

STUDY PROTOCOL

Open Access



# Design of the Building Research in CRC prevention (BRIDGE-CRC) trial: a 6-month, parallel group Mediterranean diet and weight loss randomized controlled lifestyle intervention targeting the bile acid-gut microbiome axis to reduce colorectal cancer risk among African American/Black adults with obesity

Andrew McLeod<sup>1</sup> , Patricia Wolf<sup>2</sup>, Robert S. Chapkin<sup>3</sup>, Laurie A. Davidson<sup>3</sup>, Ivan Ivanov<sup>3,4</sup>, Michael Berbaum<sup>1</sup>, Lauren R. Williams<sup>5</sup>, H. Rex Gaskins<sup>6,7,8,9,10,11</sup>, Jason Ridlon<sup>6,7,8,9</sup>, Jen Sanchez-Flack<sup>1,12,13</sup>, Lara Blumstein<sup>1</sup>, Linda Schiffer<sup>1</sup>, Alyshia Hamm<sup>14</sup>, Kate Cares<sup>14</sup>, Mirjana Antonic<sup>1</sup>, Beatriz Penalver Bernabe<sup>15</sup>, Marian Fitzgibbon<sup>1,12,13\*</sup> and Lisa Tussing-Humphreys<sup>1,13,14\*</sup>

## Abstract

**Background** Among all racial/ethnic groups, people who identify as African American/Blacks have the second highest colorectal cancer (CRC) incidence in the USA. This disparity may exist because African American/Blacks, compared to other racial/ethnic groups, have a higher prevalence of risk factors for CRC, including obesity, low fiber consumption, and higher intakes of fat and animal protein. One unexplored, underlying mechanism of this relationship is the bile acid-gut microbiome axis. High saturated fat, low fiber diets, and obesity lead to increases in tumor promoting secondary bile acids. Diets high in fiber, such as a Mediterranean diet, and intentional weight loss may reduce CRC risk by modulating the bile acid-gut microbiome axis. The purpose of this study is to test the impact of a Mediterranean diet alone, weight loss alone, or both, compared to typical diet controls on the bile acid-gut microbiome axis and CRC risk factors among African American/Blacks with obesity. Because weight loss or a Mediterranean diet alone can reduce CRC risk, we hypothesize that weight loss plus a Mediterranean diet will reduce CRC risk the most.

**Methods** This randomized controlled lifestyle intervention will randomize 192 African American/Blacks with obesity, aged 45–75 years to one of four arms: Mediterranean diet, weight loss, weight loss plus Mediterranean diet, or typical diet controls, for 6 months (48 per arm). Data will be collected at baseline, mid-study, and study end. Primary outcomes include total circulating and fecal bile acids, taurine-conjugated bile acids, and deoxycholic acid. Secondary

\*Correspondence:

Marian Fitzgibbon  
mlf@uic.edu

Lisa Tussing-Humphreys  
ltussing@uic.edu

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

outcomes include body weight, body composition, dietary change, physical activity, metabolic risk, circulating cytokines, gut microbial community structure and composition, fecal short-chain fatty acids, and expression levels of genes from exfoliated intestinal cells linked to carcinogenesis.

**Discussion** This study will be the first randomized controlled trial to examine the effects of a Mediterranean diet, weight loss, or both on bile acid metabolism, the gut microbiome, and intestinal epithelial genes associated with carcinogenesis. This approach to CRC risk reduction may be especially important among African American/Blacks given their higher risk factor profile and increased CRC incidence.

**Trial registration** ClinicalTrials.gov [NCT04753359](https://clinicaltrials.gov/ct2/show/NCT04753359). Registered on 15 February 2021.

**Keywords** Cancer health disparities, Colorectal cancer, Mediterranean diet, Weight loss, Gut microbiome, Bile acids, Nutrition

## Administrative information

<b>Funding</b>	National Cancer Institute (R01CA250390-01A1)
<b>Sponsor</b>	National Cancer Institute 9000 Rockville Pike Bethesda, MD 20892
<b>Role of Sponsor</b>	The study sponsor shall have no influence over the design, implementation, conduct, results or interpretations of the results of this trial.

The completed SPIRIT checklist is attached as Additional file 1.

## Introduction

Colorectal cancer (CRC) is the third most common cancer in the USA, with 151,030 new cases and 52,580 CRC-related deaths expected in 2022 [1]. Among all racial/ethnic groups, people who identify as African American/Blacks (AA/Bs) have the second highest CRC incidence and highest CRC mortality in the USA [1]. This disparity may exist in part because AA/Bs, compared to other racial/ethnic groups, have a higher prevalence of risk factors for CRC, such as higher body fatness [2, 3], lower dietary fiber consumption [3, 4], poorer diet quality [5, 6], and higher intake of saturated fats [7, 8] and animal proteins [3, 8, 9].

While there are several proposed mechanisms linking these modifiable risk factors with CRC, one potential mechanism is the bile acid (BA)-gut microbiome axis [10–12]. This is the two-way interaction between the gut microbiome and BAs that may explain the etiology of several diseases [13]. High levels of dietary fat and animal protein increase fecal and circulating BAs and promote a shift in amino acid conjugation of the BA pool from glycine to taurine [14, 15]. The gut bacteria *Bilophila wadsworthia* and *Clostridium scindens* can increase their abundance from, and metabolize taurine-conjugated BAs to increase fecal levels of the tumor-promoting/secondary bile acid deoxycholic acid (DCA) and to increase concentrations of hydrogen sulfide ( $H_2S$ ), a carcinogen that can trigger dysregulated

cell-cycle progression and DNA repair responses [12]. Importantly, *B. wadsworthia* is significantly more abundant in uninvolved mucosa of CRC cases versus controls among AA/Bs but not non-Hispanic Whites [7]. Furthermore,  $H_2S$  dampens production of butyrate [16], an anti-neoplastic microbial metabolite [17]. Hence, the gut microbiota may explain to some extent why higher intakes of fat and animal protein sometimes observed in AA/Bs compared to other race/ethnicities place them at increased risk of CRC.

The gut microbiota may also help explain how weight loss (WL), increased fiber, and decreased fat and animal protein consumption may decrease CRC risk in this subgroup. Aleman et al. demonstrated that a VLCD producing 10% WL reduced fecal BAs and DCA, and yielded a higher ratio of glycine-conjugated BAs, which correlated directly with gut microbial compositional changes [18]. Biemann et al. showed that a 5–15% WL from a lifestyle intervention significantly reduced circulating BA levels compared to a control group [19]. Lastly, in O’Keefe et al., increasing fiber consumption and decreasing saturated fat and animal protein in AA/Bs for 2 weeks increased fecal butyrate and reduced *B. wadsworthia* abundance, fecal DCA levels, and colonic histological biomarkers of CRC [8].

Based on these prior studies, we hypothesize that the largely plant-based and fiber-rich Mediterranean diet (MedDiet), which is associated with lower circulating DCA [20], higher abundance of fecal bacteria capable of producing butyrate [21], and decreased CRC risk in observational studies [22], will have robust positive effects on the BA-gut microbiome axis among persons with elevated levels of colonic bacteria capable of producing sulfide, including AA/Bs. Furthermore, because weight loss and a healthful diet may have an additive effect on disease risk reduction [23], we hypothesize that WL via caloric restriction combined with a MedDiet will induce the greatest weight loss, the greatest improvements in the microbiome-BA axis, and the greatest

reduction in intestinal gene expression associated with carcinogenesis.

