. We acknowledge the efforts of the Channing Division of Network Medicine at the Brigham and Women's Hospital.

**Author Contributions —** Drs. Nguyen, Huttenhower, and Chan had full access to all study data and are responsible for data integrity and accuracy of the data analysis. Study concept and design: LHN, WSG, EBR, JI, CH, ATC Acquisition of data: LHN, CH, ATC. Analysis and interpretation of data: all coauthors. Drafting of the manuscript: LHN, CH, ATC. Critical revision of the manuscript for important intellectual content: all authors. Statistical analysis: LHN, CH, ATC. Obtained funding: WSG, EBR, JI, CH, ATC. Administrative, technical, or material support: CH, ATC. Study supervision: CH, ATC.

The authors disclose no conflicts of interest.

**Funding Grant Support —** National Institutes of Health (NIH) grants U54DE023798, UM1CA167552, U01CA167552, P01CA055075, U01CA152904, T32CA009001 (to LHN and DAD), Loan Repayment Program (to LHN and DAD), and K99DK119412 (to DW), K07CA218377 (to YC), R00CA215314 (to MS), R01CA202704 (to WSG, JI, CH, and ATC), K24DK098311 (to ATC), P30DK043351 (Pilot & Feasibility Study Award to DAD), R35CA197735 (to SO), and R01CA151993 (to SO); Crohn's and Colitis Foundation Research Fellowship Award (to LHN) and Senior Investigator Award (to ATC); American Cancer Society Mentored Research Scholar Grant (to MS); the US Department of Agriculture National Institute of Food and Agriculture Hatch Multistate Research Capacity Funding Program grant W4122 (to JI); National Health and Medical Research Council (to KLI); Cancer Research UK Grand Challenge Award (C10674/A27140; to WSG, SO, and CH); Dana-Farber Harvard Cancer Center Nodal Award (to SO); STARR Cancer Consortium; and the Massachusetts General Hospital Stuart and Suzanne Steel Research Scholars Award (to ATC).