Protocol Title. Diet modulation of bacterial sulfur & bile acid metabolism and colon cancer risk
Funding sources: National Cancer Institute/National Institute of Minority Health and Health Disparities

UIC Investigators:
PI: Lisa Tussing-Humphreys, PhD, MS, RD, Assistant Professor
Department and Institution: University of Illinois at Chicago Department of Medicine, Division of Academic Internal Medicine
Address and Contact Information: 1747 W. Roosevelt Road, Rm. 558, Chicago, IL 60608
Phone: 312-355-5521, Email: ltussing@uic.edu

Co-I: Barbara Jung, MD, Associate Professor
Department and Institution: University of Illinois at Chicago Department of Medicine, Division of Gastroenterology and Hepatology
Address and Contact Information: 840 S. Wood Street, 718-E, Chicago, IL 60612
Phone: 312-3996-1565, Email: bjung@uic.edu

Co-I: Seijuong Kim, PhD, Associate Professor
Department and Institution: University of Illinois at Chicago School of Public Health, Division of Health Policy and Administration
Address and Contact Information: 823 SPH-PI, 1603 W. Taylor St. Chicago, IL 60612
Phone: (312) 413-3544, Email: skim49@uic.edu

Co-I: Cemal Yazici, MD, Fellow
Department and Institution: University of Illinois at Chicago Department of Medicine, Division of Gastroenterology and Hepatology
Address and Contact Information: 840 S. Wood Street, 718-E, Chicago, IL 60612
Phone: 312-996-6651, Email: cyazic2@uic.edu

Co-I: Sally Freels, PhD, Associate Professor
Department and Institution: University of Illinois at Chicago, School of Public Health, Division of Epidemiology and Biostatistics
Address and Contact Information: 953 SPHIP
Phone: 312-996-0064, Email: sallyf@uic.edu

Co-I: Carol Braunschweig, PhD, Professor
Department and Institution: University of Illinois at Chicago, Department of Kinesiology and Nutrition
Address: 1919 W. Taylor St. Rm 650
Phone: 312-996-2575; Email: braunsch@uic.edu

UIC Research Staff:

Neda Khudeira, Graduate Student
Department and Institution: University of Illinois at Chicago, Department of Kinesiology and Nutrition
Address: 1919 W. Taylor Street, 6th Floor
Email: nkhude4@uic.edu

Bazil Brainard, Graduate Student
Department and Institution: University of Illinois at Chicago, Department of Kinesiology and Nutrition
Address: 1919 W. Taylor Street, 6th Floor
Email: bbrain2@uic.edu

Research Protocol V 8.0 11/1/2017
Teresa Schering, Graduate Student  
Department and Institution: University of Illinois at Chicago, Department of Public Health  
Address: 1603 West Taylor Street, Chicago  
Email: tscher2@uic.edu

Taneisha Griggs, Graduate Student Roosevelt University  
Address: 1747 W. Roosevelt Rd, Chicago, Room 558  
Email: griggsdorshea@gmail.com

Phyllis Petrelli, Research Chef  
Department and Institution: University of Illinois at Chicago, Department of Medicine  
Address: 1919 W. Taylor Street, 6th Floor

Andrew McLeod, Graduate Student  
Department and Institution: University of Illinois at Chicago, Department of Kinesiology and Nutrition  
Address: 1919 W. Taylor Street, 6th Floor  
Email: amcleo2@uic.edu

Claudia Garcia  
Cancer Center, Diet and Behavior Shared Resource  
1747 W. Roosevelt Road, Chicago, IL 60608  
claudiag@uic.edu

Alyshia Hamm, RD  
Research Specialist, Department of Medicine  
1747 W. Roosevelt Road, Chicago, IL 60608  
Ahamm6@uic.edu

Shirley Spencer, PhD  
P20 Cancer Center Trainee/Governor's State University  
1747 W. Roosevelt Road, Chicago, IL 60608  
SSpencer@govtst.edu

Betty Li  
UIC Medical Student  
Bettyli2@uic.edu  
1747 W. Roosevelt Road, Chicago, IL 60608

Nancy Krett, Research Coordinator  
Department and Institution:  
University of Illinois at Chicago Department of Medicine, Division of Gastroenterology and Hepatology  
Address: 840 S. Wood Street CSB 718E  
Phone: 312-413-7330; Email: nkrett@uic.edu

Robert Carroll, MD  
Department and Institution: University of Illinois at Chicago Department of Medicine, Division of Gastroenterology and Hepatology  
Address and Contact Information: 5137 COMRB, MC 716, Chicago, IL 60612  
Phone: 312-355-4198; Email: rcarroll@uic.edu

Research Protocol V 8.0 11/1/2017

Page 2 of 49
Brian Boulay, MD, MPH
Department and Institution: University of Illinois at Chicago Department of Medicine, Division of Gastroenterology and Hepatology
Address: 840 S. Wood Street CSB 718E
Phone: 312-996-6651 Email: bboulay@uic.edu

Haya, Al-Rashdan, Clinical Research Coordinator
Division of Gastroenterology and Hepatology
Address: 840 S. Wood Street, 702 CSB
Email: halras2@uic.edu

**Rush University Medical Center investigator and research staff:**
**Site PI:** Ece Mutlu, MD, MBA, MSCR, FACG, AGAF, Associate Professor
Department and Institution: Rush University Medical Center, Department of Medicine, Section of Gastroenterology, Hepatology and Nutrition
Address and Contact Information: 1725 W. Harrison, Suite 206, Chicago, IL 60612
Phone: 312-563-3880, Email: ece_mutlu@rush.edu

Prachi Chakradeo, Clinical Research Coordinator
Department and Institution: Rush University Medical Center, Department of Medicine, Section of Gastroenterology, Hepatology and Nutrition
Address and Contact Information: 1725 W. Harrison, Suite 206, Chicago, IL 60612
Phone: 312-942-3644, Email: Prachi_Chakredo@rush.edu

Sandra Gomez-Perez, Assistant Professor
Department and Institution: Rush University Medical Center, Department of Clinical Nutrition
Address and Contact Information: 1700 West Van Buren, Suite 425 TOB, Chicago, IL 60612
Phone: (312) 942-5297

Sarah Goldberg, Clinical Research Coordinator
Department and Institution: Rush University Medical Center, Department of Medicine, Section of Gastroenterology, Hepatology and Nutrition
Address: Rush University Medical Center, 1725 W. Harrison St, Suite 206, Chicago, IL, 60612
Email: sarah_a_goldberg@rush.edu

**UIUC Investigators:**
**Contact PI:** H. Rex Gaskins, PhD, Professor
Department and Institution: University of Illinois at Urbana-Champaign, Carl R. Woese Institute for Genomic Biology
Address and Contact Information: 1206 W. Gregory Drive, Urbana, IL 61801
Phone: 217-244-3165, Email: hgaskins@illinois.edu

**PI:** Jason Ridlon, PhD, Assistant Professor
Department and Institution: University of Illinois at Urbana-Champaign, Carl R. Woese Institute for Genomic Biology
Address and Contact Information: 1206 W. Gregory Drive, Urbana, IL 61801
Phone: 217-625-0832, Email: jmridlon@illinois.edu

Patricia Wolf, Graduate Student
Department and Institution: University of Illinois at Urbana-Champaign, Division of Nutritional Sciences
Address and Contact Information: 1206 W. Gregory Drive, Urbana, IL 61801
Email: gondry2@illinois.edu

Research Protocol V 8.0 11/1/2017

Page 3 of 49
Specific Aims. Colorectal cancer (CRC) is the third most prevalent and the second leading cause of cancer deaths in the United States. Among all racial/ethnic groups the African American (AA) population exhibits the highest CRC incidence and mortality. Abundant evidence indicates that a diet high in red meat and animal fat poses significant risk for the development of CRC. We previously completed a strictly controlled diet intervention study in which the diets of AA (high meat and fat, low fiber) and native Africans (NA) (low meat and fat, high fiber) were switched to examine the hypothesis that diet mediates cancer risk through its effects on microbiota metabolism. These data provide unequivocal evidence that diet dramatically affects colonic microbiota composition and that microbial utilization of dietary nutrients can create a metabolic milieu in the colon that can either promote or protect from CRC risk. Most recently, we also observed that the mucosal abundance of Bilophila wadsworthia distinguished AA colon cancer patients from non-Hispanic white (NHW) colon cancer patients and controls. Our observation of this significant racial difference in the abundance of a bile acid metabolizing sulfidogenic bacterium involved in bile acid metabolism motivates our objective of determining the extent to which this colonic microbial niche is affected by race and diet.

We therefore hypothesize that taurocholic acid (TCA) is a key diet-controlled metabolite whose metabolism by Bilophila wadsworthia and Clostridium scindens yield a carcinogen and a tumor-promoter, respectively. We also hypothesize that the colonic microbiota (including the latter two bacteria) can be specifically modulated by altering dietary intake of taurine and saturated fat (found in high red meat diets) in AA subjects at elevated risk for colon cancer. In the liver, bile acids are conjugated to taurine or glycine before gallbladder storage; the contents of which are secreted into the small bowel to aid in lipid absorption. Diets high in red meat and animal fat induce greater bile secretion, and shifts bile acid conjugation from glycine to taurine. Bile acids are a selective agent determining microbiome composition and Bilophila, as its name implies, thrives in bile. Culture-based and animal model studies show that Bilophila liberates cholic acid (CA) and consumes the sulfite of taurine in anaerobic respiration, releasing genotoxic hydrogen sulfide (H2S). Reciprocal diet exchange studies demonstrate that B. wadsworthia is stimulated by animal-based diets. C. scindens metabolizes CA, releasing the tumor-promoter, deoxycholic acid (DCA). Our published data highlight the role of hydrogen sulfide (H2S) as a bacterial-derived intestinal insult that leads to activation of proinflammatory pathways and genomic instability—two hallmarks of CRC. Specifically, we have demonstrated that: (a) H2S induces proliferative, proinflammatory and DNA repair pathways in nontransformed intestinal epithelial cells, (b) exogenous H2S is a potent genotoxin at concentrations many-fold less than those measured in the human colon, (c) H2S induces free radical-based DNA damage in the absence of cellular metabolism, and (d) colonic mucosa of healthy human subjects is persistently colonized by sulfidogenic bacteria which utilize both inorganic and organic sources of sulfur including the taurine respiring bacterium B. wadsworthia. Most recently we observed that the mucosal abundance of B. wadsworthia distinguished AA but not non-Hispanic white (NHW) colon cancer patients and controls. A large body of literature encompassing epidemiological studies, as well as mechanistic studies demonstrates that DCA promotes CRC. DCA is a logical candidate for CRC promotion for the following reasons: (a) DCA is found in high levels (>100 μM) in fecal water; (b) it can cross biological membranes via passive diffusion; (c) it activates mammalian cell signaling pathways that are known to be involved in promoting colon carcinogenesis; and (d) serum DCA levels are strongly correlated with colon cancer risk. By examining serum bile acids, we can observe diet-induced changes in taurine-conjugation (bile salts returning from ileum to liver), and DCA levels accumulating in the biliary pool (passive diffusion from colon into portal circulation).
Our aims are as follows:

Aim 1: Determine the extent of differences in endpoint bacterial metabolism of taurine and bile acids and compare to colonic inflammation between AA and NHW subjects at elevated risk (with adenomatous polyps) for CRC. We will measure colonic mucosal abundance of bacterial genes associated with sulfur and bile acid metabolism, stool bile salt hydrolase and bile acid 7a-dehydroxylating activities, and serum taurine:glycine primary bile acid ratios and secondary bile acids absorbed from the gut and colonic expression of pro-inflammatory cytokines and Cox-2 as measures of inflammation.

Aim 2: Determine in the context of a diet-intervention trial the role of taurocholic acid metabolism by gut bacteria in AA subjects at elevated risk for CRC. Two isocaloric diets, an animal-based High in Taurine and SATurated fat (HT-HSAT) and a plant-based, Low in Taurine and Low SATurated fat (LT-LSAT) will be used to determine the extent to which the relationship between diet (independent variable) and mucosal markers of CRC risk including epithelial proliferation, oxidative stress, DNA damage, and primary and secondary bile acid pools and biomarkers of inflammation (dependent variables) is explained by the abundance of sulfidogenic bacteria and H2S concentrations &/or DCA and DCA-producing C. scindens (mediator variables).

These studies will generate novel information on a mechanistically targeted nutrient (taurine) that can be used to develop effective cancer prevention interventions based simply on diet that may contribute to a reduction in the unequal colon cancer burden in AA men and women.

Significance.

Epidemiology of Colon Cancer: Cancer is among the most significant contributors of health care spending in the United States (US) resulting in estimated overall costs of 201.5 billion dollars. Colorectal cancer (CRC) is the third most frequent cancer worldwide with 142,882 new CRC cases and 50,830 CRC related deaths in the US in 2013. Among all racial/ethnic groups the African American (AA) population exhibits the highest CRC incidence and mortality. Both genetic factors and environmental exposures play significant roles in colorectal carcinogenesis, which develops through a multistep mutational process with many associated epigenetic changes. While the incidence of CRC is higher in AA populations with a more aggressive manifestation of the disease, very few studies have addressed both biological and environmental mechanisms potentially contributing to CRC in AAs. Studies have observed extraordinary low incidence of CRC in native Black Africans relative to AA, Caucasian Africans, and Caucasian Americans. Berg observed CRC incidence similar to Western populations in 1st and 2nd generation descendants of native African (NA) immigrants who eschewed the traditional diet. Also, NA eating their traditional diet (high maize; low red meat) consumed a third of the animal protein and fat than Westerners, and had significantly lower levels of fecal H2S, H2S-producing B. wadsworthia. As illustrated in Fig. 1, we question how diets high in animal protein and fat lead to higher fecal H2S and enhanced microbial bile acid metabolism and how this relates to the etiology of CRC. Recent work by PI Gaskins has shed considerable light on the influence of diet, microbial metabolites, and CRC risk between healthy NA and AA. NA were found to have significantly lower total fecal bile acids and deoxycholic acid (DCA) than AA. A two-week diet exchange study in which NA consumed an animal-based diet and AA consumed a plant-based diet demonstrated rapid reciprocal changes in levels of bacteria responsible for H2S production (B. wadsworthia, Desulfovibrio spp) and protective butyrate production, as well as levels of DCA producing bacteria (Clostridium scindens) and DCA in stool. These results confirm a recent short-term diet exchange focusing on animal- and plant-based diets, which found enhanced DCA formation and increased levels of taurine-utilizing

Research Protocol V 8.0 11/1/2017

Page 5 of 49
Diet, Sulfide, and CRC: Fecal H2S levels have been shown to be significantly higher in patients with CRC and those who have undergone surgery for sigmoid cancer. Two independent studies demonstrate associations between cysteine-utilizing sulfidogenic bacteria and the tumor surface in a subset of CRC patients. There is also clear evidence for a relationship between dietary animal protein and sulfide production in the human colon. In both a human feeding study and a model culture system, Magee and coworkers observed a significant, dose-related increase in fecal sulfide concentrations with meat intake, with H2S generation in fecal batch cultures supplemented with both bovine serum albumin and casein correlating with protein digestion. Two studies in rats also demonstrate that increased protein intake and increased protein fermentation to be associated with genotoxicity. Our data showing H2S is genotoxic at doses found in the colon provide a reasonable explanation for these observations, and indicate that H2S could be a significant diet-modulated metabolite that initiates colon cancer.