To our knowledge, no studies have tested this hypothesis. To fill this gap, we will conduct a randomized controlled trial (RCT) enrolling AA/B, middle to older aged adults with obesity to examine the effects of a MedDiet alone (Med-A), WL through lifestyle intervention alone (WL-A), or a calorie-restricted MedDiet for WL (WL-Med) compared to typical diet controls (TDC), on BA metabolism, the gut microbiota, and intestinal epithelial markers associated with carcinogenesis.

## Methods

### Study design

The Building Research in CRC prevention (BRIDGE-CRC) trial is a parallel, four-arm, superiority RCT that will enroll 1922 adults aged 45–75 years old with obesity and who self-identify as AA/B. We will randomly assign 48 participants to each arm: Med-A, WL-A, WL-Med, and TDC. The interventions are 6 months in duration,

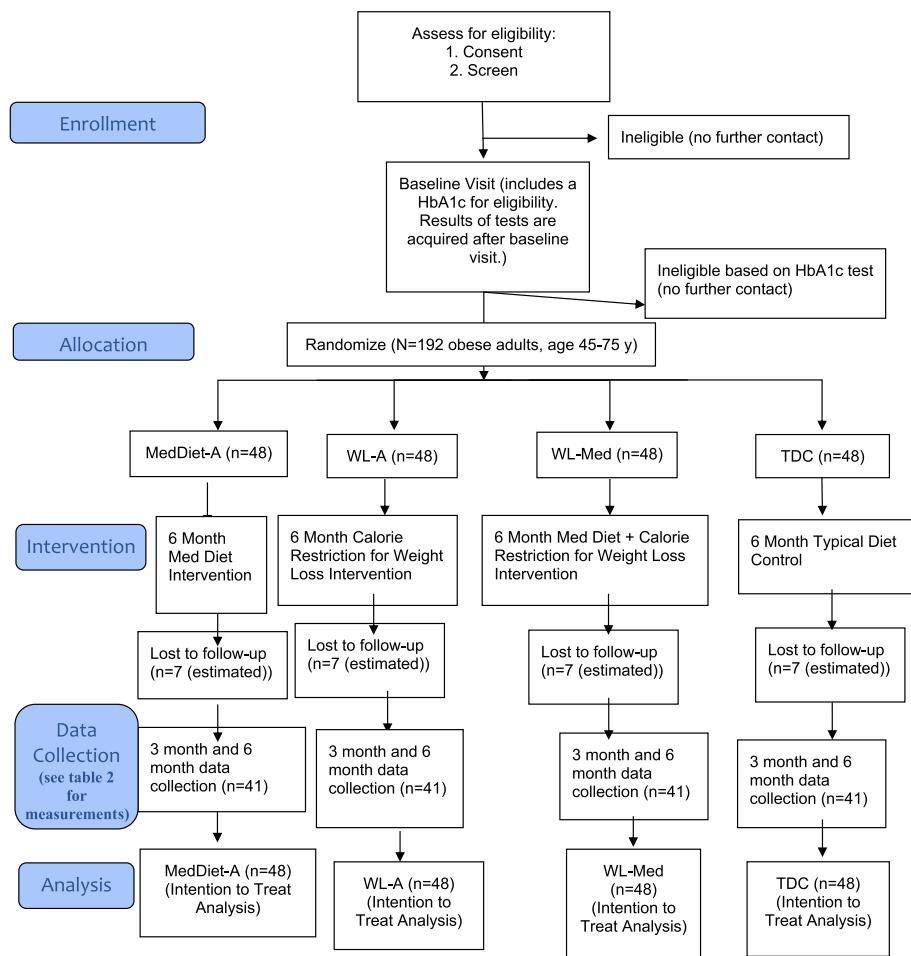
with assessment of study outcomes occurring at baseline, month 3, and study end. See SPIRIT figure (Fig. 1) for participant flow.

### Recruitment

Participants will be identified through electronic medical records (EMR) of the University of Illinois Health and Hospital System (UI Health). They will also be recruited through email listservs, social media, presentations at senior centers, and flyers at UI Health and health fairs. We will aim to recruit 30% male and 70% female participants given the underrepresentation of males within dietary and weight loss interventions [24]. We will recruit participants in three separate, approximately equal cohorts.

### Inclusion criteria

Potentially eligible participants will be screened in two phases—the first, via telephone and the second,



**Fig. 1** SPIRIT figure

in-person. Inclusion criteria to be assessed over these two phases include age (45–75 years), self-identify as AA/B, body mass index (BMI) 30–50 kg/m<sup>2</sup>, willing and able to participate in all procedures, willing to be randomized to 1 of the 4 study arms, and willing to provide informed consent. Participants must also understand English, have access to a smartphone, be familiar with digital meeting platforms, and plan to reside in the Chicago area for the duration of the intervention. Doctor's approval will be required for the following conditions: recent-onset chest pain, tightness, or pressure during physical activity; falling, feeling unsteady, or using an assistive device while standing or walking; other health concerns about starting an exercise program; or systolic blood pressure  $\geq 160$  mmHg or diastolic blood pressure  $\geq 100$  mmHg.

#### Exclusion criteria

Exclusion criteria include the following: hypertensive crisis (defined as a blood pressure  $> 180/115$  mmHg with symptoms such as chest pain and headache); renal disease; autoimmune disorders; immunodeficiency; malabsorptive disorders; significant gastrointestinal and hepatic diseases; severe ischemic heart disease; unable to exercise due to emphysema; chronic bronchitis or asthma; bariatric surgery (excluding lap bands not filled with saline); regularly drinks 3–4 alcoholic drinks per day or more; illicit drug use (other than marijuana based on self-report); combustible tobacco use; uncontrolled diabetes based on hemoglobin A1c (HbA1c)  $> 9.0\%$  (assessed after baseline data collection); eating disorder; any cancer treatment within the past 12 months; history of CRC; genetic predisposition to CRC (e.g., Lynch syndrome); difficulty walking; weight  $> 450$  lbs. (weight limitation of the DXA scanner); currently adhering to a MedDiet based on the 21 point Mediterranean Eating Pattern for Americans (MEPA)-III screener [25] (adherence defined as a score  $\geq 13$ ); self-reported WL of 10 lbs. in the past 3 months; preparing for gastric bypass surgery or other bariatric surgery; unable or unwilling to adopt a MedDiet; oral or IV antibiotic use in the 3 months prior to data collection; takes prebiotics, synbiotics, dietary fiber supplements, or laxatives  $\geq 3$  times per week; night-shift work, unwilling to stop use of prebiotics, probiotics, synbiotics, dietary fiber supplements, or laxatives from 4 weeks prior to baseline data collection up until the last data collection period; pregnancy; or active COVID-19 infection within 5 weeks prior to data collection.

#### Informed consent

We will obtain an IRB waiver of consent to identify potential research participants using the University of Illinois

Health and Hospital System EMR. Written, informed consent (see Additional file 2 for consent form) and HIPAA authorization will be obtained in-person or via video conferencing by research staff trained and certified to obtain informed consent. The written consent will be reviewed in detail and used as a guide for the person obtaining the consent. Participants will be asked if they consent or not to having their biospecimens archived for use in future, ancillary studies. Participants are encouraged to ask questions throughout the process. The informed consent procedure generally takes 15–30 min to complete. Participants are given a signed copy of the consent to bring home and informed again that they are free to withdraw at any time. Participants will be paid an honorarium for their participation in the full study. \$70 will be provided for baseline data collection, \$70 for month 3, and \$100 for final assessment plus \$240 either in food delivery service or as a financial payment for the TDC participants.

