Published in final edited form as:

Gut. 2017 November ; 66(11): 1983–1994. doi:10.1136/gutjnl-2016-313321.

Race-dependent association of sulfidogenic bacteria with colorectal cancer

Cemal Yazici¹, Patricia G Wolf², Hajwa Kim³, Tzu-Wen L Cross², Karin Vermillion², Timothy Carroll¹, Gaius J Augustus⁴, Ece Mutlu⁵, Lisa Tussing-Humphreys⁶, Carol Braunschweig⁷, Rosa M Xicola⁸, Barbara Jung¹, Xavier Llor⁸, Nathan A Ellis⁴, and H Rex Gaskins^{2,9}

¹Division of Gastroenterology and Hepatology, University of Illinois College of Medicine at Chicago, Chicago, Illinois, USA

²Division of Nutritional Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

³Center for Clinical and Translational Science, University of Illinois at Chicago, Chicago, Illinois, USA

⁴Department of Cellular and Molecular Medicine, University of Arizona, Tucson, Arizona, USA

⁵Division of Digestive Diseases and Nutrition, Rush University Medical Center, Chicago, Illinois, USA

⁶Division of Academic Internal Medicine and Geriatrics, University of Illinois College of Medicine at Chicago, Chicago, Illinois, USA

⁷Department of Kinesiology and Nutrition, University of Illinois at Chicago, Chicago, Illinois, USA

⁸Section of Digestive Diseases, Yale University, New Haven, Connecticut, USA

⁹Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

Abstract

Objective—Colorectal cancer (CRC) incidence is higher in African Americans (AAs) compared with non-Hispanic whites (NHWs). A diet high in animal protein and fat is an environmental risk factor for CRC development. The intestinal microbiota is postulated to modulate the effects of diet in promoting or preventing CRC. Hydrogen sulfide, produced by autochthonous sulfidogenic

To request permissions go to: <http://group.bmj.com/group/rights-licensing/permissions>

Correspondence to, Dr H. Rex Gaskins, University of Illinois at Urbana-Champaign, 1206 W. Gregory Drive, Urbana, IL 61801, USA; hgaskins@illinois.edu OR Dr Nathan A Ellis, University of Arizona Cancer Center, 1515 N Campbell Ave, Tucson, AZ 85724, USA; naellis@email.arizona.edu.
CY and PGW share co-first authorship.

Contributors Study concept and design (NAE, HRG, PGW and CY); acquisition of data (KV, PGW and CY); analysis and interpretation of data (GJA, NAE, HRG, LTH, EM, T-WL, PGW and CY); drafting manuscript (NAE, HRG, PGW and CY); critical revision of manuscript (NAE, HRG, LTH, BJ, T-WL, XL, PGW, CY and RMX); statistical analysis (GJA, HK and EM); obtained funding (NAE, LTH, BJ, XL, EM and HRG); administrative, technical or material support (CB, TC, XL, LTH, BJ and RMX); study supervision (XL, RMX and HRG).

Competing interests None declared.

Ethics approval Institutional review boards of each participating medical centre.

Provenance and peer review Not commissioned; externally peer reviewed.

bacteria, triggers proinflammatory pathways and hyperproliferation, and is genotoxic. We hypothesised that sulfidogenic bacterial abundance in colonic mucosa may be an environmental CRC risk factor that distinguishes AA and NHW.

Design—Colonic biopsies from uninvolved or healthy mucosa from CRC cases and tumour-free controls were collected prospectively from five medical centres in Chicago for association studies. Sulfidogenic bacterial abundance in uninvolved colonic mucosa of AA and NHW CRC cases was compared with normal mucosa of AA and NHW controls. In addition, 16S rDNA sequencing was performed in AA cases and controls. Correlations were examined among bacterial targets, race, disease status and dietary intake.

Results—AAs harboured a greater abundance of sulfidogenic bacteria compared with NHWs regardless of disease status. *Bilophila wadsworthia*-specific *dsrA* was more abundant in AA cases than controls. Linear discriminant analysis of 16S rRNA gene sequences revealed five sulfidogenic genera that were more abundant in AA cases. Fat and protein intake and daily servings of meat were significantly higher in AAs compared with NHWs, and multiple dietary components correlated with a higher abundance of sulfidogenic bacteria.

Conclusions—These results implicate sulfidogenic bacteria as a potential environmental risk factor contributing to CRC development in AAs.

INTRODUCTION

Colorectal cancer (CRC) is the third most frequent cancer in the USA with an estimated 134 490 new cases and 49 190 deaths in 2016. African Americans (AAs) have a CRC incidence 10%–30% higher than other races and ethnicities, and even greater disparities in mortality rates.¹ AAs are diagnosed at earlier ages and have more proximal CRC compared with non-Hispanic whites (NHWs).² Although some types of CRC are influenced by genetic risk factors or disease, the majority occur sporadically as a consequence of environmental exposures that promote colonocyte hyperproliferation, loss of apoptotic control and epithelial barrier integrity, and excess production of proinflammatory immunomodulatory factors.³

Significant differences in faecal and colonic microbial composition in healthy subjects and patients with cancer have been identified,^{4–13} but the impact remains uncertain. Compared with tumour-free controls, CRC subjects have higher levels of faecal hydrogen sulfide (H₂S).¹⁴ Higher levels of faecal H₂S have been detected in patients with colitis, who have an increased susceptibility of CRC.^{15–18} Exogenous H₂S is genotoxic at levels commonly found in the colon. This bacterial derived gas induces proliferative and proinflammatory pathways and inhibits β -oxidation of butyrate—an anticarcinogenic metabolite and favoured substrate of colonocytes.^{19–22} Multiple bacterial species produce H₂S, including taurine-respiring *Bilophila wadsworthia*, cysteine-using *Fusobacterium nucleatum* and sulfate-reducing bacteria (SRB) such as *Desulfovibrio* spp.^{23–25} Additionally, higher levels of faecal bile acids have been observed in CRC cases than controls,¹⁴ which could be associated with increased abundance of taurocholate-using *B. wadsworthia*.

Because exposure to H₂S could promote CRC development, we hypothesised that sulfidogenic bacteria contribute to a micro-environment conducive to colorectal

carcinogenesis through production of H₂S. Accordingly, we examined sulfidogenic bacterial abundance in uninvolved mucosa of subjects with CRC compared with tumour-free controls. Of particular interest was whether differences in sulfidogenic bacteria between AAs and NHWs might contribute to racial differences in CRC incidence rates.