The rationale for this hypothesis is supplemented by our long-term study of mammalian cell responses to the effects of exogenous H2S. In a series of prior publications, we demonstrated that H2S induces proliferative and inflammatory pathways in nontransformed intestinal crypt epithelial cells and provided unequivocal evidence that H2S is genotoxic at concentrations less than measured previously in the human colon leading to cell-cycle arrest; and DNA damage, which is produced by oxidative stress. More recently, we confirmed the genotoxic properties of H2S in nontransformed human intestinal epithelial cells and showed that H2S modulates the expression of genes involved in cell-cycle progression, triggering both inflammatory and DNA repair responses. Altogether, our observations constitute an extensive collection of data supporting the hypothesis that H2S production by mutualistic bacteria residing within the colonic mucosa serves as a key environmental carcinogen contributing to CRC risk. Our studies will be the first to examine H2S as a diet-responsive microbial metabolite that could alter CRC risk in humans in vivo. Characterization of the mechanisms by which microbial-derived H2S is a link between dietary intake (especially of animal protein and fat) and CRC is our primary objective and has not been studied before.

Western Diet, DCA, and CRC: During the 1970’s epidemiological data pointed to higher fecal concentrations of bile acids in high risk populations of Americans on Western diets relative to low-risk groups (Japanese, Chinese, Seventh Day Adventists), that CRC and patients with adenomas have higher fecal bile acids, and that this was a result of animal protein and saturated fat. The role of DCA as a tumor promoter was established in animal models during this period. Single-dose intrarectal infusions of the carcinogen N-methyl-N-nitro-N-nitrosoguanidine followed by repeated intrarectal doses of DCA significantly induced colorectal neoplasms, while DCA infusion alone was insufficient to induce tumor formation in both germ-free and conventional rats. Thus, a genotoxin, such as H2S, is required to cause cancer; however, DCA

Research Protocol V 8.0 11/1/2017
appears to accelerate and exacerbate the problem. Later, it was shown that DCA was significantly elevated in serum of patients with colorectal adenomas,\textsuperscript{45,46} that serum DCA correlates with DCA in fecal water,\textsuperscript{47} and that colonic mucosal proliferation is correlated with DCA in serum.\textsuperscript{48}

The mechanisms by which the bile acid DCA promotes CRC have been largely elucidated. The detergent properties of DCA (but not CA) cause membrane perturbations leading to the release of arachidonic acid, which is converted by the enzymes cyclooxygenase 2 (COX-2) and lipoxygenase, to pro-inflammatory and pro-angiogenic prostaglandins, and reactive oxygen species which damage DNA and inhibit DNA repair enzymes.\textsuperscript{21} DCA also induces COX-2 expression through transactivation of the epidermal growth factor receptor,\textsuperscript{49,50} and activates the 1-catenin cell-signaling pathway resulting in colon cancer cell proliferation and invasiveness. Notably, a large body of published work demonstrates the effectiveness of non-steroidal anti-inflammatory drugs (NSAIDs) in significantly reducing polyp formation and reducing CRC risk by inhibiting the activity of COX-2.\textsuperscript{52} Mutations at the adenomatous polyposis coli (APC) locus are a common and early somatic event in polyp formation and CRC.\textsuperscript{53} Conspicuously, COX-2 can be downregulated by wild type but not mutant APC.\textsuperscript{54,55} APC regulates 1-catenin levels in the cell, failure of which results in accumulation of transcriptionally active 1-catenin, which both binds to the COX-2 promoter region activating transcription, and COX-2 mRNA increasing stabilization.\textsuperscript{54,56} In addition, DCA can activate proteosomal degradation of the tumor suppressor p53 selecting for cells resistant to apoptosis in spite of DNA damage.\textsuperscript{57} Intriguingly, PI Gaskins and coworkers observed that H2S upregulated COX-2 mRNA expression nearly 8-fold in human nontransformed epithelial cells.\textsuperscript{17} Thus, metabolism of TCA by gut bacteria releases the proinflammatory and genotoxic gas H2S that may lead to events such as APC mutations, loss of control of which allows DCA to exacerbate the progression from adenoma to adenocarcinoma by further inducing pro-inflammatory pathways, reactive oxygen species, inactivation of p53 and loss of cell-to-cell adhesion.

Combined Metabolic Potential for Carcinogen and Tumor-promoter by TCA: Another mechanism by which an animal based diet may increase proinflammatory and genotoxic H2S production is through taurine respiration by \textit{B. wadsworthia} or related taurine-respiring organisms. \textit{B. wadsworthia}, a member of the \textit{Desulfovibrionaceae} family, is normally present in low abundance in healthy individuals, but often overrepresented in pathological conditions like appendicitis and other intestinal inflammatory disorders.\textsuperscript{58,59} It generates H2S through the utilization of taurine as a terminal electron acceptor in anaerobic respiration. Recent studies suggest an intriguing link between dietary fat and the utilization of taurocholic acid as a source of taurine for \textit{B. wadsworthia}.\textsuperscript{13,62} Compared to mice fed a low fat or high polyunsaturated fat diet, mice fed a high milk fat diet demonstrated an overrepresentation of \textit{B. wadsworthia}, which was mediated by taurine-conjugated bile acids.\textsuperscript{13} To our knowledge, the impact of taurine, regardless of the source, or \textit{B. wadsworthia} abundance of in human colonic mucosa has not been studied, nor has the distribution of taurine utilization among colonic bacteria. Furthermore, the relationship of taurine utilization to potential inflammatory and genotoxic pathways important in CRC development is unknown in humans and these studies will generate data that cannot be generated in rodent models, due to different taurocholate metabolism compared to humans. Taurine conjugation in rodents is genetically determined and nearly all murine bile acids are taurine-conjugated; however, in humans, bile acid taurine:glycine conjugation ratio is \textit{diet-dependent}. Feeding taurine, but not glycine can substantially alter the glycine:taurine ratio.\textsuperscript{10} NA consuming their traditional diet have taurine:glycine ratios of 1:9,\textsuperscript{63} while ratios in those consuming high meat and seafood diets were 10:1.\textsuperscript{10} Animal-based diets favoring taurine-conjugation also lead to enhanced bile acid secretion, providing substrate to support the growth of DCA-producing bacteria. Finally, metabolism of TCA was shown to stimulate DCA production in vitro through production of H2S.\textsuperscript{64} \textit{Taken together, our hypothesis provides a probable link between diet, metabolism of TCA, and CRC, a link plausibly higher for AAs who develop CRC based on our preliminary data. No study to date has measured the effect of diet on taurine-conjugation of bile acids.
and *B. wadsworthia* abundance in humans, especially in AAs in whom the mucosal abundance of this taurine respiring bacterium is significantly greater in CRC cases. This is both a major gap in our knowledge, and the central focus or our proposal.

We have shown in published work and preliminary data (Fig. 2) that sulfidogenic bacteria, which use organic sources of sulfur are present in colonic mucosa of normal healthy subjects. Thus, we predict that bacterial H2S production is common in the colon of persons who consume diets high in organic sulfur. In future studies, we will extend this hypothesis to study the role of host genetic background in dictating the degree to which H2S production promotes CRC, especially in AAs. We envision that combinations of common polymorphisms in genes operative in a variety of pathways responsive to this bacterial catabolite, including those involved in H2S detoxification (also poorly understood), regulation of inflammation, cell-cycle control and responsiveness to DNA damage create a predisposing genetic background that imparts CRC risk in response to a Western diet high in animal protein and fat. The present proposal seeks to initiate examination of this multifactorial model of CRC risk in humans. This research has the potential to identify multiple mechanisms in inflammatory or genotoxic pathways by which diet and sulfidogenic bacteria interact to contribute to CRC risk. If particular sulfidogenic bacteria or their taurine degradation pathways are linked with specific genotoxic or inflammatory mechanisms in humans and especially in AAs, these bacteria or their metabolic pathways could be targeted in future prevention efforts to alter sporadic CRC risk. Similarly, the identification of specific mechanistic links between sulfidogenic bacteria and CRC risk in AAs would provide additional measures to determine success of specific prevention interventions and potential surrogate markers of colonic and overall health especially for those AAs at high risk for CRC development. Ultimately, these studies will generate novel information on a mechanistically targeted nutrient that can be used to develop effective cancer prevention interventions based simply on diet that may contribute to a reduction in the unequal CRC burden in AAs.

**Figure 2:** *B. wadsworthia* specific dsrA is more abundant in non-involved mucosa samples of African American colorectal cancer cases (n=100) compared to healthy African American controls (n=100). p<0.01

**Innovation.** Proposed are innovative studies designed to investigate the biological basis of increased CRC risk independently associated with being AA or consuming a high red meat and saturated fat diet. The focus follows longstanding work on this health disparity in the academic institutions from Chicago.

We will establish the extent of baseline differences in endpoints of bacterial taurine and bile acid metabolism between elevated risk AA and NHW subjects with elevated risk of CRC (defined as those subjects having adenomatous polyps). This objective is based on the entirely novel observation that mucosal abundance of taurine metabolizing bacterium *B. wadsworthia* is a strong predictor of CRC cases in AA but not NHW subjects. Bacterial deconjugation of the bile acid taurocholate provides another major source of taurine, and once deconjugated, free primary bile acids are further metabolized to genotoxic and proinflammatory secondary bile acids. The possibility that racial differences exist in modes of bacterial bile acid metabolism has not been addressed. Thus, Aim 1 will provide a wealth of novel data on this question in the context of potential differences between high and low risk AAs and their NHW counterparts.

**Aim 2** is designed to determine mechanistically why consumption of a high red meat and saturated fat diet imparts risk for CRC development and to demonstrate that primary microbial risk factors (sulfidogenic and bile acid metabolizing bacteria) are modifiable by diet. The focus is on taurine, an overlooked sulfur amino acid (SAA) that is abundant in red meat or provided by bacterial deconjugation of the bile salt TCA, which is increased in subjects consuming a diet high in saturated fat. Rationale for focusing the diet intervention study on AAs comes from the previously mentioned observation that a taurine respiring bacterium distinguished AA
but not NHW CRC patients from healthy controls, and the previous work by PI Gaskins in AA subjects focused on mechanisms underlying the increased risk for CRC associated with consumption of a Western type diet.

Our strong collection of past publications and new preliminary data support our hypothesis that dietary sources of organic sulfur increase the abundance of microbes that generate H2S through taurine metabolism and that H2S activates proinflammatory pathways and serves as a genotoxin in the colonic mucosa. **We’re examining, for the first time bacteria that utilize taurine, which can be provided directly from red meat or indirectly through TCA in response to saturated fat.** Our study will be the first to examine the consequences of such specific dietary manipulation on genotoxic or inflammatory pathways implicated in CRC development in at-risk AAs.

Our results will provide novel information regarding the in vivo interactions between diet and cancer that heretofore have not been explored in humans, particularly AAs. Food taurine content is not currently provided in either the University of Minnesota Nutrition Data System for Research (NDSR) or the USDA Standard Reference (USDA SR) nutrient databases, which are the gold standard sources for the nutrient content of food. Evidence that taurine is capable of inducing biomarkers of CRC risk through promoting growth of Sulfidogenic *B. wadsworthia* or other untargeted bacteria would be an important novel observation justifying the addition of this SAA to these nutrient databases. If our hypothesis is substantiated, simple vigilance of taurine intake might diminish susceptibility to CRC in all individuals, especially AAs at elevated risk. Further, if our hypothesis is upheld, it might be possible to reduce risk not only by dietary intervention but also by microbiota modification (potentially through pre-, pro- or symbiotics). Finally, if our study reveals particular modes of bacterial sulfur or bile acid metabolism correlating with epithelial proliferation or inflammation in AAs, the endpoints identified can potentially predict non-invasively elevated risk individuals who should be: a) advised on specific dietary interventions (those investigated herein); b) offered specific therapy to reduce risk; or c) counseled on regular colonoscopic screening.

**Approach.** First we will examine the extent of baseline differences in mucosal abundance of bacterial genes associated with sulfur and bile acid metabolism, colonic inflammation, and serum and stool markers of bile metabolism between at-risk AAs and NHWs compared to controls. We will then test the working hypothesis that a diet high in animal protein and saturated fat creates a metabolic milieu in the colon that promotes CRC risk in at-risk AAs by increasing the abundance and activity of colonic bacteria that generate both a carcinogen and a tumor promoter through metabolism of the primary taurine conjugated bile salt TCA. The study will be a prospective randomized crossover feeding trial that compares a high taurine and saturated fat animal-based diet, to a plant-based diet low in taurine and saturated fat. Each subject will receive each of the two diets in a crossover design thereby serving as their own control; an experimental approach we’ve confirmed to be optimal when studying the highly individualistic and variable colonic microbiota. A mediation model will be used to determine the extent to which the relationship between diet [independent variable] and mucosal markers of CRC risk and DNA damage and repair [dependent variables] is explained by colonic microbiota and their functions [mediator variables].

**Aim 1)** Compare the extent of differences in endpoint bacterial metabolism of taurine and bile acids and baseline colonic inflammation between AA and NHW subjects with elevated CRC risk and controls. A case-control study will be conducted with 200 subjects (100 AA and 100 NHW) with history of adenomatous polyps (APs) (i.e., those with elevated CRC risk) and controls without APs to assess values for: 1) mucosal abundance of bacterial genes associated with sulfur and bile acid metabolism; 2) stool bile salt hydrolase and bile acid 7a-dehydroxylating activities; 3) measurement of serum bile acids to indicate the extent of taurine-conjugation of bile acids; 4) ratio of conjugated: unconjugated bile acids; 5) levels of secondary bile acids absorbed from the gut; and 6) expression of pro-inflammatory cytokines and Cox-2 as measures of colonic inflammation. Habitual and recent dietary intakes will be determined by food frequency questionnaire (FFQ), food records, and 24 hour diet recall.
**Rationale and Hypothesis:** This aim is motivated by the clear evidence that the AA population exhibits the highest CRC incidence and mortality among all racial/ethnic groups in the US.\(^3\) The objective is to follow up on our recent observation supporting the intriguing possibility that CRC disparity in AA may reflect to some extent microbiologic mechanisms operative in the colonic mucosa of AA cases. Specifically, we compared the abundance of taurine utilizing *B. wadsworthia* in non-involved colonic mucosa samples of AA CRC cases (n=100) to healthy AA controls undergoing colonoscopy (n=100). Intriguingly, *B. wadsworthia* was significantly more abundant in colonic mucosa of the AA-CRC cases than AA controls (Fig. 2) with an odds ratio of 180, independent of the effects of age, gender, site of sampling (left vs. right colon), BMI, income and education level. The mucosal abundance of this sulfidogenic bacterium was not significantly different between an equal number of NHW CRC cases and controls. The fact that this association was found in the non-involved mucosa of AA-CRC cases suggests that this is not simply due to the presence of cancer. If there is a causal link between *B. wadsworthia* and CRC, then an increase in *B. wadsworthia dsrA* abundance should also be apparent in patients with precursor lesions to CRC (i.e. those subjects who are at an elevated risk of CRC), such as those patients with a history of APs. Furthermore, in such a causal link, there should also be an observable dose-response relationship between presence of APs and *B. wadsworthia*. Since no data exists on the abundance of sulfidogenic bacteria in patients with APs in general and any racial/ethnic group differences, this aim will fill a significant gap in our understanding of racial disparities in CRC. Aim 1 is also motivated by the profound deficiency of knowledge of the effect of race/ethnicity on bile acid metabolism. There is some literature on racial dependent functional polymorphisms in genes encoding hepatic bile acid transporters, and ethnic differences in fecal 7a-dehydroxylating activity,\(^71\) but we were unable to identify any studies comparing colonic or serum concentrations of primary or secondary bile acids in AA relative to any other racial/ethnic groups. The emerging bile acid-gut microbiome axis and its role in health and disease, accentuates the importance of filling this gap in knowledge.\(^69,70\)

**Subjects.** 200 subjects (100 AA and 100 NHW) between the ages of 45 - 75 years old will be recruited from the GI endoscopy labs of UIC (n=100) and RUMC (n=100) (adjacently located on the Illinois Medical District campus) at the time of a scheduled colonoscopy. From both institutions, 4 subject groups (n=50/group) will be targeted for a case control design:

The following will be targeted at UIC:

- **Group 1 (n=25)** - AA subjects with elevated risk for CRC. These subjects are defined as having 3 or more APs or those with an AP >1 cm on previous (within 5 years) or incident colonoscopy (confirmed following pathology) at UIC, a well-accepted definition of increased risk.\(^72\)
- **Group 2 (n=25)** - AA controls, defined as having no polyps or if polyp(s) present, NOT meeting the definition of elevated risk as defined for Groups 1 and 3 following pathology on incident colonoscopy.
- **Group 3 (n=25)** - NHW at elevated risk for CRC (see above).
- **Group 4 (n=25)** - NHW controls (see above).