#### Randomization

The data manager, who has no contact with participants, will use the REDCap randomization module to randomly assign participants to (1) Med-A, (2) WL-A, (3) WL-Med, or (4) TDC, in a 1:1:1:1 allocation ratio. Stratified block randomization will be adopted to ensure good balance among the four study groups with respect to the selected covariates: sex, age (45–60 or 61–75 years), and baseline BMI (30–39.99 or 40–50 kg/m<sup>2</sup>). The randomization allocation will be created in SAS and imported into the REDCap randomization module.

#### Dissemination policy

Results will be made public via timely publications in the best journal possible and at academic conferences. Authorship will be defined per authorship definitions provided by the International Committee of Medical Journal Editors [26].

#### Interventions

##### Behavior change models

The interventions incorporate constructs of the Social Cognitive Theory [27] and Self-Determination Theory [28] that include self-efficacy, social support, motivation, and barrier reduction. Med-A, WL-A, and WL-Med use dietary monitoring, individualized goal setting, and problem solving to increase mastery and goal achievement in small incremental steps to enhance self-regulatory skills. For the MedDiet interventions (Med-A, WL-Med), learning to prepare MedDiet-friendly foods is a major aspect of dietary pattern adoption mastery and goal achievement. The intervention is implemented in 24 weekly one-on-one sessions and through a private Facebook group containing videos and posts. If a participant does not

want to participate in the social media group, the social media content is sent via email. An overview of the first 5 weeks of the intervention by study arm is included in Additional file 3: Table S1, Additional file 4: Table S2, and Additional file 5: Table S3.

#### Intervention instructors

The intervention instructors have experience conducting nutrition interventions and community-based nutrition education, with the lead interventionist having registered dietitian credentials. There is a separate interventionist for each study arm.

#### Individual sessions

Participants in Med-A, WL-A, and WL-Med will meet with their interventionist once weekly in-person or via Zoom (based on participant preference) for 24 sessions. Two of these sessions will be “floating” to allow for holidays that may interfere with the once-weekly plan. Weekly session content is manualized and includes time to address successes and barriers and problem solving using motivational interviewing techniques. Week one is planned as a 1-h session, with the remaining 23 weeks planned as 30-min sessions. For participants randomized to Med-A and WL-Med, the interventionist will instruct on and support the adoption of an eating pattern consistent with a MedDiet using an individualized MedDiet exchange list and companion guide [29, 30]. The exchange list will contain several categories of foods and beverages characteristic of a MedDiet (e.g., low-fat dairy). Participants will be advised to consume a specific quantity of food and beverage from each category, but will be free to choose which foods and beverages from each category they would like to consume.

#### Social media private groups

The WL-A, Med-A, and WL-Med interventionists will each lead their own private Facebook group. Each week, the interventionist will host a Facebook Live event providing an overview of the topics and content to be discussed in individual sessions and posted to the social media page. The live event will also be an opportunity for participants to ask questions. The Facebook Live content will be posted for participants who are unable to join in real-time. Weekly content will also include instructional videos (e.g., how to count MedDiet exchanges), existing web content pertinent to the study arm (e.g., YouTube cooking videos), and cooking demonstrations created by the study team and infographics. The number of weekly postings will be equivalent across the study arms. Participation at live events and views and completion analytics are available at the individual level and will be used to evaluate interaction with the social media content. Those

opting out of the social media group will be asked to confirm receipt of the weekly content via email.

#### Intervention calorie and nutrient recommendations

The Med-A and WL-Med arms will receive similar recommendations regarding macro- and micronutrient composition, with an emphasis on consuming fruits, vegetables, whole grains, plant proteins such as nuts and legumes, olive oil as the main dietary fat, and red wine. Fast foods, red and processed meats, fried foods, pastries, refined carbohydrates (e.g., white bread and white rice), sugary beverages, and high-fat dairy products will be discouraged. Recommended daily exchanges will be based on individual caloric needs to maintain weight (Med-A) or to achieve approximately 5% WL from baseline at study end (WL-Med). For participants randomized to WL-A, the focus will be on daily calorie restriction (−500 to −750 kcal/day) to achieve approximately 5% WL at study end. The content for the WL-A arm will be adapted from the Diabetes Prevention Program open-source intervention materials [31]. Energy needs for the WL-Med and WL-A arms will be determined using the Mifflin St. Jeor equation with an activity factor based on baseline physical activity level [32]. Balancing alcohol's effect on BA metabolism while respecting participants' dietary preferences, interventionists will recommend that alcohol comprise at most 1 drink per day for females and males and when possible consumed with meals. See Table 1 for more details regarding the intervention diet recommendations.

#### Food delivery

The Med-WL and Med-A groups will receive food boxes with ingredients to make the recipe for the dish prepared in the cooking skills video content for the corresponding week. The food will be sourced and delivered by Top Box Foods (Chicago, IL), a community-based non-profit with a mission to increase access to affordable and healthy foods. The deliveries will begin during week 2 of the intervention and continue on a biweekly basis until week 14, at which point deliveries will occur every fourth week, with a total of 9 deliveries throughout the 24-week study. Additionally, participants will be provided with a pantry box in the first session which will include shelf stable items to be used in recipes for later sessions.

#### Physical activity

The Med-A and control groups will be asked to maintain their baseline physical activity pattern, and the WL-A and WL-Med groups will be asked to progress toward achieving 10,000 steps daily. A fitness instructor will engage with participants through the FitBit platform to monitor step counts and to encourage and motivate

**Table 1** Intervention calorie and nutrient recommendations

Dietary component	Med-A	WL-Med	WL-A
<b>Calories</b>	No caloric deficit	500–750 kcal/day deficit	500–750 kcal/day deficit
<b>Total fat</b>	~35% of total calories	~35% of total calories	~ 30% of total calories
<b>Saturated fat</b>	~5% of total calories	~5% of total calories	< 10% of total calories
<b>Monounsaturated fat</b>	~20% of total calories	~20% of total calories	No specific recommendation
<b>Polyunsaturated fat</b>	~10% of total calories	~10% of total calories	No specific recommendation
<b>Carbohydrates</b>	~40% of total calories	~40% of total calories	No specific recommendation
<b>Protein</b>	~20% of total calories	~20% of total calories	No specific recommendation
<b>Alcohol</b>	≤1 drink per day females and males (if alcohol is usually consumed)	≤1 drink per day females and males (if alcohol is usually consumed)	No specific recommendation
<b>Sodium</b>	Limit sodium	Limit sodium	Limit sodium
<b>Fiber</b>	25–35 g/day	25–35 g/day	No specific recommendation

participants to increase daily steps. The fitness instructor will also create and post social media content (described more below) to provide strategies and motivation around increasing daily steps. The WL-A and WL-Med interventionists will also post challenges to the private Facebook groups to encourage increasing daily steps among their participants.

### Control (TDC) participants

A research team member will meet individually with the TDC participants at the start of the intervention. They will not receive information about making diet or physical activity changes and will be asked to maintain current diet and physical activity behaviors. Additionally, they will receive weekly health newsletters about non-diet related health topics (e.g., flu prevention). All WL-Med materials will be offered to the TDC participants in a self-guided format following the intervention.