MATERIALS AND METHODS

Human subjects

The Chicago Colorectal Cancer Consortium prospectively ascertained incident CRC cases from surgery and endoscopy units and tumour-free controls undergoing routine screening colonoscopy from five Chicago medical centres, including University of Illinois Hospital and Health Sciences System, Jesse Brown Veterans Administration, John H. Stroger Hospital of Cook County, University of Chicago Medicine and Rush University Medical Center, over a 2-year period (2011–2012), as previously described.²⁶ The parent study received approval for human subjects research from the institutional review boards (IRBs) of each participating medical centre; the parent protocol was administered by the IRB at University of Illinois Hospital and Health Sciences System (2010–0168). Colonic tissue biopsies from uninvolved or healthy mucosa were studied from 329 subjects after routine bowel preparation. The cases were persons with adenocarcinoma of the colon or rectum detected during patient workup due to GI symptoms or collected at the time of surgical resection. Control subjects were persons undergoing colonoscopy in whom tumours were not identified. In CRC cases, biopsies were collected approximately 10 cm away from the tumour location. Samples were taken either with endoscopic biopsy or surgical forceps. Included were 97 AA and 56 NHW CRC cases, and 100 AA and 76 NHW controls. Persons with a personal history of cancer, IBD, adenomatous or benign polyps, polyposis or tumours with microsatellite instability were excluded.²⁶ Clinical data were extracted from medical records, and epidemiological and demographic information were obtained using subject questionnaires as previously described.²⁶ Factors analysed included age, sex, race; tumour or biopsy site, tumour stage; body weight, body mass index (BMI); education and income level; usage of tobacco products and alcohol; and usage of omega-3 fatty acids and non-steroidal anti-inflammatory drugs (NSAIDs) (table 1 and online supplementary table S1). Significantly different factors were used as covariates for potential confounder adjustment.

Quantitative PCR (qPCR) analysis

Mucosal biopsies were placed into RNeasy lysis buffer and stored at 4° C overnight, after which they were stored at –80° C until extraction. DNA was isolated from mucosal biopsies using commercial DNA extraction kits (Promega, Madison, Wisconsin; Mobio, Carlsbad, California, USA), quantified and diluted to 5 ng/μL. Examination of bacterial target abundances of NHW cases prepared with the two kits yielded similar results, indicating that DNA extraction method did not introduce bias. Sulfidogenic bacteria abundance was quantified in triplicate with a 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, California, USA). Small subunit rRNA genes of SRBs including *Desulfobacter* spp. (DSB), *Desulfobulbus* spp. (DBB), *Desulfotomaculum* spp. (DFM) and *Desulfovibrio* spp. (DSV) were targeted and validated as previously described.²⁷ In addition, functional

genes from the H₂S production pathway, pan-dissimilatory sulfite reductase A (*dsrA*)²⁸ harboured by all SRB²⁹ and *B. wadsworthia*-specific *dsrA* (*dsrA*-BW)²² were quantified.

Statistical analysis

To test differences between groups, Wilcoxon rank-sum test or t-test for continuous variables and χ^2 or Fisher's exact test for dichotomous variables were used. Because bacterial targets were not normally distributed, data were log-transformed. For binary outcome analyses (AA vs NHW and case vs control), logistic regression models were used adjusting for selected covariates. In all models, age, sex and biopsy location were included. For sulfidogenic bacterial analysis in AAs, BMI, education level and annual income level were included. In NHWs, BMI, education level and NSAID use were included. Spearman correlation analysis was used to test associations between sulfidogenic bacteria. Statistical tests were conducted using SAS V.9.3 (SAS, Cary, North Carolina, USA).

16S rDNA sequence analysis and bioinformatics

Mucosal DNA samples were processed using Zymo Research DNA Clean and Concentrator, quantified using a Qubit fluorimeter (Life Technologies, Carlsbad, California, USA), and the V4 region of 16S rDNA was amplified according to Fluidigm protocols. Products were subjected to a second round of amplification with Illumina linkers and barcodes. Final products were quantified, quality tested with a Fragment Analyzer (Advanced Analytics, Ames, Iowa, USA) and size selected using a 2% agarose E-gel (Life Technologies). Quality of cleaned and extracted product was verified using an Agilent Bioanalyzer (Agilent Technologies, Santa Clara, California, USA), and sequencing was carried out using an Illumina MiSeq (Illumina, San Diego, California, USA).

Sequencing of the V4 region resulted in 4 885 578 raw reads which were demultiplexed, quality filtered to remove low abundant operational taxonomic units (OTUs) and singletons, and clustered using a 97% similarity threshold through QIIME 1.8.0 using the Greengenes 13_8 reference database, as previously described.^{30–32} Quality filtering and rarefaction to 5571 (max=67 359) sequences per sample resulted in 2 689 301 reads from 155 samples (61 AA CRC cases and 94 AA controls), with a sample mean of 17 350 sequences and an average base pair length of 252. α -diversity and β -diversity were assessed using the `core_diversity_analysis.py` workflow in QIIME 1.8.0. To confirm core diversity findings were not an artefact of sequence rarefaction and to obtain an estimate of robustness, comparable-sized jackknifed sample datasets were created, and β -diversity was reanalysed in QIIME 1.9.1.³¹

Linear discriminant analysis (LDA) effect size was calculated with the LEfSe algorithm to identify biomarkers for the differences between CRC cases and controls.³³ A secondary analysis was performed to examine the differences between subjects based on tumour location, comparing left-sided cases (n=33) and left-sided controls (n=52) and comparing right-sided cases (n=27) and right-sided controls (n=42); and based on age, comparing cases <50 (n=11) and controls <50 (n=20) and comparing cases <50 (n=11) and cases $\geq$ 50 (n=48). No further adjustments for multiple testing were required.³³

Analysis of dietary intake

Dietary intake was obtained using the Block Brief 2000 Food Frequency Questionnaire (BBFFQ) from a subset of subjects, including 50 AA and 31 NHW CRC cases and 30 AA and 24 NHW controls.³⁴ The questionnaires were administered within 3 weeks of the diagnostic procedure to reduce recall bias. NutritionQuest (Berkeley, California, USA) processed completed BBFFQs. The Diet and Behavior Shared Resource of the University of Illinois-Chicago Cancer Center prepared a dataset for statistical analysis.

Dietary intake data from outliers were excluded (<600 kcals (n=11) or >3500 kcals (n=1)). Because energy intake is highly correlated with macronutrient and micronutrient intake,³⁵ nutrient or food group was adjusted per 1000 cal. Density variables were created for total fat (g), saturated fat (g), monounsaturated fat (g), polyunsaturated fat (g), protein (g), total carbohydrate (g), total fibre (g), cysteine (mg), calcium (mg), iron (mg), alcohol (g), omega-6 fatty acids (g), and daily food group servings of meat, vegetables, fruit, total grains and dairy. Percent kilocalories from fat, protein and carbohydrate were also examined.