**Inclusion criteria.**

1. African American or Non-Hispanic White;
2. Age between 45 -75 years old;
3. Undergoing a clinically indicated screening or surveillance colonoscopy at UIC;
4. Women only: Post-menopausal (natural or surgical) defined as no menstruation in the past 6 months
5. Pathological absence (Groups 2 and 4) or NOT meeting criteria for elevated risk on incident colonoscopy OR presence of 3 or more adenomatous polyps or adenomatous polyp > 1cm on previous or incident colonoscopy at UIC/ (Groups 1 and 3).

**Exclusion criteria.**

1. Race other than Non-Hispanic White or African American
2. Women only: at least one menstrual period in the past 6 months
3. Current malignancy or malignancy in the past 5 years except non-melanoma skin cancer that has been removed
4. History of colorectal cancer.
5. Current gastrointestinal (GI) illness other than GERD or hemorrhoids (such as celiac disease, inflammatory bowel disease etc.)
6. Malabsorptive bariatric procedure (gastric bypass)
7. Chronic liver (hepatitis, elevated liver tests > 3 times the normal, etc.) or kidney disease (end stage liver disease or chronic kidney disease with creatinine above 2.0 mg/dl)
8. History of cardiac disease (such as admission for congestive heart failure within the past 5 years, on anticoagulants for heart disease, ejection fraction < 25%, coronary artery disease, peripheral artery disease (amputation, not able to walk, and/or taking anticoagulant/blood thinner), cerebrovascular disease, etc.)
9. Positive genetic test for inherited polyposis syndromes (such as familial adenomatous polyps, hereditary non-polyposis colon cancer syndromes, etc.)
10. Alcoholism or illicit drug use
11. Antibiotic use within the past 2 months
12. Regularly taking medications that may interfere with normal digestion (such as acarbose, cholestyramine, aspirin doses that exceed 81mg/day or 325 mg every other day)
13. Taking pre or pro-biotic supplements regularly (at least once per week)
14. Anticoagulant use or other factors that increase endoscopic risks
15. Non-English speaking
16. Pregnant or breastfeeding
17. Any medical condition, which, in the opinion of the investigator, could adversely affect the subject’s participation in the trial, or affect the trial integrity (e.g., amputation, severe mental disorder, immunocompromised, autoimmune disease)

**Recruitment.** Subjects will be recruited from the UIC GI lab just prior to scheduled colonoscopy using procedures previously approved by the UIC IRB for the Chicago Colorectal Cancer Consortium (UIC IRB #2010-0168). Subjects recruited at UIC will be from individuals who have had at least one prior colonoscopy with previously documented APs (within 5 years) undergoing a surveillance colonoscopy (Groups 1 and 3) OR from individuals who are undergoing their first colonoscopy at UIC (risk not known so could be classified to group 1, 2, 3, OR 4). Research staff will review UIC GI lab schedules and patient medical records to pre-screen potentially eligible subjects. Subjects deemed eligible during pre-screening will be called or emailed prior to their procedure or approached by research staff on the day of the colonoscopy prior to sedation. On the phone, via email or in person, the study will be described and if interested, and eligibility confirmed, written informed consent will be obtained by a research team member. Persons interested in the study after receiving the email will be asked to call the study staff for additional study details and eligibility screening. Subjects recruited at RUMC will be asked to sign a UIC consent form for the phone-based diet recall conducted by UIC staff and the DXA scan at UIC. Recruitment will be such that numbers of men and women and mean age and body mass index will be comparable between AA and NHWs subjects. Groups will be frequency matched for gender and age (± 5 years).

**The visit schedule and procedures and methods for Aim 1 are as follows:**

**Day of scheduled colonoscopy at UIC:** After the subject has reviewed the informed consents, asked questions, and signed it, he/she will be administered a questionnaire by research staff which will ask questions about age, race, gender, previous colonoscopies, and recent behaviors related to diet, physical activity, tobacco use and alcohol consumption, and medications and supplements the subject is taking over the counter. A blood sample (60 mL of blood or approximately 4 tablespoons) will be obtained by research staff. Two sterile oral swabs will be brushed against the inside of the subject’s cheeks to obtain buccal cells. The subject will be asked to spit into a cup to collect saliva (5 ml). He/she will collect a clean catch mid-stream urine
sample into a sterile cup (varied amount).

Then, the physician (UIC endoscopy clinicians on service) will begin performing the colonoscopy. The physician will obtain research biopsies of healthy colonic mucosa during the procedure. During the scope, 10 double bites of healthy colonic mucosa, 5 left side of colon and 5 right side of colon will be obtained using standard biopsy forceps during the clinically indicated procedure. For each pass of the biopsy forceps, 2 small biopsies (double bite) will be obtained (each about the size of a 1/3 of a rice kernel) and samples will be passed by the physician to the research staff for immediate processing. Obtaining the research biopsies is expected to prolong the colonoscopy procedure by less than five minutes. If the bowel preparation is poor, the physician will not be able to see a full view of their colon. In this case, we will not collect biopsies and the subject will be withdrawn from the study.

Procedures for risk stratification.

For Aim 1, we will know ahead of time the persons at elevated risk based on our screening of the GI lab schedule and patient medical records. However, it is possible that during the procedure, a growth (i.e., polyp, cancer) may be found and biopsied or removed and sent for pathology. If we find that the mass is cancerous following pathological assessment, the consented person will no longer be eligible. The person will be notified of their ineligibility by phone and collected samples and surveys will be discarded.

For Aim 1, it is also possible that during the endoscopic procedure a person identified as a control is found to have a growth (i.e., polyp or cancer). If the patient has a hyperplastic polyp or is considered low risk based on our risk stratification criteria, the patient will remain a control participant. However, we may find that following pathological examination of the polyp, the patient is now considered high risk based on our classification scheme. If this is the case, the patient will be stratified to our high-risk group. If we find that the mass is cancerous following pathological assessment, the consented person will no longer be eligible. The person will be notified of their ineligibility by phone and collected samples and surveys will be discarded.

For Aim 1, it is also possible that toward the end of study (i.e., Years 4/5), that we reach the participant number for patients at elevated risk but we are still seeking controls. We may consent an individual for the control group and find during the endoscopic procedure that the person has multiple polyps that would likely place him/her into the elevated risk category. If this case, the physician would make the decision to not take research biopsies. The patient would be made aware following the procedure and all biospecimens and the survey collected prior to the procedure would be discarded.

Lastly, it is possible that controls who acknowledge having a colonoscopy at another institution and agree to allow Dr. Jung or Dr. Yazici to request their colonoscopy report, have details in their report that would indicate that they belong to the elevated risk group. If this happens, these individuals would be reclassified as elevated risk.

If healthy colonic mucosa research biopsy samples are obtained, prior to being discharged, the subject will be given a food amounts booklet that will help him/her during the phone-based diet recall portion of the study. He/she will also be given a stool collection kit (including an ice pack, zip lock bags, and instructions) to collect stool samples, which he/she will bring to the follow up visit scheduled approximately 1 month later. Data will be extracted from the GI procedure report and pathology summary following the endoscopic procedure and data from previous colonoscopy reports in the EHR. Note: if the subject reports a previous colonoscopy at another institution, a request will be made by Dr. Barbara Jung or Dr. Cemal Yazici to obtain these records.

RUMC subjects: Subjects recruited and consented for Aim 1 at RUMC will also be consented for UIC-based procedures during the initial visit. This includes a phone call at 21 days post-colonoscopy to conduct a 24 hour diet recall and provide reminders regarding stool collection and follow-up research appointment (see detail of
Research Protocol V 8.0 11/1/2017

Phone call (21 days after colonoscopy): The UIC study staff will call ALL subjects (UIC and RUMC) approximately three weeks after the procedure to collect information about what he/she has eaten in the last 24 hours (also called a detailed 24-hour diet recall). The subject will be asked to use the food amounts booklet given at the time of the colonoscopy to aid the diet interview. The subject will also be reminded how and when to collect the stool sample and to confirm the in-person follow up visit date and time.

Research follow-up visit at UIC 25 to 35 days after colonoscopy. ALL subjects (UIC and RUMC) will return to UIC about a month after the initial clinically indicated colonoscopy for a research follow up visit. This visit will last approximately 180 minutes (see ancillary study for timed required for additional assessments). If the subject was required to take antibiotics between Visit 1 and Visit 2, we will have to reschedule their Visit 2, two months from the last day of their antibiotic dose. We ask that they notify the study staff, if the subject has taken antibiotics between Visit 1 and Visit 2, so we can reschedule their Visit 2. During this visit, he/she will drop off the stool collected (varied amount) in the previously provided cooler and collection kits. Then, the study staff will collect a non-fasting blood sample (50 mL of blood or approximately 3 ½ tablespoons). 20 microliters of blood obtained from the tubing tail of the BD needle collection kit will be used to test point of care CRP. Two sterile oral swabs will be touched onto the inside of the subject’s cheeks to obtain buccal cells. The subject will be asked to spit into a cup to collect his/her saliva (5 ml). The subject will provide a urine sample (varied amount). Subject’s height and weight will be measured. We will then collect information about subject’s habitual (long term exposure) dietary intake using a two Food Frequency Questionnaires (FFQ) and repeat the 24-hour diet recall (short term exposure) during this visit. The subject will also be asked about recent behaviors related to diet, physical activity, tobacco use and alcohol consumption, recent gastrointestinal symptoms, his/her and his/her family’s medical history, medication and supplement use, his/her usual physical activity and his/her mood, anxiety, childhood traumatic events, and stress levels. We will review any clinical information such as medication list, previous diagnoses and family history. Then, research staff will assess the amount of fat, bone, and muscle tissue the subject has using a whole body DXA scan. We will also assess the abdominal fat through an abdominal ultrasound machine, which will take about 5-10 minutes. The ultrasound machine does not use any radiation.

Sample processing and storage procedures. For all stool samples, we will follow the existing human microbiome project collection procedures. Blood samples will be processed for serum. Buccal cells, urine, and saliva will be kept frozen until analysis. All bio-specimens for Aim 1 will be stored at RUMC using the data bank system, FreezerWorks. A data dictionary will be built, shared between the three institutions, and PIs will be given access to the electronic central data repository.

Methodology/analyses. Mucosal abundance of bacterial genes associated with sulfur and bile acid metabolism will be measured by qPCR of 16S rRNA and functional genes with biopsy DNA. Bile salt hydrolase and bile acid 7a-dehydroxylating activities will be measured in stool samples. The proportion of glycine versus taurine conjugated bile salts, ratio of conjugated: unconjugated bile acids and concentrations of secondary bile acids will be measured in serum. Degree of baseline colonic inflammation will be determined by measuring mRNA expression of TNF-a, IL-11, IL-6, and Cox-2 using our published protocols. Bacterial, bile acid and colonic inflammation assays will be conducted in the Gaskins and Ridlon laboratories at UIUC.

PCR amplification of 16S rRNA & functional gene targets. qPCR will be performed with a 7900HT Fast Real-Time PCR System (Applied Biosystems) using DNA isolated from biopsies. Functional gene targets for taurine utilization, taurine:pyruvate aminotransferase (tpa), B. wadsworthia-specific (dsrA-Bw) and another dsrA described by Leloup et al. that detects all sulfate-reducing bacteria (SRB). Primers targeting 16S rRNA genes will be used to measure the abundance of B. wadsworthia and four SRB genera Desulfovibrio, Desulfobulbus, Desulfobacter, and Desulfotomaculum. Clostridium scindens bile acid 7a-dehydratase (baiE) gene will be quantified by forward primer 5'-TGTATTCCATAGCCCAGG-3' and reverse primer

Research Protocol V 8.0 11/1/2017

Page 13 of 49
5'TAGCCGTAGTCTCG- CTGTC-3'. All primer sets have been previously validated in-house, and DNA isolation and qPCR protocols will follow that described in our publications.

**Bile acid analyses.** Measurement of serum bile acids by electrospray-ionization mass spectrometry between diets will indicate extent of taurine-conjugation of bile acids, ratio of conjugated:unconjugated bile acids as well as the levels of secondary bile acids absorbed from the gut. Fecal bile acid analysis by HPLC will focus on dietary-induced changes in bile acids in fecal water (soluble bile acids interacting with epithelium) and total bile acids (insoluble bile acids in fecal pellet + fecal water). Correlations will be made between levels of *B. wadsworthia* in biopsies and stool and taurine-conjugated bile acids that promote their growth. Markers of bile metabolism in stool, bile salt hydrolase (BSH) activity, bile acid 7α-dehydroxylation activity (producers of tumor-promoting secondary bile acids) will be determined in stool samples. Relative and absolute levels of bile acid 7α-dehydroxylating bacteria, including *C. scindens*, will be determined by [24-14C] cholic acid assay and molecular approaches targeting bile acid-inducible genes, respectively. Markers of inflammation induced by secondary bile acids will be determined in colonic biopsies by quantification of mRNA expression of TNF-α, IL-11, IL-6, and Cox-2 using our published protocols.

Buccal cells, saliva, and urine will be also used for bacterial analysis and urine for metabolite analysis will be performed at RUMC and UIUC.