### Adherence

Several methods will be used to assess intervention adherence. First, weekly attendance at individual sessions will be tracked. Participants missing two consecutive individual sessions will be called and text messaged by the interventionist to encourage attendance. Second, engagement with weekly social media content will be monitored using Facebook analytics which is reported on the individual level. Participants not engaging with the social media content for two consecutive weeks will be reminded to engage with the content. Participants receiving social media content via email who fail to confirm receipt for two consecutive weeks will be contacted and asked to confirm receipt of the materials. Third, body weight will be monitored weekly in the WL-A and WL-Med groups. Participants will be asked to weigh in

person or transmit a photo via secure text message during their weekly one-on-one meetings. Participants who lose < 1 lb. per week for 3 consecutive weeks will discuss with their interventionist barriers to adopting a calorie-restricted eating plan and problem solve accordingly. If necessary, the interventionist will revise the calorie prescription (WL-A and WL-Med) and the number of daily MedDiet exchanges (WL-Med). For the Med-A group, we will record the weight of participants monthly to confirm adherence. Lastly, adherence to the Med-Diet will be monitored during the first 6 weeks of the intervention among the Med-A and WL-Med participants. During weekly meetings with the interventionist, participants will complete the MEPA-III screener [25]. Those unable to increase adherence to a Med Diet will discuss with their interventionist barriers to adopting a MedDiet and problem solve accordingly.

### Intervention fidelity

To assure fidelity, intervention sessions will be semi-structured in a manualized format. Given the individualized nature of the sessions, the interventions cannot be fully scripted. A semi-structured approach will allow for use of a fidelity checklist itemizing the expected general content for sessions. The interventionists will be observed by senior staff once weekly, and the fidelity checklist will be completed. Based on the fidelity checklist, senior staff will meet with the interventionists to provide feedback and to review progress, concerns, and address deviations from the intended intervention content. This approach has worked well in our previous trial [33].

### Study outcomes and data collection procedures

#### Outcomes

The primary study outcomes are total circulating and fecal BAs, taurine-conjugated BAs, and DCA. Secondary

outcomes include body weight, body composition, dietary change (i.e., MedDiet adherence), physical activity (i.e., steps per day), metabolic risk, circulating cytokines (e.g., IL-6, TNF $\alpha$ ), gut microbial community structure and composition, fecal short chain fatty acids (SCFAs) (i.e., butyrate, acetate, propionate), and exfoliated intestinal cell gene expression levels. Data collection will take place at UIC and by telephone at baseline, month 3, and study end. All data collection staff will be trained and certified in informed consent, testing procedures, specimen processing, and data management. All procedures will be written and stored electronically. Major changes to study design will first be reviewed and approved by the IRB before implementation. The primary data collectors and lab technicians will be blinded to randomization group assignment by having the assignment hidden from view in their individual REDCap account.

#### Participant confidentiality

The confidentiality of participants will be safeguarded using unique participant identification numbers/codes. Participant biospecimens will be kept in a locked lab until analysis. The biospecimens will be labeled with ID, date, and specimen type only. All survey data will be directly entered into the Research Electronic Data Capture (REDCap) data management system using data structure developed by the study team [34]. Only the investigators and members of the research team will have access to participant data files, biospecimens, and REDCap data structure. Datasets used for analysis will not contain any direct identifiers.

#### Measurements

See Table 2 for an outline of measurements taken at each data collection time point.

#### Anthropometrics

Height will be measured, in duplicate, to the nearest 0.1 cm at the beginning of the study using a fixed stadiometer (Seca, Birmingham, UK). Weight will be measured, in duplicate, to the nearest 0.1 kg using a digital scale (Tanita, Arlington Heights, IL, USA). Body composition will be measured with dual energy X-ray absorptiometry (DXA) on a General Electric Lunar iDXA machine (GE Healthcare, Chicago, IL, USA).

#### Blood-based biomarkers

Participants, who have fasted for at least 12 h, will undergo a blood draw, and we will retain whole blood and derive serum, plasma, and buffy coat using centrifugation at 3000 RPM for 15 min at 4°C. Blood products

will be sent to Quest Diagnostics (Wood Dale, IL, USA) immediately or stored in a -80°C freezer for later analysis.

#### Inflammatory markers

Systemic inflammation will be operationalized with CRP and measured with an immunoturbidimetric assay at Quest Diagnostics (Wood Dale, IL, USA). Circulating cytokines (GM-CSF, IL-2, IL-4, IL-6, IL-8, IL-10, TNF $\alpha$ , and IFN $\gamma$ ) will be measured with the Bio-Plex human cytokine assay (Bio-Rad, Hercules, CA, USA) at the UIC Research Resources Center (Chicago, IL, USA).

#### Metabolic risk

Glucose, insulin, and HbA1c will be measured by Quest Diagnostics (Wood Dale, IL, USA) using spectrophotometry, an immunoassay, and an enzymatic assay, respectively. Homeostatic model assessment of insulin resistance (HOMA-IR) will be calculated according to a published formula: (fasting glucose [mmol/l]  $\times$  fasting insulin [ $\mu$ U/ml]) / 22.5 [35].

#### Blood pressure

Blood pressure will be measured with an electronic, automated blood pressure machine (Omron 3 Series, Kyoto, Japan). Participants will rest for a minimum of 5 min before having their blood pressure measured twice in a seated position. Participants will be instructed to rest for 1 min between the two blood pressure readings.

#### Dietary intake

Habitual diet will be assessed with the validated, online, VioScreen food frequency questionnaire (FFQ) [36]. Recent dietary habits will be collected and analyzed using the Automated Self-Administered 24-hour (ASA24) Dietary Assessment Tool, version (2020), developed by the National Cancer Institute, Bethesda, MD, USA [37]. The ASA24 will be completed twice at each data collection point. While both instruments are designed to be self-administered, they will instead be administered by trained and certified research staff to reduce recall bias. A MedDiet adherence score first created by Panagiotakos et al. [38] and later adapted to an American population by Tangney et al. [39] will be calculated from both the FFQ and ASA24 data. Lastly, the MEPA-III MedDiet screener [25], developed by Tangney et al., will be used not only as a screener but to quickly assess MedDiet adherence at each data collection timepoint for the Med-A and WL-Med arms. It consists of 21 items divided into 9 categories (e.g., fruits, vegetables, nuts and seeds) and assesses