Differences in intake of nutritional variables between AAs and NHWs were assessed using the Wilcoxon signed-rank test or t-test, as appropriate. To analyse the relationship between selected dietary intake variables (cysteine, total fat and daily servings of dairy) and sulfidogenic bacterial abundance, linear regression models were created using R V.3.3.1 and adjusted for age, sex, race, disease status and BMI (R Foundation for Statistical Computing, Vienna, Austria. Packages: tidy-verse v.1.0.0, gridExtra v.2.2.1). Due to the high degree of co-linearity between these variables and other dietary factors of interest, it was necessary to isolate the independent effect of these intake variables. Consequently, we residualised each variable by obtaining the residuals for a model fit with all correlated variables >0.50 as explanatory and used these residuals in our final model.

RESULTS

Race-specific differences in mucosal sulfidogenic bacteria

Comparison of sulfidogenic bacterial targets between races revealed pan-*dsrA*, a measure of SRB abundance across a range of species, to be 10 times higher in AAs compared with NHWs ($p<0.001$), irrespective of disease status (figure 1A). Abundance of individual SRB genera was higher in AAs compared with NHWs for all species except DBB; DBB levels were mostly undetectable in AAs. Similarly, the abundance of *B. wadsworthia* was 2.5 times higher in AAs compared with NHWs, irrespective of disease status ($p<0.001$). These differences in abundance remained statistically significant after adjustment for clinical and epidemiological variables that were different between races (see online supplementary table S1).

Comparison of sulfidogenic bacterial targets between CRC cases and controls within each racial group revealed the mean abundance of pan-*dsrA* to be 1.8 times higher in AA controls than AA CRC cases ($p<0.001$) (figure 1B). By contrast, this target was 2.9 times higher in NHW CRC cases than NHW controls ($p<0.001$). These racial differences remained significant after adjusting for age, sex, biopsy site, BMI, education, income level and NSAID use (table 1).

The taurine-respiring sulfidogenic bacterium *B. wadsworthia* was 1.9 times more abundant in AA CRC cases than AA controls ($p<0.001$), and remained significant after adjusting for demographic factors mentioned previously. On the other hand, abundance of *B. wadsworthia* did not differ between NHW cases and controls.

Consistent with the pan-*dsrA* results, the abundance of DSB ($p<0.001$) and DFM ($p<0.001$) was significantly higher in AA controls than AA CRC cases. On the other hand, abundance of DSB ($p=0.018$) and DBB ($p<0.001$) was significantly higher in NHW CRC cases than NHW controls. These data demonstrate that AA and NHW cases and controls display reciprocal differences in SRB abundance, which remained after adjusting for tumour stage. Additionally, Spearman's correlation analysis revealed that SRBs correlated positively with *B. wadsworthia* in NHWs but not in AAs (see online supplementary table S2).

16S rDNA sequencing analysis in AAs

Due to the disparity in CRC risk in the AA compared with the NHW population, a 16S rRNA gene sequence analysis was conducted to validate sulfidogenic abundance trends observed by our qPCR analysis and to compare case versus control differences in the taxa not tested by qPCR analysis. The overall microbial population was represented by 10 phyla, 70 families and 137 genera (see online supplementary figure S1). Relative abundance differences in phyla between AA CRC cases and AA controls were not statistically significant, and measures of α -diversity were not different. Analysis of UniFrac distance matrices using Monte Carlo permutations ($n=999$) revealed significant differences in β -diversity between AA CRC cases and AA controls for both unweighted ($p=0.01$) and weighted ($p=0.01$) UniFrac distances (see online supplementary figure S2A, B). To ensure results were not due to sequence rarefaction, this analysis was further validated using jackknifed estimates, and similar results were obtained (see online supplementary figure S3). Overall, these data indicate that AA CRC cases displayed differences in taxon composition for both highly abundant and rare taxa compared with AA controls.

To determine which bacterial taxa were primary drivers of microbiome differences between AA CRC cases and AA controls, LDA was used to calculate the effect size of different bacterial taxa (figure 2A). LDA analysis revealed the genera *Faecalibacterium*, *Pseudomonas* and an unknown genus of *Peptostreptococcaceae* were strongly associated with AA controls; and the genera *Alistipes*, *Delftia*, *Erysipelotrichaceae*, *Micrococcus*, *Pyramidobacter*, *Ruminococcus* and unidentified genera associated with the *Bacteroidales*, *Christensenellaceae* and *Mogibacteriaceae* families were associated with AA CRC cases. The mean abundance of genera identified by LDA in AA CRC cases and controls is shown in figure 2B. We note that many of the subjects had undetected levels for the bacteria identified by this analysis.

Analysis of genera-specific differences of sulfidogenic bacteria revealed increased relative abundance of *Bilophila* and *Desulfovibrio* in AA CRC cases compared with AA controls, although these differences were not significant after adjusting for false discovery rate (figure 2C). OTUs for DSB, DFM and DBB were not identified. *Fusobacterium* was the most abundant sulfidogenic bacterium identified in the analysis, but the relative abundance of *Fusobacterium* was similar between cases and controls. Apparent differences in the relative

abundance of *Desulfovibrio* as estimated by 16S rDNA sequencing likely reflects the greater number of null reads in control subjects.

LDA was used to identify differences in the relative abundance of bacterial genera for right-sided and left-sided colonic biopsies (figure 3A, B). Effect size scores identified *Pyramidobacter*, *Weissella* and unknown genera of the *Bacteroidales*, *Mogibacteriaceae* and *Pseudomonadaceae* families as being significantly associated with right-sided AA CRC cases. The genera *Faecalibacterium*, *Erwinia*, *Klebsiella*, *Variovorax*, *Caulobacteraceae* and *Delftia* were associated with right-sided AA controls. *Coprobacillus*, *Alistipes*, *Christensenellaceae*, *Cc 115 Erysipelotrichaceae*, *Porphyromonas*, *Ruminococcus*, *Odoribacter* and unknown genera belonging to *YS2* were significantly associated with left-sided AA CRC cases. The genera *Oxalobacteraceae*, *Erwinia* and unknown genera of *Peptostreptococcaceae* were associated with left-sided AA controls.