**Sample size.** We address power for separate comparisons of the Group 1 (AA with APs) to each of the other three groups (AA controls; NHW APs; NHW controls) with samples of n=50 in each group. We will have 80% power to detect effect sizes of 0.6 or greater for each comparison. Observed effect sizes in our preliminary data comparing means of log-dsrA, range from 0.75 to 1.3. We will have 96% power or greater to detect the same effect sizes. Power is calculated for two-tailed t-tests at α=.05 using software PASS version 11.0.8.

**Data analysis.** Primary analysis will be similar to our preliminary data (Fig. 2) and will look at differences in log B *wadsworthia dsrA* abundance between the 4 study groups using t-tests for log values of dsrA. Secondary analyses will examine: 1. Log levels of bacterial gene targets associated with sulfur and bile metabolism; 2. Correlations between polyp size and polyp number for index and past colonoscopies with log B. *wadsworthia dsrA* and other targets; 3. Differences in serum DCA, ratio of primary:secondary and unconjugated:conjugated bile acids, and taurine:glycine conjugated bile acid levels between the 4 groups and their correlations to the bacterial gene targets, colonic inflammatory markers and clinical variables such as polyp size and number; 4. Differences in bile salt hydrolase and bile acid 7α-dehydroxylation activities in stool between the 4 groups, and their correlations to the bacterial gene targets, colonic inflammatory markers and clinical variables such as polyp size and number; 5. Logistic regression with Group 1 versus others as the dependent variable will be used to analyze to analyze multiple predictors; 6. To look for associations between bacterial gene targets and metabolites, target-metabolite networks, principal component analyses, and canonical correspondence analyses will be used as previously performed by members of our group; and 7. Multivariate Analysis by Linear Models (MaAsLin), (which is a multivariate statistical framework that finds associations between clinical metadata and microbial community abundance or function and is provided as open source software that allows for deconfounding the effects of confounding metadata such as diet, age, sample origin, etc, adjusting with the false discovery rate method) will be used to look for associations for continuous clinical metadata (for example age and weight), Boolean metadata (sex, stool/biopsy), or discrete/factor metadata (cohort groupings and phenotypes).

**Figure 3:** Metabolism of bile salt, taurocholic acid, by gut microbiome. Gut microbes in the ileum and colon deconjugate bile salts to free bile acids by bile salt hydrolase. Taurine, because of sulfite moiety, can provide pathobionts in the gut with a terminal electron acceptor, allowing for their
growth and expansion in the gut. High-fat diet is associated with increased taurine-conjugation in humans [10]. Free primary bile acids are further metabolized to toxic secondary bile acids that can accumulate in the bile acid pool in humans and alter host physiology.

**Expected results:** Similar to our initial finding that dsrA-Bw distinguishes AA and NHW CRC cases from controls; we expect to observe a similar difference between AA and NHW with elevated CRC risk and their respective controls. Though not investigated previously, racial differences in taurocholate metabolism provides a plausible explanation for our observation. Thus, we expect to find significantly higher serum DCA levels, higher bile acid taurine:glycine conjugation ratios, higher total fecal bile acids and DCA levels, and higher fecal scindens and bile acid metabolism rate in AA relative to NHW and between patients with APs vs. controls in both groups, with the observed effect being most predominant in AAs with high risk for CRC.

**Potential pitfalls and alternatives:** Given that all endpoints have been previously validated and communicated in our publications, we do not anticipate any issues with execution of the biochemical and molecular biology assays. There are, however, several issues to consider regarding the subject population and likelihood of detectable differences among the study groups. For example, it is highly probable that some of the subjects in our elevated risks groups (those with history of APs) will not have any APs on the second colonoscopy (index colonoscopy that we are performing at the time that they are recruited). As stated above, we will conduct analyses based on the results of both the past colonoscopy and the index colonoscopy. We expect to see an increased abundance of our primary target, *B. wadsworthia* dsrA, regardless of the presence of current polyps, primarily because our original findings demonstrated differences in *B. wadsworthia* dsrA from healthy mucosa of AA-CRC cases, suggesting the microbial risk exists regardless of the tumor itself. However, this may not be true in polyp cases that we are recruiting. To analyze for this possibility, we will conduct a preliminary evaluation of our primary target after recruitment of the first 20 subjects in the AA-elevated CRC risk group (Group 1). Within Group 1, we will compare subjects with APs at the index colonoscopy with those without. If we note a significant trend or difference, we will recruit AP subjects at their first screening colonoscopy, which will assure that all subjects at the index colonoscopy will have 3 or more APs or an AP >1 cm for Groups 1 and 3. Lastly, although we do not anticipate problems with recruiting subjects at time of endoscopy (this was done for CCCC both at UIC and at RUMC), we can recruit subjects at our clinics prior to colonoscopy should problems arise. Last year, RUMC performed 11449 endoscopic procedures with an estimated 70% of these being screening colonoscopies and 13 attending gastroenterologists saw over 11000 patients in clinic (with 30% being AA). At UIC, we perform over 5000 colonoscopies a year with on average 30% in high-risk individuals and sees 12000 patients in clinic by 9 attending gastroenterologists. Our patient population is 40% AA, and another 30% NHW. We have a large referral stream for high-risk individuals for colonoscopies, and recruitment as proposed is entirely feasible.

**Aim 2:** Determine in the context of a diet-intervention trial the role of taurocholic acid metabolism by gut bacteria in CRC risk. Two isocaloric diets, an animal-based High in Taurine and SATurated fat (HT-HSAT) and a plant-based, Low Taurine and Low SATurated fat (LT-LSAT) will be used to determine the extent to which the relationship between diet (independent variable) and mucosal markers of CRC risk including epithelial proliferation, oxidative stress, DNA damage, and primary and secondary bile acid pools and biomarkers of inflammation (dependent variables) is explained by the abundance of *B. wadsworthia* and genotoxic H2S concentrations and C. scindens and concentrations of its metabolic end product the tumor-promoter DCA (mediator variables).

**Rationale and Hypotheses:** This aim will test the hypothesis that TCA is a key diet-controlled metabolite whose metabolism by *B. wadsworthia* and *C. scindens* yield a carcinogen and a tumor-promoter, respectively as illustrated in Fig. 3. Through a collaboration between PI Gaskins and PI Mutlu, we have observed significant positive correlations between high saturated fat intake and increased abundance of *B. wadsworthia* in a recent pilot study using DNA from sigmoid biopsies and FFQ collected from 9 healthy subjects over 4 time points (0, 3, 6, and 9 months). These latter data demonstrate that variation of saturated fat intake can influence levels of this bacterium, supporting our hypothesis and the selection of our diets. We and others have also observed

Research Protocol V 8.0 11/1/2017

Page 15 of 49
that feeding diets high in saturated fat to mice results in greater abundance of intestinal SRB associated with compromised barrier function and local inflammation. Together these observations provide rationale for testing the hypothesis that an animal-based diet selects for *B. wadsworthia*, which generates H2S through its metabolism of taurine.

We examined the genotoxicity using alkaline single-cell gel electrophoresis (SCGE or COMET assay) in nontransformed human intestinal epithelial FHs 74 Int cells. A marked and dose-dependent genotoxic effect was observed at concentrations as low as 250 μM (similar to that found in human colon) indicating H2S generates single-stranded DNA breaks (Fig. 4A). Cell death was not observed at 24 h of treatment (Fig. 4A, top). Sulfide damages directly, in absence of cellular metabolism, at concentrations as low as 1 μM in naked nuclei from CHO cells; nuclei exposed to H2S concentrations greater than 50 μM were completely degraded (Fig. 4B). Sulfide-induced DNA damage of naked CHO nuclei was quenched by co-treatment with the free-radical scavenger butyl-hydroxyanisole (BHA) (Fig. 4C), indicating genotoxicity is mediated by free radicals. Incubation of H2S-treated naked nuclei with the bacterial enzyme [fapy]-DNA glycosylase, a DNA repair enzyme that specifically recognizes and de-purinates 8-oxo-guanine residues in DNA,85 produced substantial increases in tail moments vs. nuclei not treated with enzyme (Fig. 4D). These data indicate that H2S treatment can produce oxidative DNA damage, specifically increased 8-oxo-guanine residues in DNA. We have also demonstrated that H2S modulates the expression of genes involved in cell-cycle progression, triggering both inflammatory and DNA repair responses in nontransformed rat16 and human17 intestinal epithelial cells. Together our observations constitute an extensive collection of data supporting the hypothesis that H2S production by mutualistic bacteria residing in the colonic mucosa serves as a diet responsive insult contributing to CRC risk.86 We will now examine in vivo the extent to which an animal- based diet rich in taurine and high in saturated fat induces biomarkers of DNA damage and repair in colonocytes through an increase of sulfidogenic bacteria and H2S.

![Figure 4](image-url)

**Figure 4:** Sulfide is a genotoxic agent. A) SCGE analysis demonstrates that sulfide produced genotoxic (bottom) but not cytotoxic (top) damage in nontransformed human intestinal epithelial FHs 74 Int cells. B) Analysis of the same with naked nuclei from CHO cells demonstrates that the genotoxic damage is generated directly without cellular metabolism. C) Sulfide-induced DNA damage is quenched by the hydroxyl scavenger BHA. D) Sulfide induces an increase in base lesions recognized by [Fapy]-DNA glycosylase. Experimental details can be found in Attene-Ramos et al.17-19

In humans, two primary bile acids are synthesized by the liver, the trihydroxy-bile acid CA, and the dihydroxy-bile acid, chenodeoxycholic acid (CDCA). Conversion of CDCA to the mono-hydroxy secondary bile acid lithocholic acid (LCA) by gut bacteria results in a water-insoluble bile acid unlikely to be found in abundance in fecal water. This is reflected in the fact that LCA does not accumulate in the biliary pool of humans, unlike DCA15, and that LCA is sulfated by the liver (and likely epithelial cells) detoxifying LCA and enhancing efficiency of excretion.87 Our hypothesis thus focuses on TCA due to the production of a water-soluble secondary bile acid, DCA, whose high levels in fecal water interact with colonocytes. Aim 2 will also test the hypothesis that due to its high saturated fat content, the HT-HSAT diet will increase the concentration of taurocholic acid, which will be metabolized to the tumor-
promoter DCA by C. scindens.

**Study design.** The proposed study will utilize a randomized, crossover, controlled feeding trial design composed of two experimental diets: (1) animal-based, high in taurine and saturated fat (HT-HSAT) and (2) plant-based, low in both taurine and saturated fat (LT-LSAT) (each consumed for 21 days).

**Subjects.** AA subjects between 50 - 70 years old at elevated risk for CRC, defined as a history of 3 or more APs or those who had an AP >1 cm on a prior colonoscopy at UIC/RUMC in addition to being obese (body mass index 30 – 50 kg/m^2) and circulating C-reactive protein (CRP) level > 3 mg/L (assessed during a research eligibility screening visit at RUMC as briefly described below) will be recruited for Aim 2. UIC AA subjects from Aim 1 that meet the elevated risk and BMI 30 – 50 kg/m^2, will be invited to participate in Aim 2 if they provided consent to future contact in AIM pertaining to new research opportunities. We will also actively recruit participants for this Aim in the UIC GI lab as described for Aim 1 (Groups 1 and 3).

**Inclusion criteria:**
1. Adult African American;
2. Obese (defined as BMI 30 – < 50 kg/m^2);
3. Age between 50 - 70 years old;
4. Patients with an increased risk for CRC, defined as 3 or more adenomatous polyps or adenomatous polyp > 1 cm within 5 yrs of enrollment;
5. An elevated C-reactive protein (CRP) (defined as > 3 mg/l)
6. Participants must be in good general health, not expecting major lifestyle changes in the next 6 months and willing to maintain their current activity level throughout the duration of the study.
7. Women only: Post-menopausal (natural or surgical) defined as no menstruation in the past 6 months.

**Exclusion criteria:**
1. BMI < 30 or > 50 kg/m^2 (for those interested and eligible, verify BMI by measuring weight and height, complete the screening consent form before assessing these measures)
2. Weight > 450 lbs. (max weight for the DXA scanner)
3. Race other than African American
4. Women only: at least one menstrual period in the past 6 months
5. Current malignancy except non-melanoma skin cancer that has been removed
6. Current gastrointestinal (GI) illness other than GERD or hemorrhoids (such as celiac disease, inflammatory bowel disease, malabsorptive bariatric procedures, etc.)
7. Chronic liver or kidney disease (elevated liver tests >3 times normal or creatinine above 2.0 mg/dl)
8. History of cardiac disease (such as admission for congestive heart failure within the past 5 years, or being on anticoagulants for heart disease, or having an ejection fraction <25%, etc.)
9. Positive genetic test for inherited polyposis syndromes (such as familial adenomatous polyps, hereditary non-polyposis colon cancer syndromes, etc)
10. Alcoholism or illicit drug use
11. Antibiotic use within the past 2 months
12. Regularly taking medications that may interfere with normal digestion (such as acarbose, cholestyramine, aspirin doses that exceed 81mg/day or 325 mg every other day)
13. Anticoagulant use or other factors that increase endoscopic risks
14. Non-English speaking
15. Pregnant or breast feeding
16. Dietary supplement use including pre- or probiotics within the past month
17. History of intestinal cancer, inflammatory bowel disease, celiac disease, or malabsorptive bariatric surgery
18. Inflammatory or connective tissue diseases (such as lupus, scleroderma, rheumatoid arthritis, etc.)
19. Diverticulitis
20. Prior perforation at colonoscopy or gastrointestinal bleeding due to biopsies of the colon
21. Therapeutic or vegetarian diet
22. Food allergy/aversion to any foods included in the trial
23. Claustrophobia
24. Any medical condition, which, in the opinion of the investigator, could adversely affect the subject’s participation in the trial, or affect the trial integrity

**Rationale for subject selection.** Participation has been limited to individuals with exaggerated risk for developing CRC (i.e., APs, obesity and systemic inflammation)\(^{88-91}\) to improve the power to detect the influence of the dietary interventions on our outcome measures. To reiterate, the rationale for focusing the diet intervention study on AAs comes from the recent observation mentioned above regarding a taurine respiring bacterium distinguishing AA but not NHW CRC patients from healthy controls and the previous work by PI Gaskins in AA subjects that focused on mechanisms underlying the increased risk for CRC associated with consumption of a Western type diet.

**Rationale for study design and diet choices.** The randomized crossover design will be used because it enables each participant to serve as his/her own control, which is optimal given inter-individual variation in the microbiota. The random assignment of diet order ensures any effect observed is due to the specific diet rather than treatment order. Because each participant acts as their own control for comparison it is imperative they enter each experimental diet interval in an identical state. To ensure subjects gut environment has returned to their baseline status, colonic sigmoidoscopies will be completed prior to and upon completion of each experimental diet. Rationale for selecting the two diets: a HT-HSAT diet has been chosen because: 1) animal protein, in particular red meat, is the most important single diet item shown to affect CRC risk,\(^{26}\) 2) animal protein is a particularly rich source of taurine, and 3) an animal-based diet is typically also high in saturated fat enabling an opportunity to examine how the simultaneous exposure of these nutrients influences primary outcomes. The LT-LSAT diet has been selected as the control diet because it is low in saturated fat and taurine, and contains no red meat, thus enabling efficient detection of the influence of high taurine with high saturated fat on the microbiota. The two diets are comparable for total fat and total fiber content. We aimed to hold total fiber constant across the two diets so that outcomes, mediated by animal protein and saturated fat, are not influenced by differential rates of colonic fermentation.