**Table 2** BRIDGE-CRC trial measurements

Data type	Baseline	Month 3	Month 6
<i>Anthropometrics</i>			
Height	X		
Weight <sup>a</sup>	X	X	X
Body composition (DXA)	X		X
<i>Inflammatory markers</i>			
Systemic inflammation (hs-CRP)	X	X	X
Circulating cytokines (GM-CSF, IL-2, IL-4, IL-6, IL-8, IL-10, TNF $\alpha$ , and IFN $\gamma$ )	X	X	X
<i>Metabolic risk</i>			
Glucose, mg/dL	X	X	X
Insulin, $\mu$ U/mL	X	X	X
Hemoglobin A1c, %	X	X	X
HOMA-IR (calculated from glucose and insulin)	X	X	X
Blood pressure, mmHg	X		X
<i>Dietary intake</i>			
VioScreen food frequency questionnaire (FFQ)	X		X
24-h diet recall (two per timepoint) (ASA24)	X	X	X
MEPA-III MedDiet screener	X	X	X
Medication use	X	X	X
Physical activity (steps per day, FitBit)	X		X
Microbiota composition (16S rRNA)	X	X	X
Predicted gene abundance, metabolic pathways, and microbial network analysis	X	X	X
Quantification of the following genes via PCR: <i>dissimilatory sulfite reductase A</i> ( <i>pandsrA</i> ), <i>B. wadsworthia</i> -specific <i>dsrA</i> , 16S rRNA, <i>Desulfovibrio</i> spp.-specific 16S rRNA, BA 7 $\alpha$ -dehydratase from <i>C. scindens</i> , <i>butyryl-CoA transferase</i> gene ( <i>BcoA</i> )	X	X	X
Fecal short chain fatty acids (uM) (e.g., butyrate, acetate, and propionate)	X	X	X
Fecal and circulating bile acids (uM) (i.e., total BAs, total unconjugated BAs, taurine- and glycine-conjugated BAs, and CA, CDCA, LCA, and DCA)	X	X	X
Exfoliated intestinal cell gene expression analysis	X	X	X
<i>Surveys</i>			
Demographics	X		
Health history, including bowel health	X		X
Sleep quality	X		X
Food away from home	X		X
<i>Psychosocial variables</i>			
Discrimination	X		
Food security	X		X
Depressive symptoms	X		X
Stress	X		X
Recent trauma	X		X
Anxiety	X		X

**Abbreviations:** BA bile acid, CA cholic acid, CDCA chenodeoxycholic acid, DCA deoxycholic acid, DXA dual energy X-ray absorptiometry, GM-CSF granulocyte-macrophage colony-stimulating factor, HOMA-IR Homeostatic Model Assessment for Insulin Resistance, hs-CRP high-sensitivity C-reactive protein, IFN interferon, IL interleukin, LCA lithocholic acid, MedDiet Mediterranean diet, MEPA Mediterranean Eating Pattern for Americans, TNF tumor necrosis factor

<sup>a</sup>Weight will also be recorded as needed throughout the intervention to monitor compliance

frequency and quantity of intake of foods encouraged and discouraged by a MedDiet eating pattern.

#### Medication use

We will catalog prescription medications, over the counter drugs, and dietary supplements taken in the past 30 days.

#### Physical activity

Participants will wear a fitness tracker (FitBit, San Francisco, CA, USA) for 7 days. We will access steps, heart rate, and sleep data from the FitBit with a third-party application, FitaBase (San Diego, CA, USA).

### Stool collection

Stool will be collected by participants 0–48 h prior to each of the three data collection visits using a kit provided by study staff. They will be instructed to store collected stool in a commercial refrigerator and to transport the stool on an ice pack to each visit. The stool will be stored in a refrigerator during the study visit. After the visit, it will be processed immediately for exfoliomics analysis (see below), with the remaining stool being aliquoted, flash frozen in liquid nitrogen, and then stored at –80°C until further analysis.

### Fecal microbiota composition

Microbial DNA will be extracted with the Qiagen DNeasy PowerLyzer Powersoil Kit (Qiagen, Hilden, Germany) as done previously [7]. The UIC Genomics core will polymerase chain reaction (PCR) amplify genomic DNA with primers CS1\_515F and CS2\_806R (modified from the set used by the Earth Microbiome Project) targeting the V4 region of microbial small subunit ribosomal RNA (rRNA) genes. Amplicons will be generated using a two-stage PCR protocol, described previously [40, 41]. The V4 region of the 16S rRNA gene will be sequenced with the Illumina MiSeq platform to generate 2×250 bp paired end reads per sample [42]. Environmental controls will be included in the sequences to distinguish from any contaminants in reagents or the lab environment.

### Fecal microbiome analysis

16S rRNA reads will be processed with DADA2 [43]. After quality filtering and removing low abundance (<0.5%) reads, Silva v132 reference sequence database [44] will be used for taxonomical assignment of the Amplicon Sequence Variants (ASVs).

### Predicted gene abundance, metabolic pathways, and microbial network analysis

Employing the software PiCRUSt2 [45], we will predict gene abundance and most probable metabolic pathways using ASVs as input. The software SPIEC-EASI will be used to model microbial network configurations [46].

### Quantification of microbial gene targets

Quantitative PCR (qPCR) will be performed with a 7500Fast Real-Time PCR system (Life Technologies, Carlsbad, CA, USA) and validated primers, as described previously [7, 8, 47], using microbial DNA isolated as explained above. Functional gene targets for sulfide production will be measured including degenerate dissimilatory sulfite reductase A (pan-dsRA) and *B. wadsworthia*-specific dsRA (dsRA-Bw). Primers targeting 16S rRNA genes will be used to measure total bacterial abundance as well as *Desulfovibrio* spp. The BA

7α-dehydratase (*baiE*) gene from DCA producing *C. scindens*, and *butyryl-CoA transferase* gene (*BcoA*) will also be quantified. An internal standard will be used.

### Fecal SCFAs

Flash frozen stool will be transferred into plastic tubes; 2,2-dimethylbutyric acid will be added at 1 mmol/L as internal standard, vortexed and centrifuged, and supernatant filtered through a Millex-GS filter (Millipore, Burlington, MA, USA) and refiltered with a microconcentrator (Millipore, Burlington, MA, USA). The final filtrate will be analyzed using an Agilent Technologies 6890N Network GC System (Santa Clara, CA, USA) with a flame ionization detector for SCFAs. A mixed-SCFA standard solution will be prepared using high purity (0.99%) reagents (Sigma, St. Louis, MO, USA). Fecal SCFA concentrations will be computed by using a peak area ratio of the sample profile relative to the internal standard. Analysis will be conducted by the UIC Mass Spectrometry, Metabolomics and Proteomics Facility.

### Fecal and circulating bile acids

All major primary (e.g., CA and CDCA) and secondary fecal and circulating BAs (e.g., DCA and LCA) as well as their glycine and taurine conjugates will be assessed. Total BAs and total unconjugated BAs will also be measured. Samples will be extracted, and supernatants will be dried and resuspended for liquid chromatography/mass spectroscopy analysis following validated and published methods [48]. Authentic reference BAs will be purchased from Sigma-Aldrich (St. Louis, MO, USA) and Steraloids (Newport, RI, USA). Blind duplicate samples will be used to assess inter- and intra-batch variability. An internal standard will be used.

### Stool-based exfoliated intestinal epithelial cell analysis—exfoliomics

From stool preserved in Ambion Denaturation Solution, eukaryotic polyA+ RNA will be isolated using the Active Motif mTRAP Maxi kit followed by DNA removal with DNAFree (Invitrogen, Waltham, MA, USA) as previously described [49]. Illumina libraries will be created with 100 ng RNA using the Lexogen QuantSeq 3' mRNA-Seq Library Prep Kit for Illumina (FWD) (Lexogen, Vienna, Austria) and unique molecular identifiers. The libraries will be generated following the adjusted protocol from the manufacturer for low input RNA and quantified using the Qubit 1x dsDNA HS assay kit (ThermoFisher Scientific, Waltham, MA, USA) and a Bioanalyzer High Sensitivity DNA Analysis kit (Agilent, Santa Clara, CA) following manufacturer's instructions. Sequencing will be performed on an Illumina NovaSeq 6000 platform using standard Illumina protocols. RNA reads will be mapped

with the Bowtie2 [50] using optimized parameters to the Ensembl GRCh38 human reference. Reads will be examined for quality control using FastQC and quantified using HTSeq-count [51, 52]. Sequencing reads will be filtered to remove any gene with less than 2 reads in more than 33% of the samples [49].