Additionally, LDA was used to identify differences in relative abundance of bacterial species comparing CRC cases and controls <50 years of age (figure 3C). The genera *Bilophila*, *Parabacteroides*, *Odoribacter* and an unknown genus of the *Barnesiellaceae* family were associated with early-onset CRC, and the genus *Acidaminococcus* was associated with controls <50 years of age. We also compared CRC cases below 50 and CRC cases 50 and older, revealing two genera (unknown genera of the families *Barnesiellaceae* and *Pseudomonadaceae*) associated with cases above the age of 50, and 11 genera (*Staphylococcus*, *Coprobacillus*, *Lactococcus*, *Corynebacterium*, *Cetobacterium*, *Megamonas*, *Anaerococcus*, *Sneathia*, *Odoribacter*, and unknown genera of the *Methylobacteriaceae* and *Mogibacteriaceae* families) associated with cases under 50 years of age (figure 3D).

Associations between diet and sulfidogenic bacteria

To examine whether diet differed between AAs and NHWs, dietary intake was analysed with the BBFFQ, focusing on nutrient and dietary factors associated with the typical mixed Western-style diet, namely intake of total fat, animal fat and protein. Overall, dietary fat and protein intake per 1000 kcals was significantly greater in AA compared with NHW subjects; on the other hand, calcium intake and servings of dairy per 1000 kcals were lower (see online supplementary table S3). To examine the effect of dietary intake on sulfidogenic target abundance, these dietary variables were examined using linear regression models (figure 4 and online supplementary figure S4, online supplementary table S4). As observed in the larger dataset (figure 1), the three major sulfidogenic bacteria tested (pan-*dsrA*, *B. wadsworthia* and DSV) strongly associated with race—the main variable that explained differences between AAs and NHWs (figure 4). In the models, pan-*dsrA* abundance also significantly associated with residuals for cysteine ($p=0.016$), total fat ($p=0.011$) and daily servings of dairy ($p=0.007$), as well as male sex ($p=0.0032$). DSV abundance was not significantly associated with dietary variables, but trends were observed with cysteine and male sex. Similarly, abundance of *B. wadsworthia* was not significantly associated with dietary variables, but there was a strong association with age (figure 4; $p=0.001$).

DISCUSSION

The present results demonstrate that the microbiome of uninvolved mucosa in CRC cases is distinct from tumour-free controls and implicate sulfidogenic bacteria as potential contributors to CRC development. One of the more striking observations was the greater abundance of SRBs and *B. wadsworthia* in AAs compared with NHWs, irrespective of the disease status. Pan-*dsrA* was 10-fold and *B. wadsworthia dsrA* 2.5-fold more abundant in AA colonic mucosa, and overall, analysis of individual sulfidogenic bacterial species confirmed this difference. Second, reciprocal differences in sulfidogenic bacterial abundance were observed in both AA and NHW CRC cases compared with controls. Notably, similar reciprocal differences were observed during diet exchange between rural South African blacks and AAs living in Pittsburgh.¹⁴

We demonstrated previously that human colonic mucosa is persistently colonised by sulfidogenic bacteria.³⁶ Constitutive bacterial derived sulfide may have a protective antimicrobial effect, based on the susceptibility of resident microbes to sulfide.³⁷ Sulfide production may become deleterious when the provision of organic sources of sulfur promotes growth of taurine and cysteine using bacteria, increasing sulfide to proinflammatory and genotoxic concentrations. Taurine provided by bacterial deconjugation of taurocholic acid acts as a substrate for anaerobic respiration by *B. wadsworthia*, generating genotoxic H₂S.³⁸ Once deconjugated, free primary bile acids are further metabolised by colonic bacteria to genotoxic and proinflammatory secondary bile acids. Specifically, the secondary bile acid, deoxycholic acid, acts as a tumour promoter, causing membrane perturbations leading to arachidonic acid release. Arachidonic acid is converted by cyclo-oxygenase-2 and lipo-oxygenase to proinflammatory and proangiogenic prostaglandins and reactive oxygen species, damaging DNA and inhibiting DNA repair enzymes.³⁸ Accordingly, these data support the hypothesis that differences in sulfidogenic bacteria between AAs and NHWs may contribute to variations in CRC incidence between these two populations.

Examination of microbial abundance in uninvolved mucosa from AA CRC cases and controls using 16S rDNA analysis underscored a number of prevailing taxonomic trends associated previously with cancer development in NHWs.^{4–13} For example, in agreement with previous analyses,^{8–10,12} relative abundance of *Mogibacteriaceae*, *Leptotrichia*, *Porphyromonas*, *Erysipelotrichaceae* and *Ruminococcus* was significantly higher in AA CRC cases. Similar to *F. nucleatum*, *Mogibacteriaceae* was first isolated from the human oral cavity and associated with oral and extraoral abscesses.³⁹ Likewise, species of the genera *Sneathia*,⁴⁰ formally known as *Leptotrichia sanguinegens*,⁴¹ and *Porphyromonas* are prevalent in periodontal disease,⁴² and recently members of the oral microbial community have been associated with CRC.⁴³ Organisms from the family *Erysipelotrichaceae* have been linked with IBD, are consistently positively associated with high-fat diets and dyslipidemia, and may be modulated by changes in host cholesterol metabolites.⁴⁴ while some species of the genus *Ruminococcus* appear to be protective,^{45,46} a greater abundance of *Ruminococcus obeum* was observed in the faeces of rats with 1,2-dimethyl hydrazine-induced precancerous lesions.⁴⁷

Several reports note increased abundance of sulfidogenic microbes in CRC-related dysbiosis, but few have connected the production of H₂S with CRC development.^{5–13} Our study is the first to look at differences in sulfidogenic bacteria, specifically in uninjured mucosa of individuals with CRC compared with controls. LDA revealed several genera associated with sulfur metabolism as significant indicators of disease, including *Mogibacteriaceae*, an organism previously observed to co-occur with volatile sulfur compound-producing organisms like *F. nucleatum*,³⁹ and the sulfidogenic genera *Lactococcus*, *Porphyrromonas*, *Odoribacter*, *Bilophila* and *Pyramidobacter*.^{2348–51} Similar to *F. nucleatum*, species of *Lactococcus* and *Poryphyromonas* generate sulfide from cysteine metabolism.⁴⁸⁴⁹ The growth of *Odoribacter splanchnicus*, another species from the *Porphyrromonadaceae* family, is enhanced with the addition of bile, and this bacterium produces an abundance of sulfide.⁵⁰ While one other study observed higher abundance of *Pyramidobacter* in proximal colon cancer tissue, the authors did not make the association with sulfide.⁶ Our study also corroborates previous evidence that microbial composition of proximal and distal cancers differ and that *Pyramidobacter* is enriched in proximal cancers.⁶ Thus, *Pyramidobacter* spp. could contribute to the propensity for proximal tumours in AAs.² The present report is the first to demonstrate that *B. wadsworthia* is associated with CRC.