**Power and sample size.** Our power calculations are based on effect sizes generated from, \(B.\ wadsworthia\) and \(dsrA-Bw\) (bacterial gene markers for saturated fat exposure) in our previous feeding trial\(^7\) of a high animal protein diet and high saturated fat exposure (similar to the proposed HT-HSAT diet) in NAs that normally consume a diet very low in animal protein and saturated fat. The 14-day intervention resulted in a mean effect size for \(B.\ wadsworthia\) (16S rDNA), and \(dsrA-Bw\) of 1.11 and 0.63, respectively. **Our study is designed for at least 40 subjects to complete the trial.** On the basis of these results and a significance level of 5%, a sample of 40 subjects will provide >97% power to detect effect sizes at these levels. Assuming a 10% loss to follow-up, an initial recruitment sample of 44 will yield 40 participants that complete the trial. Moreover, the recent diet swamp study between NAs and AAs by PI Gaskins and colleagues substantiate that even 20 subjects provide sufficient power to detect diet effects for many of the Aim 2 endpoints (bacterial functional gene targets, Ki67, CD3, CD68, DCA and other secondary bile acids).

**The visit schedule and procedure and methods for Aim 2 are as follows:**

We will conduct a 10-week randomized controlled crossover feeding trial in 44 African American subjects with a history of adenomatous polyps, obesity, and elevated CRP as described above. The study visits will be conducted collaboratively at both RUMC and UIC. All persons found eligible during the screening eligibility visit at RUMC and willing to participate will be invited for the baseline study visit conducted at both UIC/RUMC. Subjects are asked to attend 5 total visits at RUMC and up to 8 visits at UIC. Details pertaining to the UIC related visits are detailed below.

**Visit 2/7 (Baseline before 1st or 2nd diet):** The subject has already been deemed eligible by RUMC staff and...
has provided RUMC written informed consent during a screening eligibility research visit. This visit will start at the UIC Integrative Physiology Lab/Westside Research Office Building, then the second half of the visit will happen at RUMC. The subject will be reminded in advance to fast (no eating or drinking liquids besides water) for 12 hours prior to this research visit by the RUMC team. At the first portion of the visit at UIC informed written consent for UIC procedures will be obtained. Then, measurement of Resting Energy Expenditure via indirect calorimetry will be completed. This will measure how many calories the subject's body needs at rest. We will use this data when we create the subject's isocaloric meal plans. For this assessment, the subject will rest for about 30 minutes and then wear a transparent hood that is connected to a metabolic monitor. The subject will need to lie still under the hood for about 30 minutes. Next, at UIC, the subject will have a Dual Energy X-ray absorptiometry (DXA) whole body composition scan which takes about 15 minutes. A DXA scanner is a machine that uses x-rays to measure body composition (percent of body fat, total body fat, and muscle mass and bone mass) that involves exposure to very low amounts of X-ray radiation. If the subject is female and pregnant, they cannot have their body composition tested using a DXA scan because it emits a small amount of radiation. For this reason, we will preselect post-menopausal women or women who have not had a menstrual period for at least six months to participate in this study. During this visit we will also ask the subject about the foods eaten in the last 24 hours, which is also called a 24 hour dietary recall. We will have the subject fill out a food frequency questionnaire (FFQ) that asks about the subject's dietary intake over the past year. The subjects will also be handed a food log that we will ask them to fill out at home over the 1 - 7 days (dependent on when Visit 3/8 is scheduled). The procedures at UIC will take up to 2 hours. After the visit at UIC, the subjects will go to RUMC to complete the remainder of assessments.

**Visit 3/8 (First food pick, Diet 1 or Diet 2 and non-fasting blood and stool collection):** This visit takes place at UIC within 7 days of Visit 2. This will be the first food pick up visit for Diet 1 OR 2. All food pickups will take place at the Applied Health Sciences Building located at 1919 W. Taylor Street, Chicago, IL, 60612 on the UIC campus. The subject will give the research staff their stool sample (varied amount) and will be asked about the method used to store the sample. The subject will also be asked to turn in their food log and we will review the log with the subject. The subject will be asked about medication use. If the subject has not provided a stool sample at this visit, we will not be able to give them the study foods and a new pick up date will be scheduled. If the subject has started antibiotics, we will also not be able to give them the study food and they will no longer qualify for the study. For persons returning a stool sample, 7 days of breakfast, lunch, dinner, beverages, snacks and condiments will be provided. The subject will be instructed on how to keep a daily food log and will be asked to return this log at the next pick up. For the first week, the subject will be asked to save any uneaten foods and beverages in the original packaging and bring this to the next food pick-up visit (visit 4/9). A non-fasting blood sample (50 mls or 3½ tablespoons) will also be collected to measure bile acids and other factors related to bacteria in the intestines and subject body weight will be measured. This visit will take approximately 30 minutes.

**Visit 4/9 (Second food pick up for Diet 1 or Diet 2):** This visit takes place at UIC one week after visit 3/8. This will be a food pick up visit. When the subject returns for the visit 4/8 food pickup, staff will collect any uneaten foods, measure body weight, and review the subject’s daily food logs including any deviations from the study provided food items. A new 7 day allotment of meals will be provided. The subject will receive new daily food logs and we will review how to keep a daily food log once again; the subjects will be asked to return this log at the visit 5/10 food pick-up. The subjects may be asked to continue to collect uneaten foods and beverages if they consumed < 75% of the food items or consumed non-study foods. This visit will take approximately 20 minutes.

**Visit 5/10 (Third food pick up for Diet 1 and 2):** This visit takes place at UIC one week after visit 4/9. This will be the last food pick up for a given diet (diet 1 or 2). When the subject returns for the visit 5 food pickup, staff will collect any uneaten food (if asked at visit 4 to being back uneaten foods), measure body weight, and review daily food logs including any deviations from the study provided food items. We will collect a non-fasting blood sample (50 mls or 3½ tablespoons) for the bile acid analysis. A new 7 day allotment of meals will be provided. The subject will then receive new daily food logs and we will review how to keep a daily food log.
once again and answer any questions; the subject will be asked to return this log at the visit 6/11 visit at RUMC. The subject may be asked to continue to collect uneaten foods and beverages. The research staff will provide the subject with stool collection supplies and instructions to take home. The subject will be asked to collect a stool sample and bring it to their next visit which will be visit 6/11 at RUMC. This visit will take approximately 30 minutes.

**Randomization/blinding.** Following baseline data collection, UIC Co-I Freels (biostatistician) will randomize subjects to their diet sequence (Diet 1 then 2 OR Diet 2 then 1) using a computer generated list of random numbers. She will be blinded to all aspects of the subjects’ profiles and will be given a subject ID only. Subjects will be blinded to their diet sequence; however they will likely realize when they receive a LT-LSAT vs. the HT-HSAT diet. UIC study personnel will not be blinded to the subjects’ diet sequence.

**Diet formulation and provision.** UIC PI Tussing-Humphreys will ensure timely preparation and distribution of the study meals. Both experimental diets will be formulated by the UIC PI (Tussing-Humphreys) and prepared, packaged and distributed by the diet coordinator in the UIC research kitchen. Both experimental diets are isocaloric and designed so that participants maintain their body weight within +/- 2.0% of the measures obtained during the baseline visit for each diet at RUMC/UIC (Visit 2/7). The subject will start on a diet within no more than 7 days of the baseline visit at UIC/RUMC (Visit 2 and 7). Seven days of breakfast, lunch, dinner, beverages, and condiments will be provided at a time. The experimental diets will be prepared at four calorie levels (1500, 2000, 2500 and 3000 calories). For individuals requiring an intermediate calorie level (e.g., 2100 calories), 100 calorie portions of foods with minimal protein and fat will be increased/decreased to reach the subject’s individual calories needs.

Water, non-caloric non-cola soda, and coffee will be provided consistently across the three diets. The taurine content is more than 17 times higher in the animal based diet (HT-HSAT) compared to the plant-based diet (LT-LSAT). Total protein is lower in the LT-LSAT diet compared to the HT-HSAT diet. The animal to vegetable protein ratio in the animal-based diet is approximately 4:1 and 1:3 in the plant-based diet. The HT-HSAT diet will have >15% of energy from saturated fat and the LT-LSAT diets will have < 8% of energy as saturated fat. While the type of fat varies, the overall percentage of energy from fat is approximately 36 - 39% in each of the diets. The dietary fiber content is similar across the two diet at 23 (HT-HSAT) or 27 (LT-LSAT) grams/day. Total carbohydrates are higher in the plant-based LT-LSAT diet compared to the animal-based diet. The relatively short exposure to safely prepared (UIC diet team has food sanitation training and certification) and nutritionally balanced diets are not expected to pose a risk. Careful regular monitoring will assure that any unanticipated adverse events, due to dietary changes, are immediately brought to the attention of the UIC/RUMC PIs.

**Measurement of daily calorie needs.** To ensure accuracy of the calorie prescription for each subject, REE will be measured using indirect calorimetry as previously described. The participant’s REE along with an activity factor consistent with the self-reported habitual physical activity obtained baseline (visit 2 at RUMC) will be used to calculate daily calorie needs. Participants are asked to maintain their habitual level of daily physical activity throughout the trial.

**Length of interventions and washout intervals.** The validity of a crossover feeding trial requires (a) subjects enter each dietary intervention without carryover effects from the preceding intervention (b) length of the intervention is adequate to induce the hypothesized change and (c) no other appreciable changes have occurred in the subject’s overall nutritional status. The 21-day interval for both the washout and dietary intervention was selected based on the fact that we observed changes in our previous study with a 14-day diet intervention. We have extended the interval to 21-days in the current trial to improve maximum response to the dietary intervention exposure given the budget and subject involvement (i.e. free living) perspective. To assess stability of nutritional status during the washout period, body weight and body composition via DXA (skeletal muscle and total regional fat mass) will be assessed at the beginning of each of the two diets.
**Dietary compliance.** Standard methods for assessment of dietary compliance throughout the trial include: 1) unannounced 24 h dietary recall; 2) weekly in-person body weight assessments conducted by the UIC diet coordinator when patients retrieve their meals; and 3) daily records of consumption of foods provided when on experimental diets and diaries of additional foods consumed not included in the diet. Dietary compliance will be closely monitored by PI Tussing-Humphreys who is experienced with running dietary intervention studies. Adherence counseling will be provided to all at the start of the trial and otherwise as needed.

**Sample processing and storage procedures.** All bio-specimens will be processed at RUMC. For all stool samples, we will follow the existing human microbiome project collection procedures. Blood samples will be processed for serum. Buccal cells, urine, and saliva will be kept frozen until analysis. All bio-specimens for Aim 2 will be stored at RUMC using the data bank system, FreezerWorks. A data dictionary will be built, shared between the three institutions, and PIs will be given access to the electronic central data repository.

**Experimental plan.** At baseline and post-diet (Fig. 5), we will measure the outcomes analyzed in Aim 1: (1) mucosal abundance of bacterial genes associated with sulfur and bile acid metabolism; 2) stool bile salt hydrolase and bile acid 7a-dehydroxylating activities; 3) measurement of serum bile acids to indicate the extent of taurine-conjugation of bile acids; 4) ratio of conjugated: unconjugated bile acids; 5) levels of secondary bile acids absorbed from the gut; and 6) expression of pro-inflammatory cytokines and Cox-2 as measures of colonic inflammation. In addition for Aim 2, we will measure in the colon biopsies concentrations of mucosal H2S and DCA and these will be related to colonocyte proliferation (Ki-67) and inflammation (CD3, CD68) by immunohistochemistry, and DNA damage. We will measure DNA damage by COMET fluorescence hybridization assay together with in situ staining for 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxodG), an oxidative DNA lesion, and the expression of its repair enzyme 8-oxoguanine DNA-glycosylase (OGG1; ab91421), XRCC1 [33-2-5] ab1838), which coordinates the action of DNA ligase III, polymerase beta, and poly-ADP-ribose polymerase in the BER pathway and the apurinic/apyrimidinic endonuclease Ape1 (ab2717), a multifunctional protein that protects cells from oxidative stress via its DNA repair, redox, and transcription regulatory activities. Proteins will be co localized to spots of oxidative DNA damage via immunofluorescence with co-localization for 8-oxodG with anti-8-oxodG 2Q2311 (ab64548). Also, bile acids in the fecal water (soluble bile acids) and total bile acids (fecal water + insoluble bile acids in fecal pellet) will be measured and correlated with levels of *Clostridium scindens* and with bile acid 7a-dehydroxylation [24-14C] CA activity assay. Buccal cells, saliva, and urine will be also used for bacterial analysis and urine for metabolite analysis. All assays will be performed in the Gaskins and Ridlon laboratories following our published protocols.

**Statistical analysis.** Our statistical model is shown in Fig. 6. The specific variables under study will be:  

**Independent variables:** diet composition, diet compliance. **Mediator variables:** A) Microbial genes associated with H2S and DCA production: *B. wadsworthia*-specific dissipatory sulfate reductase gene (*dsrA*), taurine:pyruvate aminotransferase (*tpa*); SRB dissipatory sulfate reductase gene (*dsrA*), *C. scindens* bile acid CoA-ligase (*baIB*) and bile acid 7a- dehydratase (*baiE* gene), which encodes the rate- limiting enzyme in DCA production; and 16S rRNA genes for *B. wadsworthia*, *C. scindens*, *C. hylemonae*, *C. sordellii*, and *C. hiranonas*; B) Mucosal H2S concentrations, and fecal and serum DCA. **Dependent variables:** A) Mucosal markers of CRC risk-number of positive cells in each biopsy tissue section: epithelial proliferation (Ki-67); Inflammation markers by immunohistochemistry (CD3, CD68) and qRT-PCR (TNF-a, IL-11, IL-6, and Cox-2). B) DNA damage (SCGE)-average median SCGE tail moment values (2 to 4 microgels/biopsy and at least 3 biopsies will be used for each subject at each treatment interval), 8- oxodG and BER protein localization; and C) Systemic inflammation marker CRP. Paired t-tests comparing each mean change score to zero will be used to determine how the dietary interventions impact dependent variables. The contrasts between the two diets (HT-HSAT vs. LT-LSAT) for the change in outcomes will also be tested with paired t-tests. Next we will explore how the dietary interventions influence the mediator variables with paired t-test for within and between the diets. Finally we will test for correlations between changes in mediators and changes in outcomes to determine if mediators account

---

**Research Protocol V 8.0 11/1/2017**

Page 21 of 49
for effects of dietary interventions. If distribution of change scores or contrasts between change scores shows strong departures from normality, the Wilcoxon-signed rank test will be used to test medians against zero, and sample ranks rather than original values will be used for linear regression modeling.

**Figure 6:** A mediation model will be used to determine the extent to which the relationship between diet [independent variable] and mucosal markers of CRC risk and DNA damage and repair [dependent variables] is explained by intestinal microbiota and their functions [mediator variables].