### Surveys

All non-dietary surveys will be electronic and self-administered. The measures include demographics (e.g., sex, household income), health history including bowel health (Bowel Health Questionnaire from NHANES [53]), sleep quality (Pittsburgh Sleep Quality Index [54]), discrimination (Everyday Discrimination Scale [55]), food security (US Household Food Security Survey Module [56]), depressive symptoms (Patient Health Questionnaire-8 (PHQ-8) [57]), perceived stress (Perceived Stress Scale (PSS-10) [58]), recent trauma (Recent Traumatic Events Scale [59]), anxiety (Generalized Anxiety Disorder scale, GAD-7 [60]), and food (eaten) away from home [61].

## Statistical analysis plan

### Sample size

We calculated the sample size to detect group differences in three circulating and fecal BA outcomes (total BAs, taurine-conjugated BAs, and DCA) using results reported by Biemann et al. [19] and Aleman et al. [18] as a guide. In Biemann et al. [19], data were summarized as medians, interquartile ranges, and sample sizes; we estimated the corresponding means and standard deviations [62]. Estimated effect sizes for diet intervention on serum total BAs and taurine-conjugated BAs are 0.75 and 1.2 (assuming the correlation coefficient between pre and post measurements is 0.9 as commonly observed in repeated assays). Estimated effect size on fecal DCA is 0.67 using the raw data provided by Aleman (assuming change in the control group is 0) [18, 62]. For each dependent variable, the overall level of type I error is set at  $\alpha = 0.05$ ; that is, no adjustment is made for correlation among dependent variables and hence tests of these variables are not independent. The sets of variables examined in Hypotheses 1 and 2 should be interpreted as related. With 4 groups, there are 6 pairwise contrasts among cell means at a given occasion; we plan to test 3 of these using *t*-tests at the Bonferroni-adjusted  $\alpha' = 0.05/3 \approx 0.017$  significance level. For the smallest effect size 0.67, we need 41 subjects per group to afford power of 0.80 at  $\alpha' = 0.017$ . Note that in all cases these are one-sided tests of superiority in that change in one group's mean is hypothesized to be greater than change in that of another group [63, 64]. For the smallest effect size, we need 164 subjects to complete the study, which offers ample power for the other effect sizes. Assuming 15% dropout at 6 months

based on our previous studies, the recruitment target is 192 subjects (48 per arm).

Normal continuous variables will be presented as mean  $\pm$  standard deviation. Continuous variables which cannot be transformed to achieve normality will be presented as median (interquartile range). Discrete variables will be presented as percentages.

Statistical analysis will be conducted using SAS v9.4 (Cary, NC). Independent and dependent *t*-tests will be employed to investigate differences in outcome measures between and within study groups, respectively. Any non-normal variables which cannot be transformed will be analyzed with the Mann-Whitney *U* and Wilcoxon signed-rank tests. The effect of the interventions on change in primary outcomes at mid- and post-intervention will be examined using repeated-measures linear models with restricted maximum likelihood estimation and a fully specified covariance matrix in SAS [65–67]. Such modeling can handle missing data as long as the data is missing at random, making this method ideal for our intention-to-treat analysis. Baseline covariates including the stratification variables for randomization will be adjusted for in all models. We will determine Spearman's rank correlations between changes in circulating and fecal BAs and changes in fasting glucose, insulin, HOMA-IR, and circulating pro-inflammatory cytokines. We will also determine partial correlations for the above by adjusting for pertinent covariates, such as BMI.

### Statistical analysis of the 16S rRNA data and predicted microbial metabolic pathways

ASV counts will be normalized using cumulative sum scaling [68]. Differences in community richness, evenness, and structure [alpha (i.e., Shannon index) and beta diversity (i.e., Unifrac [69] and Bray-Curtis distances)] between groups will be determined using PERMANOVA from vegan R package. Generalized linear mixed models (GLMMs) will determine the phenotypic variables, cardiometabolic outcomes, and BA-associated variables associated with each microbial ASV and genus level as implemented in MasLin [70], with treatment group as the main fixed factor, adjusting for important covariates, such as age and medication usage. Multiple comparisons will be adjusted using false discovery rate (FDR) [71]. An FDR-corrected *p*-value  $\leq 0.001$  will be deemed significant. All analyses will be done with R using functions encoded in the phyloseq package [72]. ASVs will be employed to determine the most likely associated microbial genes that encode for enzymes and associated microbial metabolic pathways using PiCRUST2 [45]. GLMMs will be employed to determine the genes and metabolic pathways associated with each intervention arm with

MasLin [70]. Finally, we will determine the likelihood that the ensemble of microbial genes associated with the production and consumption of DCA and H<sub>2</sub>S are affected by the interventions.

#### Statistical analysis of specific microbial target genes from qPCR and SCFAs

Microbial gene targets obtained via qPCR will be examined for distribution and transformed if appropriate to examine within and between group changes and differences at mid-study and post-intervention [7]. We will also examine associations between the microbial gene targets and serum and fecal BAs, circulating pro-inflammatory cytokines, and metabolic markers (e.g., fasting glucose) using principal component analyses, and canonical correspondence analyses as we have done previously [73]. Multiple linear regression analysis will be used to identify which determinant variable changes (e.g., fecal BAs) are independently associated with changes in microbial gene targets from baseline to mid-study and post-intervention, controlling for participant-level covariates such as age and medication usage.

#### Gene expression analysis in exfoliated intestinal epithelial cells

We will test for differential gene expression (DE) in the exfoliome of participants using the generalized linear models available in edgeR, followed by FDR correction [71]. The DE genes will be interrogated by the Ingenuity Pathway Analysis (Qiagen, Hilden, Germany) software and enriched regulatory pathways will be identified as previously described [49].

#### Feature selection based on linear discriminant analysis (LDA)

We will consider the regulatory genes expressed in the host exfoliome [74], the gut microbial ASVs, fecal BAs, and SCFAs as four categories of features in a classification problem that aims to detect sets of features (from the four feature categories) that best discriminate among the four groups. We will search all subsets of five or fewer features and test for significance using LDA [75, 76]. Sets with a low bolstered resubstitution misclassification error [77] ( $\leq 0.10$ ) will be reported.

#### Modeling of regulatory networks between the gut microbiome and human colonocytes

Our objective is to computationally model the host regulatory networks following exposure to the lifestyle interventions. Mathematical modeling of relationships in a gene regulatory network was originally proposed by Dougherty et al. [75] and further developed by members

of our team [78]. We will use the concept of coefficient of determination (CoD) [79], which has previously been used to measure multivariate nonlinear gene interaction [75], and define the expectation that a master regulatory gene, when active, will be responsible for regulating one or more cascades of subsequent changes in the activity of other members of the respective pathway(s). This approach models a probabilistic relationship between regulators and regulated and derives the CoD of the master regulatory gene relative to the regulated [80]. The validation of the model for the case of gene regulatory networks has been performed as previously described [80–82] and can be readily extended to any regulatory cascade including bacterial species and microbially-produced metabolites (e.g., SCFAs, DCA) that interact with the host gut. Comparisons between the groups of participants will allow us to model the microbial consortia regulatory actions as modifiers of signaling cascades in colonocytes. Because the CoD does not imply directionality, we will incorporate existing knowledge of regulatory networks to constrain the space of possible solutions. We will also use this approach to discover novel regulatory relationships between the features used in our model.