A meta-analysis of prospective cohorts revealed increased CRC incidence among persons who consume a diet high in red and processed meat.⁵² Consistently, the present study observed a higher intake of dietary fat and protein in AAs, who are at higher risk of CRC development compared with NHWs. A Western diet is thought to be a major contributor to CRC incidence variance in different countries. For example, rural South African blacks, who consume a diet low in fat and animal protein, have a negligible risk of CRC compared with AAs and American NHWs, who on average consume a diet high in animal products.²¹ A recent comparison of AA and rural South African blacks showed significantly higher levels of both primary and secondary bile acids in AAs, as well as reciprocal changes in secondary bile acids and colonic microbiota after diet exchange.²¹ Additionally, a recent study revealed rapid changes in ‘bile-tolerant’ organisms upon consumption of a diet rich in fat and animal protein, including two genera abundant in this study, namely, *Alistipes* and *Bilophila*.⁵³ Notably, species of *Cetobacterium*, a taxon observed to be significant in AA CRC, are also bile-tolerant.⁵⁴ Similarly, *B. wadsworthia* was significantly more abundant in mice fed a high milk fat diet and increased severity of symptoms in dextran sodium sulfate (DSS) induced colitis.⁵⁵ A high-fat diet, rich in the sulfur-containing amino acids cysteine and taurine, increases bile secretion and shifts the taurine:glycine bile acid ratio towards taurine conjugation.⁵⁶

A key advantage of the present study is that recruited subjects represent an urban American population of average to low-income persons with predominantly high school or some college education, enabling a more generalisable examination of health and dietary disparities than performed previously.¹²⁵⁴⁵⁷⁵⁸ Additionally, epithelial cell function is affected to a greater extent by mucosally associated microbes^{59–61} as measured herein, functional differences have been observed between mucosal and luminal microbiota,⁶² and mucosal but not faecal microbiota may discriminate between disease status and health.⁶⁰

Our study is limited by use of the BBFFQ to measure habitual dietary intake of subjects. It is well established that diet assessment methods which rely on self-report are prone to underreporting of nutrient intake due to socially desirable responses and recall bias.^{63,64} The BBFFQ has a reduced food list (only 70 foods/beverages) and thus inherently will result in energy and macronutrient estimates lower than actual intake. With this in mind, nutrient density (per 1000 kcals) variables were created for dietary components of interest and have been suggested to more closely correlate with 'true' intake versus estimates of absolute intake.⁶⁴ Given the limitations of self-report diet assessment methods, to best understand the impact of diet composition in the manipulation of bile acid ratios and microbial metabolites, controlled feeding studies are warranted. Additional limitations include lack of information on probiotic and antibiotic use by subjects. Use of LEfSe to identify significant genera in the 16S analysis may bias against identification of rare OTUs; however, it is one of the most commonly used analytical methods for biodiscovery. Metabolites and histological biomarkers were not examined in the present study; these variables should be evaluated in future studies.

Our analysis revealed that race was the strongest variable associated with differences in sulfidogenic bacteria in AA compared with NHWs, and that diet differences had relatively small effects; consequently, we propose some other latent variable explains this difference, identifying that latent variable is an important future direction. In the meantime, it is clear a greater understanding of sulfidogenic bacteria associated with the colonic mucosa and development of novel CRC prediction models using key sulfidogenic bacteria may provide substantial knowledge for prevention and early detection of CRC particularly in the AA population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The Fluidigm and Illumina assays were performed by the Functional Genomics and DNA Sequencing Cores at the Carver Biotechnology Center, University of Illinois Urbana-Champaign. Statistical analysis was performed by Design and Analysis Core at the Center for Clinical and Translational Science, University of Illinois at Chicago, Chicago, IL.

Funding This work was supported by grants from the National Cancer Institute (U01 CA153060 and P30 CA023074, NAE; RO1 CA204808, HRG, EM, LTH; RO1 CA141057, BJ) and the American Cancer Society Illinois Division (223187, XL). PGW was supported by a predoctoral fellowship from Mayo Clinic and University of Illinois Alliance for Technology-Based Healthcare. CY was supported by a training grant from National Institute of Health (5T32DK007788-15). GJA was supported by a Cancer Biology Training Grant (T32CA009213).

References

1. American Cancer Society. Cancer Facts & Figures 2016. Atlanta: American Cancer Society; 2016.
2. Carethers JM. Screening for colorectal cancer in African Americans: determinants and rationale for an earlier age to commence screening. Dig Dis Sci. 2015; 60:711–21. [PubMed: 25540085]
3. Cho M, Carter J, Harari S, et al. The interrelationships of the gut microbiome and inflammation in colorectal carcinogenesis. Clin Lab Med. 2014; 34:699–710. [PubMed: 25439270]
4. Schulz MD, Atay C, Heringer J, et al. High-fat-diet-mediated dysbiosis promotes intestinal carcinogenesis independently of obesity. Nature. 2014; 514:508–12. [PubMed: 25174708]