**Expected results:** Our previous study demonstrated that diet switch induced reciprocal changes in biomarkers of epithelial proliferation (Ki-67) and colonic inflammation (CD3 (lymphocytes) & CD68 (macrophages)), with reductions in AAs consuming a traditional African diet and increases in NAs consuming a red meat-based diet. We expect to observe similar differences for each subject upon comparison of colonic biopsies taken after consumption of the HT-HSAT versus the LT-LSAT diets. We predict that alteration of all dependent variables of DNA damage and mediator variables of microbial targets will be greater following consumption of the HT-HSAT diet than the LT-LSAT diet, and will demonstrate increased DNA damage mediated by sulfidogenic bacteria. We expect to find disparate microbiota and DNA damage profiles associated with each dietary intervention and higher numbers of sulfidogenic bacteria and DCA-producing bacteria associated with the HT-HSAT diet. We expect to find significantly higher taurine-conjugated bile acids in serum, and higher serum and fecal total and secondary bile acids (DCA) in the HT-SAT diet. We expect to find significant upregulation of COX-2 mRNA in biopsies during the HT-SAT diet which will correlate with DCA in serum and stool and H2S in mucosa.

**Limitations, pitfalls, alternative approaches:** Human nutrition intervention studies have many complexities that can be accomplished only with veteran investigators that have access to well-validated facilities and a proven track-record in this type of research. The diet team at UIC is highly qualified in this area; exemplified by numerous publications describing nutritional intervention studies similar to those proposed including a current study PI Tussing-Humphreys is conducting examining dietary iron exposures and intestinal inflammation. Our application is also distinctively advantaged by having two PIs (Mutlu and Gaskins) experienced in human studies of the microbiome. Thus we anticipate full success with implementation and interpretation of the diet intervention component of the two aims. PI Ridlon has extensive experience with bile acid chemistry and analysis and is an expert on DCA-producing bacteria; we therefore anticipate full success measuring bile acid profiles and fecal bile acid 7a-dehydroxyating bacteria. However, if we detect unknown bile acid metabolite(s), we will consult with past and current collaborators Dr. Takashi lida, Nihon University and Dr. Shigeo Ikegawa, Kinki University or Dr. Kenneth Setchell, University of Cincinnati. It is certainly plausible that the washout period may be insufficient for a return to the subject’s baseline, although this is unlikely based on our previous observation that a 14-day period was adequate. Accordingly, we will reassess after completion of the first 5 subjects’ first washout period, using our dependent variables and part of our mediator variables that look at sulfidogenic bacteria, DCA-producing bacteria, and serum bile acid profile. If subjects do not return to baseline, we will lengthen the washout period for the rest of the subjects in the study. We have decided to focus our efforts on measuring levels of bacterial genes, which control metabolism of TCA in the gut, rather than perform extensive microbiome sequencing. However, we have this option as we will collect and store patient stool DNA, and have a strong record of microbiome analysis.

Recent collaborative work between PI Mutlu and PI Gaskins demonstrates that the functional gene markers are detectable in sigmoid biopsies. Although, the assays to measure effects of DNA damage are standard and we have prior experience, we have recruited two experts for assistance with assay implementation and data interpretation. Dr. Nathan Ellis,
Program Leader of Cancer Biology and Genetics at the University of Arizona Cancer Center, has many years of experience working at the intersection of DNA repair and colorectal cancer susceptibility and has committed his expertise and resources to our study. Dr. Vilhem Bohr, one of the pioneers in the study of the biochemistry and cell and molecular biology of oxidative DNA damage and the BER DNA repair pathway has also offered enthusiastic support for our project and will provide advice as needed on the measurement of 8-oxodG. In the event that the collective data do not verify that BER is the primary pathway responsible for repair of the oxidative lesions induced by H2S, we will use biopsy RNA to examine other relevant DNA repair pathways measure expression of 84 key genes encoding the enzymes involved in base-excision, nucleotide excision, mismatch, double-strand break, and other repair processes using the Human DNA Repair RT² Profiler™ PCR Array (PAHS-042Z; SABiosciences).

**Timetable:** Study infrastructure will be established in months 1-6 [IRB approval; train personnel in SOP (consent and biospecimen collection, de-identification, storage); prepare subject materials (recruitment, screening tools, study folders); set up data mgmt.]. Aim 1 will be conducted between month 6 and YR 5. Aim 2 will start in YR 2 & conclude at the end of YR4. Data analyses and publications will be completed in YR 5.
Literature Cited

17. Attene-Ramos MS, Nava GM, Mueller MG, Wagner ED, Plewa MJ, Gaskins HR. DNA damage and toxicogenomic analyses of hydrogen sulfide in human intestinal...


51. Pai R, Tarnawski AS, Tran T. Deoxycholic acid activates 1-catenin signaling


68. Deplancke B, Hristova KR, Oakley HA, McCracken VJ, Aminov R, Mackie RI,


81. Leloup J. et al. Diversity and abundance of sulfate-reducing microorganisms in the


90. Huang XF, Chen JZ. Obesity, the PI3K/Akt signal pathway and colon cancer. *Obesity Rev* 10, 610-16 (2009).


97. McInnes P, Cutting M. in Core Microbiome Sampling, Protocol A, HMP Protocol # 07-
001 (2010).


Ancillary study: The influence of structural violence and individual behavior and health on the gut microbiome and colorectal cancer risk

Among all racial/ethnic groups, the African American (AA) population exhibits the highest CRC incidence and mortality, for reasons that remain poorly understood. Illinois has among the highest CRC incidence in the nation (53.9 vs. 47.7 per 100,000), with AAs residing in Cook County, Illinois, which includes the city of Chicago, suffering disproportionately (67.4 per 100,000). Both biological and socio-environmental exposures play a significant role in CRC, which develops through a multistep mutational process with many associated molecular events. Few studies have addressed both biological and socio-environmental factors, including exposure to structural violence that may contribute to disproportionately high CRC risk in AAs. The gut microbiome is emerging as a significant contributor to host health and disease. How social determinants interacting with individual factors influence the gut microbiome may be key to understanding racial/ethnic variation in CRC. AAs are more likely to live in urban poverty areas that are fraught with structural violence. Structural violence refers to the multiple ways in which social, economic, and political systems expose particular populations to risks and vulnerabilities including disproportionate exposure to poverty, crime, violence, and limited access to healthcare and healthy foods, leading to increased morbidity and mortality. Chronic exposure to these societal pressures can elicit adverse neuroendocrine [elevated glucocorticoids (GCs)] and immunologic (inflammation) responses that alter the gut microbiome. We reported that urban AAs harbor a higher abundance of sulfidogenic bacteria (SRB) in their colon compared to urban Non-Hispanic whites (NHWs), and race was the strongest predictor although diet (fat, protein) contributed modestly. This is an important finding given that SRB generate hydrogen sulfide (H2S) which is genotoxic and can activate pro-inflammatory pathways (Cox-2) and genomic instability—two hallmarks of CRC. We hypothesize that exposure to structural violence in the context of a high saturated fat and animal protein diet may explain the difference in SRB abundance in the AAs.

We suggest this given that in a mouse model, chronic social stress was associated with elevated GCs and a change in the gut microbiota, particularly an increase in the abundance of a genera that includes SRB, and a lower abundance of a genera containing bacteria that protect from intestinal inflammation; an effect that may be further exacerbated by diet. Furthermore, GCs are the most potent endogenous inhibitors of Cox-2. Colonic expression of 11β-hydroxysteroid dehydrogenase type II (11βHSD2) inactivates GCs, and inhibition of 11βHSD2 is protective against CRC through downregulation of Cox-2. The metabolites formed by gut bacterial 21-dehydroxylation of host GCs are potent inhibitors of host 11βHSD2. Therefore, we also suggest that AAs secrete higher GCs (due to exposure to structural violence), but lack the gut bacterial genes capable of producing protective 11βHSD2-inhibitors. Our overarching hypothesis is that exposure to structural violence increases psychosocial and physical vulnerability (e.g., anxiety/stress), compounded by one’s behavior (e.g., diet), which interacts with the gut microbiome in ways that result in CRC health inequality in urban AAs. Because the gut microbiome can be reshaped by diet and other factors, gaining an understanding of the complex interaction of social determinants, behavior, and biology on the gut microbiome, has promise to lead to novel strategies to reduce racial disparities in CRC.

Aim. Determine if exposure to structural violence increases psychosocial and physical vulnerability, and interacts with diet, to contribute to gut microbiome and microbial metabolite differences predictive of colonic inflammation and adenoma in urban AAs and NHWs. By leveraging our study of 200 Chicagoland adults at low and high risk for CRC (R01CA204808, study described above) we will assess: exposure to structural violence at the community and individual level; subject psychosocial and physical health, diet, GCs, cytokines, and healthy colonic mucosa Cox-2 expression; and from stool and mucosa, gut microbial
taxonomic structure (16s rRNA amplicon), abundance of targeted microbes and their functional genes (qPCR), shotgun metagenomics, and targeted stool microbial metabolites. We will ascertain differences between AAs and NHWs and relationships among the variables. We will parameterize community context, individual behavior/health, and microbial data to develop predictive models for colonic inflammation and adenoma and determine if different predictors exist for AAs and NHWs.

BACKGROUND AND SIGNIFICANCE.

**Epidemiology of colorectal cancer.** Cancer contributes significantly to healthcare spending in the US resulting in an estimated overall cost of approximately 201.5 billion dollars annually.\(^1\) Colorectal cancer (CRC) is the third most frequent cancer in both women and men, with approximately 135,000 new CRC cases, and 49,000 CRC related deaths in the US in 2016.\(^1\) Among all racial/ethnic groups, the African American (AA) population exhibits the highest CRC incidence and mortality for reasons that remain elusive.\(^2\) The state of Illinois has CRC incidence rates that exceed the national average (53.9 per 100,000 vs. 47.7 per 100,000) with AA Illinoisans disproportionately burdened.\(^3\) AAs residing in Cook County, Illinois, which includes Chicago, experience the state’s highest CRC incidence rates (67.4 per 100,000), which are higher than rates for AAs residing in suburban, small urban, and rural locations.\(^3\) Within the city limits of Chicago, estimated CRC prevalence, incidence and mortality is higher in AA neighborhoods compared to largely Non-Hispanic white (NHW) neighborhoods.\(^20\) Both biological and environmental exposures play significant roles in colorectal carcinogenesis, which develops through a multistep mutational process with many associated epigenetic changes.\(^4-6\) Although the incidence rate of CRC is higher in AA urban neighborhoods with a more aggressive manifestation of the disease, few studies have addressed socio-environmental risk by examining structural violence along with biological factors that may contribute to disproportionately high CRC rates in AAs.

**Racial/ethnic variation in the gut microbiome and CRC inequality.** The human body is home to trillions of bacteria, archaea, fungi, and viruses collectively known as the human microbiome. These microorganisms have the ability to both positively and negatively influence host health.\(^7,21\) The colon contains the majority of these microorganisms and under normal conditions, these microbes are commensal, aiding in digestion and the maintenance of the immune system.\(^17,22\) However, significant shifts to the gut microbial community structure and metabolic function have been implicated in CRC.\(^19,23\) Published work by our team indicates that racial differences exist in the type of bacteria residing in the colon.\(^11\) We compared the abundance of sulfidogenic bacteria (SRB) in healthy and non-involved colonic mucosa from 329 AA (100 controls, 97 cases) and NHW (76 controls, 56 cases) subjects undergoing colonoscopy in Chicago. Intriguingly, a comparison between races revealed that the abundance of SRB across a range of species (measured as pan-dsra), to be 10 times higher in AAs compared with NHWs (p<0.001), irrespective of disease status while controlling for age, sex, biopsy site, body mass index (BMI), education, income level and NSAID use (Figure 1). This finding is notable given that SRB produce genotoxic H\(_2\)S. Others have also reported variation in the gut microbial composition,

---

Figure 1. Differences in mean gene copy of pan-dsra comparing AAs vs. NHWs with and without CRC
predicted metabolic function, and microbial metabolite production [e.g., short chain fatty acids (SCFAs)] in AAs compared to other racial/ethnic groups. Co-I Gaskins has also described that AAs have higher levels of fecal microbial genes linked to the production of carcinogenic secondary bile acids [deoxycholic acid (DCA)]; higher DCA in stool; and lower levels of fecal butyrate, a microbiobly-derived SCFA beneficial to colonic health. Together, these studies suggest a plausible link between the gut microbiome and CRC inequalities in AAs. However, an explanation for these observed differences is not clear. To some extent nutrients associated with a high meat, high saturated fat diet may mediate the observed difference. Another factor may be host inflammation given that H2S producing bacteria thrive in an inflammatory milieu. We found that race was by far the strongest predictor of differences in the abundance of SRB in the AAs compared to the NHWs. However, this finding tells us little about the cause of microbiome diversity in our cohort. Fortenberry recently suggested going beyond the traditional racial and ethnic categories, to understand the social, behavioral, and health context of individuals that may link the microbiome to health inequality.

**Structural violence and health inequality.** Structural violence is a term used to describe the social, economic, and political structures that produce systemic inequality, discrimination, and exclusion of certain groups of people. Beyond individual lifestyle behavior and inherent biology, such upstream social determinants may contribute significantly to racial/ethnic health inequalities. Compared to other racial/ethnic groups, AAs are more likely to live in urban poverty areas that are disproportionately affected by structural violence including unequal exposure to segregation, poverty, violence, and crime and reduced access to healthy foods and healthcare. Specific to Chicago, highly segregated, predominantly black communities, such as Austin, Englewood, and Garfield Park, have long been affected by poverty, violence, and mass incarceration. and by the 2016 upsurge of shootings. The pathogenicity of these structural factors and their contribution to disease, and in particular CRC, remains poorly understood although recent studies have begun to document that area-level deprivation may affect CRC outcomes. We do know however, that chronic exposure to these types of societal pressures can elicit neuroendocrine and immunologic responses that can adversely impact health. Specifically, chronic exposure to structural violence can dysregulate the hypothalamic pituitary adrenal (HPA) axis. Chronic stress-induced dysregulation of the HPA axis is linked to altered glucocorticoid (GC) production (i.e., corticosterone/cortisol) and a propagated pro-inflammatory immune response that combined promote the development of chronic diseases including CRC.

**The gut microbiome and CRC health inequality: is exposure to structural violence the link?** The social and physical environment along with a person’s diet, physiologic and psychosocial health, genetics, and other lifestyle habits likely converge to influence the gut microbiome. We hypothesize that exposure to structural violence increases psychosocial and physical vulnerability (e.g., anxiety/stress) and interacts with one’s behavior (e.g., diet) which then disrupts the gut microbiome, resulting in CRC health inequality in urban AAs. Specifically, chronic exposure to structural violence can promote a propagated physiologic stress response that alters production of HPA axis related GCs. This increase in GC production promotes the upregulation of immune-derived pro-inflammatory cytokines. These neuroendocrine and immune changes can alter the gut ecosystem causing shifts in the structure and metabolic function of the gut microbiome, communication that is bidirectional. There are distinct differences in the gut microbiome of animals exposed to chronic social stress compared to controls. Specifically in Burokas et al., a higher abundance of the bacterial genera *Desulfovibrio*, which includes many SRB spp., and a lower abundance of the genera *Akkermansia*, which includes bacteria that protect from intestinal inflammation, was observed in stressed animals. This shift in the microbiota may be further exacerbated in the backdrop of a high fat diet. Moreover, chronic exposure to social stress is linked to lower expression of microbial metabolic pathways associated with SCFA production, an important observation...
given that higher levels of SCFAs are linked to lower colonic inflammation and reduced risk for CRC.\textsuperscript{19} Whereas in both humans and animals, modulating the gut microbiome using dietary probiotics dampened the HPA axis stress response,\textsuperscript{59,60} and in animals, conferred resilience to chronic social stress.\textsuperscript{60} Together these studies suggest that exposure to chronic social stress, along with diet, may shift the composition of the gut microbiota and alter their metabolic activity and metabolite production that favors colonic inflammation and genomic instability that promotes CRC.