#### sCCA data integration

To assess the interdependencies among fecal microbiome, BAs, SCFAs, and host intestinal physiology in the intervention arms, we will deploy sparse Canonical Correlation Analysis (sCCA) [74, 83]. The groups of host genes, microbial ASVs, and microbial metabolites (i.e., SCFAs, DCA) identified by the 3-way sCCA will indicate potential biomarker sets for the participants in each one of the three intervention arms. We will also compare the 3-way correlative structures detected in each one of those groups to the others to elucidate the effects of the dietary interventions on the interactions between the host microbiota and the human gut relevant to CRC prevention and progression.

#### Oversight and monitoring

##### Organizational structure and responsibilities

This study will be conducted with a pre-specified organization structure assuming various responsibilities. Table 3 below details staff roles and duties.

#### Data and safety monitoring

The following information will be included in the yearly reports to the Chair of our data and safety monitoring committee: study event updates (e.g., changes in personnel, recruitment strategies, testing procedures), study enrollment (actual vs. expected), baseline data on

**Table 3** Organizational roles and responsibilities

Role	Responsibilities
Principal Investigators	<ul style="list-style-type: none"> <li>- Oversee all research-related activity and personnel</li> <li>- Publish results in a timely manner</li> </ul>
Data and Safety Monitoring Committee (DSMC)	<ul style="list-style-type: none"> <li>-Review the study protocol for any major concerns prior to implementation</li> <li>-Regularly review and evaluate the accumulated study data for participant safety, study conduct, and progress</li> <li>-Make recommendations concerning the continuation, modification, or termination of the trial</li> </ul>
Data Manager	<ul style="list-style-type: none"> <li>-Organizes electronic data</li> <li>-Data verification</li> <li>-Assist with data safety</li> </ul>
Project Director	<ul style="list-style-type: none"> <li>-Manages recruitment and data collection</li> <li>-Manages budget and staff</li> <li>-Trains and certifies staff</li> <li>-Interfaces with IRB to maintain ethical standards</li> </ul>

participants, recruitment and retention flow diagram, participant retention data (loss to follow up, withdrawals), and adverse health event monitoring. An adverse health event is any untoward medical occurrence experienced by a participant during participation in the clinical study. These events can be non-serious or serious. A serious adverse event is any adverse event that results in one or more of the following outcomes:

- A life-threatening event
- Inpatient hospitalization or prolongation of existing hospitalization
- A persistent or significant disability/incapacity
- An important medical event based upon appropriate medical judgment
- Death

#### Data and safety monitoring committee

The DSMC will consist of 12 independent experts reflecting the disciplines and medical specialties (e.g., clinical trials, biostatistics, nutrition, clinical medicine) necessary to interpret data from the clinical trial and to fully evaluate participant safety. They are independent of the study sponsor and have no competing interests. The primary responsibilities of the DSMC are outlined in Table 3. The PIs will meet with the DSMC annually and will present a written and verbal report. Additional meetings will be scheduled if warranted by the data or adverse events.

#### Adverse event monitoring and reporting

The DSMC Chair will be notified of all study-related serious adverse events within 48 h of discovery. The Chair will decide if further action beyond what the research

team has proposed is warranted. Additionally, participants will be asked to report each adverse health event in person at study visits or during the intervention sessions, or via telephone/email/text message. The principal investigators will consult and determine the severity and appropriate response or action to the reported adverse health event. All serious unanticipated problems or serious adverse health events will be reported to the UIC Institutional Review Board and Office for the Protection of Research Subjects.

UIC Office for the Protection of Research Subjects requires that all other problems (non-serious adverse events) be reported within 5 days of discovery and provide sufficient information to gauge severity and to complete reports. UIC Office for the Protection of Research Subjects has procedures in place to review the reports, and the PIs are notified if further review, changes to research protocol or consent form, or other action is required. The PIs will notify NIH should UIC IRB require further action after any unanticipated problem/adverse event is reported.

UIC has not set aside money to pay participants if they get ill or injured from being in the study. The only exception to this policy is if it is proven that a participant's injury or illness is directly caused by the negligence of UIC.

#### Discussion

The study described in this protocol will address potential solutions to reduce the high risk of CRC among AA/B adults by testing the ability of a MedDiet, weight loss, or both to reduce biological risk markers of CRC. This study will also address a potential mechanistic underpinning, the gut microbiome-bile acid axis, leading the way to more precise interventions.

To achieve these objectives, our study will use innovative and rigorous methods. First, measuring the exfolome via excreted stool is less invasive than endoscopic sampling of the gut epithelium. It is also responsive to dietary interventions, can be used to measure the human gut transcriptome, and consists of epithelial cells which have been shown to become dysplastic even before stem cells do [49, 81, 84–88]. Here, we will further establish exfoliated intestinal epithelial cells in stool as non-invasive probes in human dietary lifestyle interventions. Second, we will employ statistical and computational analyses to examine the crosstalk between diet, weight loss, gut microbiome, and host intestinal physiology [89, 90]. Third, our study design allows us to determine the independent contributions of weight loss and dietary change to reducing biomarkers of CRC risk. Nearly all studies examining the effect of weight loss on CRC risk not only restrict calories or increase physical activity but also change the diets of participants [18, 19, 91–99]. Given the strong relationship between diet and CRC [3] as well as diet and the gut microbiota [100], any change in diet may mask or alter the independent effect of weight loss on cancer risk and gut microbiota. In our study, the weight loss arm is taught to lose weight only via caloric restriction and increased physical activity, not through dietary pattern change.

Our study is not without potential pitfalls and alternative outcomes. We have prepared for some of these eventualities. For example, participants who are not randomized to their preferred arm may pursue weight loss or dietary change independent of the intervention. If, for instance, a Med-A participant loses 5% of their body weight, this may obscure differences in outcomes between Med-A and any other arm. To help avoid this, we have implemented adherence measures to be taken throughout the study so that we may intervene when appropriate. We will also perform a per-protocol analysis, including only those who are adherent to their intervention arm. Similarly, given antibiotics' profound impact on the gut microbiota, we will perform a sensitivity analysis whereby those who consume antibiotics during the intervention will be excluded from analysis to determine if results change significantly. Lastly, delivering the intervention remotely and by social media/email may dissuade less tech savvy participants during recruitment or disengage them after randomization. We determined this was a necessity to improve recruitment, as many would be participants were concerned with in-person classes given the continuing COVID-19 pandemic. Also, many work and were unable to attend in-person, group classes during working hours. Besides, remote intervention delivery is more convenient than in-person classes which would have required substantial travel time for participants,

thus potentially improving retention. If successful in a remote format, our intervention could improve recruitment and retention for future studies.

## Conclusion

Our study has the potential to show that diet and/or weight loss can lead to reductions in biomarkers of CRC risk among a group of individuals under-represented in dietary and weight loss interventions but who bear a disproportionate burden of CRC incidence and mortality. It may also reveal potential mechanisms driving the effect of diet and weight loss, allowing for more precise and effectual interventions in the future. Ultimately, our findings could have a profound public health impact on reducing CRC risk among AA/Bs and translate into timely dissemination opportunities.