5. Kostic AD, Chun E, Robertson L, et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe*. 2013; 14:207–15. [PubMed: 23954159]
6. Gao Z, Guo B, Gao R, et al. Microbiota dysbiosis is associated with colorectal cancer. *Front Microbiol*. 2015; 6:20. [PubMed: 25699023]
7. Kostic AD, Gevers D, Pedamallu CS, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res*. 2012; 22:292–8. [PubMed: 22009990]
8. Warren RL, Freeman DJ, Pleasance S, et al. Co-occurrence of anaerobic bacteria in colorectal carcinomas. *Microbiome*. 2013; 1:16. [PubMed: 24450771]
9. Wang T, Cai G, Qiu Y, et al. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME J*. 2012; 6:320–9. [PubMed: 21850056]
10. Chen W, Liu F, Ling Z, et al. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PLoS ONE*. 2012; 7(6):e39743. [PubMed: 22761885]
11. Castellarin M, Warren RL, Freeman JD, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res*. 2012; 22:299–306. [PubMed: 22009989]
12. Zeller G, Tap J, Voigt AY, et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol*. 2014; 10:766. [PubMed: 25432777]
13. Zackular JP, Rogers MA, Ruffin MTVI, et al. The human gut microbiome as a screening tool for colorectal cancer. *Cancer Prev Res*. 2014; 7:1112–21.
14. O'Keefe SJ, Li JV, Lahti L, et al. Fat, fibre and cancer risk in African Americans and rural Africans. *Nat Commun*. 2015; 6:6342. [PubMed: 25919227]
15. Pitcher MCL, Beatty ER, Cummings JH. The contribution of sulphate reducing bacteria and 5-aminosalicylic acid to fecal sulphide in patients with ulcerative colitis. *Gut*. 2000; 46:64–72. [PubMed: 10601057]
16. Gibson GR, Cummings JH, Macfarlane GT. Growth and activities of sulphate-reducing bacteria in gut contents of healthy subjects and patients with ulcerative colitis. *FEMS Microbiol Ecol*. 1991; 86:101–12.
17. Roediger WE, Moore J, Babidge W. Colonic sulfide in pathogenesis and treatment of ulcerative colitis. *Dig Dis Sci*. 1997; 42:1571–9. [PubMed: 9286219]
18. Christl SU, Eisner HD, Dusel G, et al. Antagonistic effects of sulfide and butyrate on proliferation of colonic mucosa: A potential role for these agents in the pathogenesis of ulcerative colitis. *Dig Dis Sci*. 1996; 41:2477–81. [PubMed: 9011461]
19. Deplancke B, Gaskins HR. Hydrogen sulfide induces serum-independent cell cycle entry in nontransformed rat intestinal epithelial cells. *FASEB J*. 2003; 17:1310–12. [PubMed: 12738807]
20. Attene-Ramos MS, Nava GM, Muellner MG, et al. DNA damage and toxicogenomic analyses of hydrogen sulfide in human intestinal epithelial FHs 74 Int cells. *Environ Mol Mutagen*. 2010; 51:304–14. [PubMed: 20120018]
21. Attene-Ramos MS, Wagner ED, Plewa MJ, et al. Evidence that hydrogen sulfide is a genotoxic agent. *Mol Cancer Res*. 2006; 4:9–14. [PubMed: 16446402]
22. Attene-Ramos MS, Wagner ED, Gaskins HR, et al. Hydrogen sulfide induces direct radical-associated DNA damage. *Mol Cancer Res*. 2007; 5:455–9. [PubMed: 17475672]
23. Bolstad AI, Jensen HB, Bakken V. Taxonomy, biology, and periodontal aspects of *Fusobacterium nucleatum*. *Clin Microbiol Rev*. 1996; 9:55–71. [PubMed: 8665477]
24. Baron EJ. *Bilophila wadsworthia*: a unique Gram-negative anaerobic rod. *Anaerobe*. 1997; 3:83–6. [PubMed: 16887567]
25. Carbonero F, Benefiel AC, Alizadeh-Ghamsari AH, et al. Microbial pathways in colonic sulfur metabolism and links with health and disease. *Front Physiol*. 2012; 28:448.
26. Xicola RM, Gagnon M, Clark JR, et al. Excess of proximal microsatellite-stable colorectal cancer in African Americans from a multiethnic study. *Clin Cancer Res*. 2014; 20:4962–70. [PubMed: 25013126]
27. 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; 144:1181–7. [PubMed: 24919690]

28. Daly K, Sharp RJ, McCarthy AJ. Development of oligonucleotide probes and PCR primers for detecting phylogenetic subgroups of sulfate-reducing bacteria. *Microbiol.* 2000; 146:1693–705.
29. Klein M, Friedrich M, Roger AJ, et al. Multiple lateral transfers of dissimilatory sulfite reductase genes between major lineages of sulfate-reducing prokaryotes. *J Bacteriol.* 2001; 183:6028–35. [PubMed: 11567003]
30. DeSantis TZ, Hugenholtz P, Larsen N, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol.* 2006; 72:5069–72. [PubMed: 16820507]
31. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods.* 2010; 7:335–6. [PubMed: 20383131]
32. Liu TW, Park YM, Holscher HD, et al. Physical activity differentially affects the cecal microbiota of ovariectomized female rats selectively bred for high and low aerobic capacity. *PLoS ONE.* 2015; 10:e0136150. [PubMed: 26301712]
33. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011; 12:R60. [PubMed: 21702898]
34. Block G, Hartman AM, Naughton D. A reduced dietary questionnaire: development and validation. *Epidemiol.* 1990; 1:58–64.
35. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr.* 1997; 65(4 Suppl):1220S–8S. discussion 1229S–1231S. [PubMed: 9094926]
36. Nava GM, Carbonero F, Croix JA, et al. Abundance and diversity of mucosa-associated hydrogenotrophic microbes in the healthy human colon. *ISME J.* 2012; 6:57–70. [PubMed: 21753800]
37. Schroeder BO, Wu Z, Nuding S, et al. Reduction of disulphide bonds unmasks potent antimicrobial activity of human beta-defensin 1. *Nature.* 2011; 469:419–23. [PubMed: 21248850]
38. Ridlon JM, Wolf PG, Gaskins HR. Taurocholic acid metabolism by gut microbes and colon cancer. *Gut Microbes.* 2016; 7:201–15. [PubMed: 27003186]
39. Nakazawa F, Sato M, Poco SE, et al. Description of *Mogibacterium pumilum* gen. nov., sp. nov. and *Mogibacterium vescum* gen. nov., sp. nov., reclassification of *Eubacterium timidum* (Holdeman *et al.* 1980) as *Mogibacterium timidum* gen. nov., comb. nov. *Int J Syst Evol Microbiol.* 2000; 50:679–88. [PubMed: 10758875]
40. Thiesen CM, Nicolaidis M, Lökebö JE, et al. *Leptotrichia amnionii*, an emerging pathogen of the female urogenital tract. *J Clin Microbiol.* 2007; 45:2344–7. [PubMed: 17522272]
41. Collins MD, Hoyle L, Tornqvist E, et al. Characterization of some strains from human clinical sources which resemble 'Leptotrichia sanguinegens': description of *Sneathia sanguinegens* sp. nov. gen. nov. *Syst Appl Microbiol.* 2001; 24:358–61. [PubMed: 11822670]
42. Griffen AL, Becker MR, Lyons SR, et al. Prevalence of *Porphyromonas gingivalis* and periodontal health status. *J Clin Microbiol.* 1998; 36:3239–42. [PubMed: 9774572]
43. Flynn KJ, Baxter NT, Schloss PD. Metabolic and community synergy of oral bacteria in colorectal cancer. *mSphere.* 2016; 1:e00102–16. [PubMed: 27303740]
44. Kaakoush NO. Insights into the Role of *Erysipelotrichaceae* in the human host. *Front Cell Infect Microbiol.* 2015; 5:84. [PubMed: 26636046]
45. Lee JY, Arai H, Nakamura Y, et al. Contribution of the 7 β -hydroxysteroid dehydrogenase from *Ruminococcus gnavus* N53 to ursodeoxycholic acid formation in the human colon. *J Lipid Res.* 2013; 54:3062–9. [PubMed: 23729502]
46. Barrasa JI, Olmo N, Lizarbe MA, et al. Bile acids in the colon, from healthy to cytotoxic molecules. *Toxicol In Vitro.* 2013; 27:964–77. [PubMed: 23274766]
47. Wei H, Dong L, Wang T, et al. Structural shifts of gut microbiota as surrogate endpoints for monitoring host health changes induced by carcinogen exposure. *FEMS Microbiol Ecol.* 2010; 73:577–86. [PubMed: 20629751]
48. Bruinenberg PG, De Roo G, Limsowtin G. Purification characterization of cystathionine (gamma)-lyase from *Lactococcus lactis* subsp. cremoris SK11: possible role in flavor compound formation during cheese maturation. *Appl Environ Microbiol.* 1997; 63:561–6. [PubMed: 16535512]
49. Persson S, Edlund MB, Claesson R, et al. The formation of hydrogen sulfide and methyl mercaptan by oral bacteria. *Oral Microbiol Immun.* 1990; 5:195–201.