Another pathway linking structural violence to the gut microbiome and CRC disparities in urban AAs may be through microbial metabolism of host GCs. Within colonocytes, an NADP+-dependent dehydrogenase, 11β–hydroxysteroid dehydrogenase type II (11βHSD2), functions to regulate the ratio of cortisol, to cortisone, the product of 11βHSD2.\textsuperscript{17} In this way, 11βHSD2 acts as a “switch” that limits intracellular cortisol levels and can control the inflammatory response. Zhang et al.\textsuperscript{16,61} reported that overexpression of 11βHSD2 mRNA and protein in both human adenoma and intestinal adenomas in \textit{APC}\textsuperscript{min} mice was highly correlated with increased Cox-2 expression and activity. Inhibiting 11βHSD2 either by gene silencing or enzyme inhibition by oral glycyrrhetinic acid (GA) resulted in significant reduction in tumor Cox-2 activity, tumor growth, and metastasis through inhibition of Cox-2 signaling.\textsuperscript{16,61} Notably, the more potent inhibitors of 11βHSD2 are gut microbial products of host GCs including cortisol metabolite \textit{11β-hydroxytestosterone} (11β-OHT) and corticosterone/cortisol 1β-hydroxyprogesterone (11β-OHP) derivatives, termed glycyrrhetinic acid-like factors (GALFs).\textsuperscript{17,18} Indeed, the absence of 11β-OHP GALFs was suggested as a biomarker for CRC.\textsuperscript{62} Co-I Ridlon has recently discovered the functional microbial genes responsible for formation of GALFs.\textsuperscript{63} Thus, we can target gut microbial metabolites predicted to prevent CRC (GALFs) between AAs and NHWs and quantify the microbial genes and taxa responsible. Thus, we posit that AA CRC inequality is driven not only by an increased abundance/activity of mucosally-associated SRB (i.e., \textit{Fusobacterium}, \textit{Bilophila}, \textit{Desulfovibrio}, etc.) but also a defect in microbial GC metabolism specifically, lower abundance/activity of protective GALF-forming bacteria (strains of \textit{C. scindens} encoding steroid-17,20-desmolase,\textsuperscript{64} \textit{Butyricicoccus desmolans},\textsuperscript{63} \textit{C. cadaveris},\textsuperscript{63} \textit{Eggerthella lenta} strains encoding 21-dehydroxylase).\textsuperscript{65}

\textbf{APPROACH.}

We argue that exposure to structural violence increases psychosocial and physical vulnerability (e.g., anxiety/stress) and interacts with one’s behavior (e.g., diet) which then disrupts the gut microbiome, resulting in CRC health inequality in urban AAs. To test this hypothesis we will employ a translational research approach to explore concepts from the following research aim.

The ancillary aim is a clinical epidemiologic study of AA and NHW Chicagoland adults at high and low risk for CRC (defined below). In these subjects, we will measure: (1) the level of exposure to structural violence at both the neighborhood and individual level, sociodemographic characteristics and psychosocial conditions; (2) the stool and colonic mucosa microbiome; concentrations of specific microbial metabolites linked to CRC; (3) colonic mucosa inflammation; biomarkers linked to chronic stress (GC and pro-inflammatory cytokines); and (4) individual behaviors. We will ascertain: (1) if the distribution of these predictors and microbial markers differ between the AAs and NHWs; (2) if the level of exposure to structural violence is associated with the stool and colonic mucosa associated microbiome and stool metabolite production; and (3) if exposure to structural violence, microbial data, GCs/cytokines and individual health/diet are predictive of colonic inflammation and adenoma using machine learning approaches like Random Forest. We also plan to isolate bacteria from the clinical samples that are capable of producing GALFs. These isolates may serve as novel probiotics that will function to dampen inflammation through inhibition of 11βHSD2 that can be tested in future pre-clinical and human intervention trials studies. Data from this aim may also inform research within CHER conducted by ESI from the Investigator Development Core. By
simultaneously studying social/environmental context and individual behavior and health in a diverse sample of urban AAs and NHWs, this Aim will begin to unravel the complex and interrelated mechanisms responsible for persistent CRC disparities.

**Rationale and hypothesis.** This aim is driven by the unequivocal evidence that the AA population, and in particular AAs in large urban cities such as Chicago, exhibit among the highest CRC incidence and mortality rates in the US.¹ We reported that microbiologic mechanisms operative in the colon may to some extent explain CRC risk differences in urban AAs.¹¹ Specifically, we found that the abundance of SRB across a range of species (measured as pan-dsrA), to be 10 times higher in urban AAs compared with NHWs (p<0.001) (Figure 1), suggesting the potential for deleterious H₂S production in AAs. Factors contributing to the observed difference are unknown although race was the strongest predictor. AAs residing in Chicago live in urban neighborhoods that are highly segregated with vast social and economic barriers including an upsurge in gun violence. Chronic exposure to these types of societal stressors can elicit neuroendocrine and immunologic responses that alter the gut microbiome.⁹ Exposure to chronic stress in animals is associated with a shift in the taxonomic structure of the microbiota and altered microbial metabolism and metabolite production.¹³,⁵⁶,⁵⁷ We are also putting forth a novel hypothesis, supported by work led by co-I Ridlon, that gut bacterial GC metabolism may play an important role in CRC disparities among AAs. Previous work identified gut bacterial 21-dehydroxylase activity against host GCs as a biomarker, absent in CRC.⁶² We posit that gut bacterial GC metabolism may be protective through inhibition of colonic epithelial 11βHSD2 which results in downregulation of Cox-2 expression and activity. This existing evidence gives credence to our hypothesis that exposure to structural violence, in the setting of a diet high in saturated fat and animal protein, drives gut microbiome dysbiosis that favors colonic inflammation, genomic instability and colon carcinogenesis in urban AAs.

**R01CA204808, Diet Modulation of Bacterial Sulfur & Bile acid Metabolism and Colon Cancer Risk.** Our team (Gaskins, Mutlu, Ridlon, Tussing-Humphreys, and Jung) was awarded (8/2016) an NIH grant to examine colonic mucosal abundance of bacterial genes associated with sulfur and bile acid metabolism and colonic inflammation in 200 AAs and NHWs at high [i.e., history or incident adenomatous polyps (APs)] and low (i.e., no history of APs) risk for CRC. This study began enrolling subjects in February 2017. To date (5/5/2017), 15 subjects have been enrolled, 8 AA and 7 NHW, and six have completed the study. A map of Chicago detailing CRC mortality rates by city ward with the residential location of the study subjects superimposed is presented in Figure 3. The blue dots represent NHWs and the yellow dots AAs. Three NHW subjects live in the Chicago suburbs. The majority of AA subjects reside in city wards with high CRC mortality rates. For Aim 1, we will leverage the existing clinical research infrastructure and engage with the same participants accrued to our ongoing trial. We will also leverage the biospecimens and clinical metadata collected from the subjects that have completed the trial. It is important to note that the new research aims dovetail with the aims of existing trial and extend our current research in a meaningful way. Thus, capitalizing on our existing research infrastructure is highly cost-effective and reasonable.
Subjects. We are recruiting 200 subjects (100 AA and 100 NHW) between the ages of 50 - 70 years old from the GI endoscopy clinics of UIC and RUMC at the time of surveillance/screening colonoscopy. Four subject groups (n=50/group) are being targeted for a case control design: **Group 1 (n=50)** - AAs with elevated risk for CRC. These subjects are defined as having 3 or more APs or those with an AP >1 cm on previous (within 5 years) or incident colonoscopy (confirmed following pathology), a well-accepted definition of increased risk. **Group 2 (n=50)** - AA controls, defined as having no polyps or if polyp(s) present, NOT meeting the definition of elevated risk as defined for Groups 1 and 3 following pathology on incident colonoscopy. **Group 3 (n=50)** - NHWs at elevated risk for CRC (see Group 1 above). **Group 4 (n=50)** - NHW controls (see Group 2 above). Recruitment is such that the numbers of men and women, and mean age, are comparable between the AA and NHW subjects. Exclusions (detailed in the human subjects section) include: significant medical conditions (e.g., cancer; chronic gastrointestinal diseases; clinically significant heart, liver, kidney, or inflammatory diseases; and inherited polyposis syndromes); alcoholism; therapeutic or vegetarian diet; and antibiotic or dietary supplementation use including pre- or probiotics. The sample size for this proposed study is based on our ongoing trial for which we have 80% power to detect effect sizes of 0.6 or greater for each group comparison for SRB abundance and functional microbial genes related to H₂S production. Observed effect sizes in our preliminary data comparing means of log-pan dsrA (abundance of SRB across a range of species) range from 0.75 to 1.3. Power was calculated for two-tailed t-tests at α=.05 using software PASS version 11.0.8

Experimental plan. Subjects are recruited prior to colonoscopy using IRB approved procedures (UIC IRB #2016-0495; RUMC IRB #13102201). Subjects are individuals who have had at least one prior colonoscopy with previously documented APs (within 5 years) undergoing a surveillance colonoscopy (Groups 1 and 3) OR from individuals who are undergoing their first colonoscopy (risk status not known could be classified to group 1, 2, 3, OR 4). Research staff review clinic schedules and patient medical records to pre-screen potentially eligible subjects. On the phone or in person, the study is described and if interested, and eligibility confirmed, written informed consent is obtained on the day of the endoscopic procedure (T1). Following, a survey is administered (sociodemographics, recent lifestyle behaviors, and medication/supplement use) and a fasting venous blood sample is obtained. Healthy colonic mucosa pinch biopsies are taken from the right and left colon (defined as 10 cm distal to the cecum and 25 cm from the anal verge, respectively), and immediately snap frozen in liquid nitrogen, stabilized in RNAlater, or placed in culture medium. Prior to discharge, subjects are given a stool collection kit and instructed to collect a stool sample at home in about 4 weeks (+/-5 days) (this will allow the gut microbial community to reestablish following colonoscopy prep). Data is also extracted from the electronic health record GI procedure report and pathology summary following the endoscopic procedure along with previous colonoscopy reports. If the subject reports a previous colonoscopy at another institution, a request is made to obtain these records. If we find that a subject has a cancerous mass following pathological assessment, the consented person is no longer eligible and notified of their ineligibility. All subjects return to UIC about a month after the colonoscopy to return their stool sample and provide a non-fasting venous blood sample (T2). We also measure subject height and weight, whole body composition via DXA scan, habitual diet intake via Food Frequency Questionnaire (FFQ), self and family medical history, usual physical activity, depressive symptoms, anxiety, and recent and childhood traumatic events.

Structural violence. To characterize one’s exposure to structural violence, we will measure both individual and neighborhood level factors. A similar set of measures is being utilized across the three U54 projects. Co-I Kim’s work has been on neighborhood effects and health inequality. Dr. Kim will guide assessment of neighborhood characteristics, exposure to structural violence, and individual psychosocial condition. Individual level variables will be obtained during face-to-face interviews with study participants at T2 and for those subjects who
have already completed our existing trial, these measures will be completed during an additional in-person research visit (T2.2). To quantify neighborhood disadvantage and risk of structural violence, participant residential addresses will be geocoded, and neighborhood level variables will be appended. Neighborhood characteristics will be computed at the census tract level using various sources, including the US Census Bureau and Chicago Police Department. We also ask residential mobility history to take into account change in neighborhood context over time. Neighborhood level measures include: the Index of Concentrated Disadvantage; crime rate (per 100,000) specifically, serious/violent crime (e.g., murder or armed robbery) and less serious crime (e.g., prostitution or drug dealing); Social Capital Index which is a composite measure including; civic engagement, community diversity, and interaction potential; the Environmental Protection Agency’s National Air Toxics Assessment (NATA) Total Cancer Risk to estimate potential exposure to carcinogens; access to healthcare will be determined based on Medically Underserved Area (MUA) classification; and access to healthy food (or food deserts) will be measured utilizing locations of grocery stores. We geocoded all grocery stores in Chicago. We will calculate distance to a nearest grocery store, and density of grocery stores at the census tract level. To examine individual level of exposure to structural violence and its psychosocial effects, we will measure perceived stress; PTSD symptoms; exposure to violence; and experiences of discrimination. We will also collect data on some of the known mediators, including: perceived neighborhood quality, problems, and resources; fear of crime; community collective efficacy, and resilience.

Methods/analyses

Data collection for ancillary measures. For all subjects enrolled in Aim 1 of the parent study described above (UIC and RUMC subjects), 30 additional minutes will be required during the research follow up visit to complete the ancillary measures (up to 210 minutes). We will collect 3-4 strands of hair from the nape of the participant’s neck for the cortisol analysis (5 minutes). We also have participants complete additional surveys related to perceived stress; PTSD symptoms; exposure to violence; experiences of discrimination; perceived neighborhood quality, problems, and resources; fear of crime; community collective efficacy, and resilience (25 minutes). Note: For participants that have already completed Aim 1, they will be invited to a 45 – 60 minute visit to complete a consent addendum and the above procedures in addition to the ultrasound to better abdominal body fat distribution.

Stool and colonic mucosa DNA extraction, 16s rRNA amplicon sequencing, metagenomics and bioinformatics. Microbial genomic DNA will be extracted from stool and mucosa using the Qiagen DNeasy PowerSoil- DNA Isolation Kit and Zymo Research DNA Clean and Concentrator kit, and quantified using a Qubit fluorimeter as previously described. To assess the stool and mucosa microbiota taxonomic structure, we will conduct 16s rRNA sequencing of the V4 region as previously described. We will use the Illumina MiSeq platform (Illumina, San Diego, CA) to generate paired end reads with an average base pair length of 300nt. Initial processing of sequence data will follow the workflow proposed by the QIIME developers and Callahan, with modifications as needed based on changes in software and sequence technology. Briefly, raw sequence data are demultiplexed, evaluated for base quality, trimmed to remove primer sequences and low-quality base calls, and finally filtered to remove error-prone reads. Reads are further dereplicated to remove redundancy, and overlapping paired reads are then assembled into a single contiguous read. Assembled reads are then clustered into operational taxonomic units (OTUs), after which potential chimeric sequences from PCR amplification are detected and removed. OTU clustering strategies will be optimized for the specific amplicons being assessed, largely dependent on whether a sizable database exists for additional downstream analysis. For example, an open OTU clustering strategy may be used for clustering and classifying sequences derived from the V4 region of the 16s rRNA gene as suggested by QIIME documentation, but other targeted amplicons may be assessed.

Research Protocol V 8.0 11/1/2017
Page 37 of 49
using alternative approaches such as de novo OTU ‘picking’ to determine whether more information is captured.