50. Göker M, Gronow S, Zeytun A, et al. Complete genome sequence of *Odoribacter splanchnicus* type strain (1651/6^T). *Stand Genomic Sci.* 2011; 4:200–9. [PubMed: 21677857]
51. Downes J, Vartoukian SR, Dewhirst FE, et al. *Pyramidobacter piscolens* gen. nov., sp. nov., a member of the phylum ‘Synergistetes’ isolated from the human oral cavity. *Int J Syst Evol Microbiol.* 2009; 59:972–80. [PubMed: 19406777]
52. Chan DS, Lau R, Aune D, et al. Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. *PLoS ONE.* 2011; 6:e20456. [PubMed: 21674008]
53. David LA, Maurice CF, Carmody RN, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature.* 2014; 505:559–63. [PubMed: 24336217]
54. Finegold SM, Vaisanen ML, Molitoris DR, et al. *Cetobacterium somerae* sp. nov. from human feces and emended description of the genus *Cetobacterium*. *Syst Appl Microbiol.* 2003; 26:177–81. [PubMed: 12866843]
55. Devkota S, Wang Y, Musch MW, et al. Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in IL10–/– mice. *Nature.* 2012; 487:104–8. [PubMed: 22722865]
56. Willett WC, Stampfer MJ, Colditz GA, et al. Relation of meat, fat, and fiber intake to the risk of colon cancer in a prospective study among women. *N Engl J Med.* 1990; 323:1664–72. [PubMed: 2172820]
57. Fung TT, Hu FB, Wu K, et al. The mediterranean and dietary approaches to stop hypertension (DASH) diets and colorectal cancer. *AmJClin Nutr.* 2010; 92:1429–35.
58. Darfeuille-Michaud A, Boudeau J, Bulois P, et al. High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn’s disease. *Gastroenterology.* 2004; 127:412–21. [PubMed: 15300573]
59. Swidsinski A, Ladhoff A, Pernthaler A, et al. Mucosal flora in inflammatory bowel disease. *Gastroenterology.* 2002; 122:44–54. [PubMed: 11781279]
60. Zoetendal EG, von Wright A, Vilpponen-Salmela T, et al. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Appl Environ Microbiol.* 2002; 68:3401–7. [PubMed: 12089021]
61. Tang MS, Poles J, Leung JM, et al. Inferred metagenomic comparison of mucosal and fecal microbiota from individuals undergoing routine screening colonoscopy reveals similar differences observed during active inflammation. *Gut Microbes.* 2015; 6:48–56. [PubMed: 25559083]
62. Parthasarathy G, Chen J, Chen X, et al. Relationship between microbiota of the colonic mucosa vs feces and symptoms, colonic transit, and methane production in female patients with chronic constipation. *Gastroenterology.* 2016; 150:367–79. [PubMed: 26460205]
63. Kipnis V, Midthune D, Freedman LS, et al. Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am J Epidemiol.* 2001; 153:394–403. [PubMed: 11207158]
64. Subar AF, Freedman LS, Tooze JA, et al. Addressing current criticism regarding the value of self-report dietary data. *J Nutr.* 2015; 145:2639–45. [PubMed: 26468491]

Significance of this study

What is already known on this subject?

- ▶ The colonic microbiome of patients with colorectal cancer is different from healthy subjects, but the impact of these differences remains uncertain.
- ▶ A diet high in red meat and animal fat (a so-called Western diet) is associated with greater colorectal cancer risk.
- ▶ African Americans have a significantly higher incidence of colorectal cancer compared with other Americans, possibly a consequence of environmental rather than genetic risk factors. Diet may play an important role in the higher incidence of colorectal cancer in African Americans.
- ▶ African Americans are diagnosed with colorectal cancer at earlier ages and have a higher proportion of proximal colorectal cancers compared with non-Hispanic whites.
- ▶ Hydrogen sulfide, which is produced by sulfidogenic bacteria, triggers proinflammatory and hyperproliferative pathways, and it is genotoxic.
- ▶ People consuming a Western diet have a higher abundance of both primary and toxic secondary bile acids. Taurine-conjugated bile acids may serve as a substrate for the sulfidogenic bacterium *Bilophila wadsworthia*.

What are the new findings?

- ▶ The abundance of sulfidogenic bacteria is higher in colonic mucosa of African Americans compared with non-Hispanic whites, regardless of disease status.
- ▶ The sulfidogenic bacteria, *Bilophila wadsworthia* and *Pyramidobacter* spp., are significantly more abundant in African American colorectal cancer cases compared with African American controls.
- ▶ Abundance of sulfidogenic bacteria correlated positively with the components of a diet high in fat and animal protein, and negatively with the servings of dairy and calcium.

How might it impact on clinical practice in the foreseeable future?

- ▶ This study reflects health disparities and dietary differences in the African American compared with the non-Hispanic white population in Chicago. Practice aimed to make a healthy diet accessible for this population may have a significant effect on long-term health outcomes.
- ▶ Dietary and pharmaceutical interventions focused on reducing sulfidogenic bacteria and secondary bile acids, or supporting fermentative modes of metabolism by sulfidogenic bacteria, may reduce colorectal cancer risk.

- ▶ Measuring sulfidogenic bacterial abundance from biopsies during routine colonoscopy may constitute biomarkers for increased risk of colorectal cancer, particularly in African Americans.

Figure 1.

Differences in mean gene copy numbers of sulfidogenic bacteria comparing colorectal cancer cases and controls in African Americans (AAs) and non-Hispanic whites. A. Scatterplot representations of mean gene copy numbers per nanogram of DNA of sulfidogenic bacterial targets in African Americans compared with non-Hispanic whites with median and upper and lower quartiles indicated. African Americans are represented in green, and non-Hispanic whites are represented in purple. All comparisons had a $p < 0.001$. B. Scatterplot representations of mean gene copy numbers per nanogram of DNA of sulfidogenic bacterial targets in cases and controls in each racial group. African Americans

cases are represented in light green and controls in dark green. Non-Hispanic white cases are represented in light purple and controls in dark purple.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2.

Analysis of bacterial 16S rDNA data identifies differences in bacterial abundances comparing African American colorectal cancer (CRC) cases and controls. *A.* Differences identified by linear discriminant analysis. Effect sizes were calculated by the LEfSe algorithm and genera associated with effect size differences $-\log_2$ and $\log_2$ are shown. Genera associated with control samples are in green, and genera associated with CRC are in red. *B.* Scatterplot representations of the bacterial genera identified by LEfSe analysis in cases and controls. Each point is the logarithm of the proportion of each indicated bacterial genera with median and average indicated. The null measurements are not indicated on the plot. The numbers of nulls for each subject (case/control) by genera is *Faecalibacterium* 1 case/1 control, *Pseudomonas* 22/31, *Peptostreptococcaceae* 58/64, *Alistipes* 61/94, *Delftia* 62/79, *Mogibacteriaceae* 62/94, *Micrococcus* 62/94, *Pyramidobacter* 58/92, *Christensenellaseae* 54/92, Cc_115 *Erysipelotrichaceae* 60/93, *Bacteroidales S24-7* 49/81 and *Ruminococcus* 32/65. *C.* Scatterplot representations of abundances of three selected sulfidogenic bacteria—*Fusobacterium*, *Desulfovibrio* and *Bilophila*—shown as in *B.* The numbers of nulls for each subject (case/control) by genera are *Fusobacterium* 34/55, *Bilophila* 52/81 and *Desulfovibrio* 57/83. LDA, linear discriminant analysis.

Figure 3.

Linear discriminant analysis identifies differences between colorectal cancer (CRC) cases and controls. Effect size calculated by the LefSe algorithm identifies genera strongly associated with differences between CRC cases and controls in right-sided colonic biopsies (A) and in left-sided colonic biopsies (B), as well as differences based on age, comparing cases <50 and controls <50 (C) and comparing cases <50 and cases ≥50 (D). Genera associated with control samples are in green and genera associated with CRC are in red in panels A–C; genera associated with cases <50 are in red and cases ≥50 are in green in panel D. LDA, linear discriminant analysis.

Figure 4.

Comparison of selected factors that explain differences in sulfidogenic bacteria in African Americans and non-Hispanic whites. A. Scatterplot representations of residuals of linear models for selected bacterial targets, namely, pan-*dsrA*, *Desulfovibrio* spp. and *B. wadsworthia*. African Americans are represented in green, and non-Hispanic whites are represented in purple. Cases are represented as open circles and controls as closed circles. B. Forest plots of effect size estimates from the linear models. Outcome variables are log-transformed gene copy abundances of pan-*dsrA*, *Desulfovibrio* spp. and *B. wadsworthia*. Points and lines on each plot represent the point estimate and 95% CI, respectively, for each

covariate in the model. Estimates reflect predicted change in variable for one unit change in outcome. Positive associations are in blue, negative associations are in red. p value for each covariate is represented on the right, and statistically significant values after Bonferroni correction are in bold. The wide CI seen in the body mass index (BMI), underweight class is explained by the small number of observations in this class.

Table 1

Clinical and demographic characteristics of African American (AA) and non-Hispanic white (NHW) subjects

Clinical variable	AAs (n=197)			NHWs (n=132)						
	Controls	N	p Value	Controls	N	p Value				
Age, mean in years±SD	55.5±9.6	100	59.8±10.8	97	0.003	55±12.91	76	59.6±10.1	56	0.047
Biopsy, left/right large bowel *	55/45	100	55/42	97	0.886	50/25	75	42/13	55	0.248
Gender, F/M	46/54	100	40/57	97	0.566	32/44	76	18/38	56	0.28
Weight, mean in lb±SD	198.1±43.9	100	178.8±41.9	97	0.007	177.3±37.1	76	189.1±45.1	56	0.024
BMI, kg/m ² ±SD	30.4±6.1	100	28.3±6.5	97	0.043	27.5±6.5	76	28.4±5.9	56	0.046
College education, % [†]	59.6	99	35.5	97	<0.001	69.7	76	52.7	55	0.047
Annual income, %		100		96	0.032		75		55	0.384
<25 000	60	60	75	72		36	27	47.3	26	
25 000 –75 000	29	29	21.9	21		36	27	32.7	18	
750 000	11	11	3.1	3		28	21	20	11	
Omega-3, % users [‡]	27.6	87	14.6	41	0.174	50	50	27.3	22	0.07
Alcohol, % current users	31.3	99	24.2	91	0.33	48.7	76	32.1	53	0.06
Smoking, mean pack years	6.1±10.7	46	7.6±14.8	70	0.73	12±21.8	39	14.9±26.3	44	0.92
NSAID use in years [§]	4.9±6.8	93	3.03±7.8	95	0.084	7.6±9.2	75	3.1±7	51	<0.001
TNM stage [¶]				89					50	
Stage 0 and 1			28 (31.5%)					14 (28%)		
Stage 2			23 (25.8%)					16 (32%)		
Stage 3			23 (25.8%)					12 (24%)		
Stage 4			15 (16.9%)					8 (16%)		

Statistical comparisons between cases and controls within each ethnic group were performed using Wilcoxon rank-sum or t-test for continuous variables and χ^2 or Fisher's exact test for dichotomous variables. Annual income in US\$.

^{*} The right side of the large bowel includes the caecum and appendix, ascending colon, hepatic flexure and transverse colon. The left side includes the splenic flexure, descending colon, sigmoid colon and rectum.

[‡] Percent of subjects who graduated from a 4-year undergraduate scholastic programme.

[§] Percent of subjects who reported dietary supplementation with omega-3 fatty acids in the last 5 years.

[¶] NSAID used included Ibuprofen, Advil or Motrin.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

¶ The p value for the comparison of stages between AA CRCs and NHW CRCs was 0.89.

BMI, body mass index; CRC, colorectal cancer; N, number of subjects from whom information was available; NSAID, non-steroidal anti-inflammatory drug; TNM, tumour node metastasis classification of malignant tumours.