We will also construct metagenomic libraries using the NexteraXT kit (Illumina, San Diego, CA) for AAs at highest CRC risk (n=10, high exposure to structural violence, AP, colonic inflammation), lower risk AAs (n=10, low exposure to structural violence, no AP or colonic inflammation), NHWs at high risk (n=10, AP, colonic inflammation) and NHWs at low risk (n=10, no AP or colonic inflammation), which will be barcoded and multiplex sequenced on the Hiseq 4000 platform (Illumina, San Diego, CA). Metagenomic data will be processed to initially filter potential host DNA sequence and remove low quality base calls and potential artificial sequences such as adapters. Raw shotgun data can be further analyzed by initially aligning unassembled raw reads to a comprehensive non-redundant database, such as the NCBI RefSeq or ‘nr’ non-redundant protein databases, currently possible with various accelerated alignment tools such as DIAMOND. Annotation from the protein hits can then be used for both taxonomic and functional classification, currently possible using the tool MEGAN though alternative methods are also available. Parallel strategies for analysis include de novo assembly of the reads using tools such as MEGAHIT, followed by taxonomic and functional classification of the assembled contigs, taxonomic binning strategies to find related fragments, and alignment of whole data to quantitatively assess abundance of taxa. Though these analyses may exclude species present in low abundance, they can also help identify and resolve overall genic and operon structure. Furthermore, targeted assembly approaches such as Xander, can also be employed to help recruit reads from low-abundance organisms and potentially resolve targeted genic regions of interest, such as dsrA gene from sulfur-metabolizers and GALF-forming genes. The amplicon and metagenomics sequencing and the associated bioinformatics will be carried out by the UIUC High-Throughput Sequencing and Genotyping Unit and High-Performance Computing in Biology (HPCBio) group (see letters of support).

**qPCR of SRB, Akkermansia, and functional microbial genes of H\(_2\)S and GALF production pathways.** Quantitative PCR will allow us to verify abundances in metagenomic sequences and determine the levels of key bacterial functional genes as well as microbial taxa. Sulfidogenic bacteria (from our ongoing trial), Akkermansia, and functional microbial genes associated with \(H_2S\) (from our ongoing trial) and GALF production pathways will be quantified in triplicate with a 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA) with support from co-Is Gaskins and Ridlon. Small subunit rRNA genes of SRBs including Desulfobacter spp., Desulfobulbus spp., Desulfitomaculum spp. and Desulfovibrio spp. will be targeted. In addition, functional genes from the \(H_2S\) production pathway, dsrA harbored by all SRB, B. wadsworthia-specific dsrA and Clostridium scindens bile acid 7α-dehydratase gene will be quantified using primer sets that have been previously validated in-house, and DNA isolation and qPCR protocols will follow that described in our publications. The primers and probes used to identify Akkermansia muciniphila are based on 16s rRNA gene sequences validated and published by Everard and colleagues, GALFs are generated by the \(\text{desABC}(\text{E})\text{D}\) gene cluster and bacterial 21-dehydroxylase (Elen_0351-Elen_0362) from strains of *E. lenta.*

**Stool microbial metabolites.** The UIC Metabolomics core will conduct these analyses (see letter of support). Targeted metabolomics analysis for stool SCFAs (i.e., acetic acid, propionic acid, butyric acid), \(H_2S\), the GALFs 11beta-hydroxyprogesterone and 11beta-hydroxytestosterone and their 3alpha, 5alpha-reduced derivatives, and DCA, will be performed using liquid chromatography-mass spectrometry (LC-MS/MS).

**Colonic mucosal inflammation.** Degree of colonic mucosa inflammation will be determined by measuring mRNA expression of Cox-2 in healthy colonic mucosa using our published protocols (from our ongoing trial).

**Glucocorticoids and exposure to chronic stress.** 100 μL of serum will be used to measure 17-hydroxycorticosterone using similar LC-MS/MS methods as described above for stool. We
will also determine exposure to chronic stress by measuring cortisol from a small hair sample (1 cm) using a commercial ELISA kit (Alpco Diagnostics, Salem, NH). Hair cortisol has been validated against 24-hour urine cortisol as a relevant biomarker of long-term cortisol exposure.

**Circulating cytokines.** Serum from each subject will be analyzed in triplicate with the Bio-Plex® Precision Pro™ (Bio-rad, Hercules, CA) human cytokine 10-plex immunoassay to detect IL-1β, IL-6, IFN-γ and TNF-α, IL-2, IL-4, IL-5, IL-10, IL-12 (p70) and IL-13.

**Isolating GALFs, H₂S and DCA forming bacteria from stool and mucosa.** (a) **Biopsies:** Mucosal biopsies taken during colonoscopy will be transferred immediately into 1 ml anaerobic Dulbecco’s PBS containing 0.1% TWEEN 80 and vortexed 3 min. 0.1 ml aliquots will be transferred into anaerobic Laked Blood with Kanamycin and Vancomycin (LVK) medium and Bacteroides Bile Esculin agar for isolation of *Fusobacterium* spp., ATCC Medium 1490 (1% w/v taurine) and BBE for *Bilophila* and *Pyramidobacter,* and Postgate Medium N for *Desulfovibrio* spp. and other sulfate-reducing bacteria and serially diluted (1/10). **MALDI-TOF-MS Identification:** MALDI-TOF analysis will be performed in the Veterinary Diagnostic Laboratory, UIUC, on a MALDI Biotyper (Bruker Daltonics, Inc.), by standard methods developed by Carol Maddox, Ph.D. (see letter of Support). Peptide fragmentation patterns will be compared against several well-curated databases, including: Andromas of Andromas SAS (Paris, France), Biotyper of Bruker Daltonics (Bremen, Germany) and Saramis of BioMérieux (Marcy l’Etoile, France). Those colonies that cannot be identified by MALDI-TOF-MS will be identified by sequencing the full-length 16s rRNA gene. (b) **Activity-assays for GALFs, H₂S and DCA-forming bacteria:** we will assay individual cultures, identified by MALDI-TOF-MS.

**Data analysis.** In general, we will employ a suite of statistical techniques, as well as network modeling and machine learning approaches, to examine differences in exposure to structural violence, stress response (GCs/cytokines), individual behavior and health, and the gut microbiome and metabolite production between urban AAs and NHWs. For the contextual factors and spatial distributions, we hypothesize that the likelihood of exposure to structural violence is higher in disadvantaged neighborhoods (concentrated disadvantage and social capital), and the effect of structural violence are mediated through physical symptoms (stress) depending on discrimination experience, while perceived neighborhood and individual characteristics (sense of community, collective efficacy, fear of crime, resilience) moderate the relationship. We will use STATA and ArcGIS 10.2, geographic information system (GIS) software. We will employ descriptive statistics and visual mapping. We will examine the overall spatial patterns of clustering of structural violence outcome measures, in relation to the distribution of AAs and NHWs. Then, similar to our recent publication, we will examine how structural violence outcome measures relate to GCs/cytokines, abundance of bacterial genera, species, and functional microbial genes related to the production of H₂S and GALFs (from qPCR data), and stool microbial metabolites including SCFAs, H₂S, GALFs, and DCA in AAs and NHWs. We will conduct multi-level logistic and linear regressions while adjusting for pertinent covariates including age, biopsy site (when modeling the mucosa microbial data), sex, BMI, education, incomes, and NSAID use.

For the amplicon and metagenomic sequencing, all processed data are imported into downstream R/Bioconductor tools; this allows almost seamless integration across packages focused on particular data analysis techniques in a highly optimized statistical analysis environment. In particular, packages such as phyloseq can be used for general data processing and filtering, including performing multiple ordinate analyses to assess whether particular experimental factors (structural violence measures, diet, GC/cytokines) may correlate with the overall distribution of taxa across samples (or specifically in AAs and NHWs) may help pinpoint which taxa in particular may play a significant role. For example, structural violence, microbial metabolites, GCs/cytokines, and diet patterns will be tested by principal coordinate analysis (PCoA), permutational multivariate analysis of variance (PERMANOVA) and
generalized linear models (glm) to determine which organisms’ abundance, community structure, and metabolic functions, differ between AAs and NHWs. More complex statistical analyses, including supervised learning approaches such as Partial Least Squares (PLS) or Random Forest analysis, are implemented in the caret package. These approaches can be used to develop methods for prediction of possible phenotypes (colonic inflammation or adenoma) or factors for a sample with a given microbiome metabolite composition. We will run predictions overall and in AAs and NHWs separately. The phyloseq package also has direct integration with metagenomeseq, a separate toolset developed to focus on whole metagenome data analyses. This can be used to relate metagenome data such as inferred functional information, perform basic pathway analyses, and develop a more complete taxonomic and functional landscape of the microbiome.

Across all analyses, correction for false discovery rate and multiple testing comparisons will be employed as appropriate.

**Expected results and dissemination of findings.** We will present for the first time an evaluation of the influence of social determinants and individual behavior and health on the gut microbiome in Chicagoland AAs and NHWs at high and low risk for CRC. Based on our existing work, we expect to see differences in the gut microbiome of AAs and NHWs. Specifically from the 16s rRNA amplicon sequencing, we expect to observe an increased relative abundance of the genera *Desulfovibrio* and *Akkermansia* in mucosa and stool in AAs compared to NHWs. We also expect that based on the qPCR approach, we will see a higher abundance of the SRB and functional microbial genes associated with H$_2$S production in addition to a lower abundance of the functional genes associated with GALF production in mucosa and stool in the AAs compared to the NHWs. From the targeted metabolomic analyses, we also anticipate that AAs will have significantly higher H$_2$S and DCA and lower SCFAs and the GALFs 11beta-hydroxyprogesterone and 11beta-hydroxysteroid dehydrogenase and their 3alpha, 5alpha-reduced derivatives in their stool compared to NHWs. Overall, we expect that these differences will be explained by differential exposure to structural violence, GCs/cytokines, and diet. We also suggest that perceived neighborhood characteristics (sense of community, collective efficacy, and fear of crime) and self-report resilience will moderate the relationship between structural violence and gut microbiome dysbiosis. From the metagenomic data, we expect that the AAs with the greatest exposure to structural violence, AP and colonic inflammation will have a gut microbial community with altered metabolic function that is linked to inflammation. We also anticipate that the advanced machine learning algorithms will identify social, microbial, GCs/cytokines, and dietary parameters that are predictive of colonic inflammation and adenoma in AAs and NHWs. Lastly we anticipate that we will successfully culture unique bacterial strains that produce GALFs that can be cultivated and trialed in future interventions through pilot research by ESI. Taken together, we anticipate that the results will present a set of social, microbial, behavioral and physiological parameters that together will provide a deeper understanding of CRC health inequalities in AAs. Furthermore, if we see that differences in the gut microbiome are linked with exposure to structural violence, and augmented by diet, these microbes and their metabolic pathways could be targeted in future diet and probiotic (developed from the strains we will cultivate)-focused prevention efforts to decrease CRC risks disparities in urban AAs. The outcomes from this inquiry will also be disseminated through the CED to community stakeholders that can inform funding and interventions to address food access in segregated community areas in Chicago. We will co-host a CRC and diet community conversation held on the Southside of Chicago at the ACCESS Center for Discovery, operated by key community stakeholder ACCESS Community Health.

**Potential pitfalls and alternatives.** Collectively, our multidisciplinary team has extensive experience in all aspects of this project. We do not anticipate issues with any of the biochemical or molecular biology methods given that many of the assays have been communicated in our research protocol.
publications. Furthermore, we are collaborating with several well-established core research facilities on the UIC and UIUC campuses for the amplicon sequencing, metagenomics, bioinformatics, microbial metabolomics, and GC analyses. There are, however, several issues to consider regarding the subject population and likelihood of detectable differences among the study groups. For example, there is the possibility that no microbial differences will be observed with the current sample size, as the sample size is based on our ongoing trial. However, given our existing work, at a minimum we expect to see difference in SRB abundance between AAs and NHWs. It is possible that the genera *Akkermansia* will not differ based on exposure to structural violence, therefore, we will also examine abundance of the genera *Alistipes*, *Oscillibacter*, and *Lachnospiraceae* which have been linked to chronic exposure to psychosocial stressors including depression.\(^{139,140}\) For the microbial isolate analyses, it is possible that we will observe an organism of interest in a metagenome, and will not immediately culture the organism. We will utilize KEGG and SEED pathways from the genome in order to determine ways to modify the culture medium to increase our odds of successful isolation. It is likely that some bacterial isolates will produce a novel fragmentation pattern by MALDI-TOF and not be identified in the database. We will determine the identity of the organism by additional biochemical tests and 16S rRNA sequencing. Another consideration is at the time of funding, we expect that 50 subjects will have completed our ongoing trial. To date, all the subjects have consented to future contact for new research opportunities. They have also agreed to allow other studies to utilize their biospecimens and clinical metadata without additional consent. However, because Aim 1 will require additional data collection from the completed subjects, we contacted the six completers by phone to demonstrate the feasibility of this approach. We successfully contacted 100% of the subjects by phone. We pride ourselves on building a strong rapport with our subjects through numerous phone and in-person contacts. Lastly, we do not anticipate problems with recruiting subjects given our current progress the current R01 trial (R01CA204808). Last year, RUMC performed 11449 endoscopic procedures with an estimated 70% of these being screening colonoscopies and 13 attending gastroenterologists saw over 11000 patients in clinic (with 30% being AA). At UIC, we perform over 5,000 colonoscopies a year with on average 30% in high-risk individuals and sees 12,000 patients in clinic by 9 attending gastroenterologists. Our patient population is 40% AA, and another 30% NHW. We have a large referral stream (including the community partner ACCESS Health) for high-risk individuals for colonoscopies, and recruitment as proposed is entirely feasible. Although we are not opposed to recruiting additional subjects to fulfill the aims of this study if need be.

**Timeline.** The clinical research infrastructure has already been established at UIC and RUMC and main IRB approvals has been obtained. An IRB amendment will be required at both institutions for the additional subject assessments and for the return visit for the subjects that have already completed our ongoing trial. We anticipate that the amendment will be approved by month 2. All subject data will be collected between YRS 1 – 3. Culturing of bacterial isolates and MALDI-TOF analysis will occur YRS 1 – 5. qPCR, amplicon sequencing and metagenomic sequencing will be conducted YRS 3 and 4. Cytokine assays will be conducted in YR 4. Metabolomics of targeted microbial metabolites and GC will occur YRS 4 and 5. Bioinformatics and data analyses and publications will be completed in YRS 4 and 5.
REFERENCES FOR ANCILLARY AIM.

40. Doussard M, Peck J, Theodore N. After Deindustrialization: Uneven Growth and


47. Bryan M. *Gun Deaths In Chicago Reach Startling Number As Year Closes.* Chicago, IL: NPR;2016.


77. Principe DR, DeCant B, Staudacher J, et al. Loss of TGFbeta signaling promotes colon


94. Shobar RM, Velineni S, Keshavarzian A, et al. The Effects of Bowel Preparation on Microbiota-Related Metrics Differ in Health and in Inflammatory Bowel Disease and for the Mucosal and Luminal Microbiota Compartments. Clinical and translational


133. Bokkenheuser VD, Winter J, Dehazya P, Kelly WG. Isolation and characterization of