## Trial status

Recruitment began in March of 2022. Recruitment is expected to be complete by December 2024. The current Protocol version is 13, dated 9/26/22.

## Abbreviations

ASA24	Automated Self-Administered 24-hour Recall
BA	Bile acids
BMI	Body mass index
CA	Cholic acid
CDCA	Chenodeoxycholic acid
CoD	Coefficient of determination
CRC	Colorectal cancer
DCA	Deoxycholic acid
DE	Differential gene expression
DSMC	Data and safety monitoring committee
EMR	Electronic medical records
ESV	Exact sequence variant
FDR	False discovery rate
GLMM	Generalized linear mixed models
GM-CSF	Granulocyte-macrophage colony-stimulating factor
HbA1c	Hemoglobin A1c
HIPAA	Health Insurance Portability and Accountability Act
HOMA-IR	Homeostatic Model Assessment for Insulin Resistance
hs-CRP	High-sensitivity C-reactive protein
IFN	Interferon
IL	Interleukin
IRB	Institutional Review Board
LCA	Lithocholic acid
LDA	Linear discriminant analysis
Med-A	Mediterranean diet alone
MedDiet	Mediterranean diet
MEPA	Mediterranean Eating Pattern for Americans
NIH	National Institutes of Health
PCR	Polymerase chain reaction
PI	Principal investigator
RCT	Randomized controlled trial
RD	Registered dietitian
sCCA	Sparse Canonical Correlation Analysis
SCFA	Short-chain fatty acid
TDC	Typical diet controls
TNF	Tumor necrosis factor
UI Health	University of Illinois Health and Hospital System
UIC	University of Illinois Chicago
WL-A	Weight loss alone

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13063-023-07115-4>.

**Additional file 1:** SPIRIT checklist.

**Additional file 2:** Consent form.

**Additional file 3: Table S1.** Med-A: Summary of the First 5 Weeks.

**Additional file 4: Table S2.** WL-A: Summary of the First 5 Weeks.

**Additional file 5: Table S3.** Med-WL: Summary of the First 5 Weeks.

## Acknowledgements

Not applicable.

## Authors' contributions

LTH and MLF conceived of the study. LTH and APM prepared the manuscript. All authors contributed to the design of the study and approved the submitted manuscript.

## Funding

This trial is funded by grant R01CA250390 from the National Cancer Institute (NCI). APM is supported by T32CA057699 (NCI). The funder shall have no influence over the design, implementation, conduct, results, or interpretations of the results of this trial.

## Availability of data and materials

De-identified results from the clinical laboratory and stool bacterial 16S rRNA amplicon sequencing, targeted qPCR, targeted metabolomics, and exfoliated intestinal cell transcriptomics, as well as information necessary to interpret the data, such as the study protocol and modifications thereof, instruments, survey tools, and variable lists, will be submitted to relevant National Institutes of Health (NIH)-designated data repositories, according to standards set forth in the Health and Human Services Regulations for the Protection of Human Participants and in a manner consistent with the informed consent provided by study participants. Basic clinical metadata such as age will also be provided.

## Declarations

### Ethics approval and consent to participate

This trial was approved by the University of Illinois Chicago (UIC) Institutional Review Board (IRB # 2020-1342). Written, informed consent and HIPAA authorization will be obtained in-person or via video conferencing by research staff trained and certified to obtain informed consent.

### Consent for publication

We may desire to publish photos and/or videos of individuals but will only do so after receiving consent to publish from said individuals.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Institute for Health Research and Policy, University of Illinois Chicago (UIC), Chicago, IL, USA. <sup>2</sup>Department of Nutrition Science, Purdue University, West Lafayette, IN, USA. <sup>3</sup>Department of Nutrition, Program in Integrative Nutrition & Complex Diseases, and Center for Environmental Health Research, Texas A&M University, College Station, TX, USA. <sup>4</sup>Department of Veterinary Physiology & Pharmacology, and Center for Environmental Health Research, Texas A&M University, College Station, TX, USA. <sup>5</sup>Mile Square Health Center, University of Illinois Chicago, Chicago, IL, USA. <sup>6</sup>Department of Animal Sciences, University of Illinois Urbana-Champaign, Urbana, IL, USA. <sup>7</sup>Division of Nutritional Sciences, University of Illinois Urbana-Champaign, Urbana, IL, USA. <sup>8</sup>Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, IL, USA. <sup>9</sup>Cancer Center at Illinois, University of Illinois Urbana-Champaign, Urbana, IL, USA. <sup>10</sup>Department of Biomedical and Translational Sciences, University of Illinois Urbana-Champaign, Urbana, IL, USA. <sup>11</sup>Department

of Pathobiology, University of Illinois Urbana-Champaign, Urbana, IL, USA.

<sup>12</sup>Department of Pediatrics, University of Illinois Chicago, Chicago, IL, USA.

<sup>13</sup>University of Illinois Cancer Center, University of Illinois Chicago, Chicago, IL, USA. <sup>14</sup>Department of Kinesiology and Nutrition, University of Illinois Chicago, Chicago, IL, USA. <sup>15</sup>Department of Biomedical Engineering, University of Illinois Chicago, Chicago, IL, USA.

Received: 17 November 2022 Accepted: 23 January 2023

Published online: 15 February 2023

## References

1. American Cancer Society. *Cancer Facts & Figures 2022*. 2022.
2. Flegal KM, Kruszon-Moran D, Carroll MD, Fryar CD, Ogden CL. Trends in obesity among adults in the United States, 2005 to 2014. *JAMA*. 2016;315(21):2284–91. <https://doi.org/10.1001/jama.2016.6458>.
3. World Cancer Research Fund/American Institute for Cancer Research. The Continuous Update Project Expert Report 2018. Diet, nutrition, physical activity and colorectal cancer. *Contin Updat Proj*. 2018;1–62.
4. Storey M, Anderson P. Income and race/ethnicity influence dietary fiber intake and vegetable consumption. *Nutr Res*. 2014;34(10):844–50. <https://doi.org/10.1016/j.nutres.2014.08.016>.
5. Shan Z, Rehm CD, Rogers G, et al. Trends in dietary carbohydrate, protein, and fat intake and diet quality among US adults, 1999–2016. *JAMA*. 2019;322(12):1178–87. <https://doi.org/10.1001/jama.2019.13771>.
6. Park S-Y, Boushey CJ, Wilkens LR, Haiman CA, Le Marchand L. High-quality diets associate with reduced risk of colorectal cancer: analyses of diet quality indexes in the multiethnic cohort. *Gastroenterol (New York, NY 1943)*. 2017;153(2):386–394.e2. <https://doi.org/10.1053/j.gastro.2017.04.004>.
7. Yazici C, Wolf PG, Kim H, et al. Race-dependent association of sulfidogenic bacteria with colorectal cancer. *Gut*. 2017;66(11):1983–94. <https://doi.org/10.1136/gutjnl-2016-313321>.
8. O'Keefe SJD, Li JV, Lahti L, et al. Fat, fibre and cancer risk in African Americans and rural Africans. *Nat Commun*. 2015;6(May 2014). <https://doi.org/10.1038/ncomms7342>.
9. Smit E, Nieto FJ, Crespo CJ, Mitchell P. Estimates of animal and plant protein intake in US adults: results from the Third National Health and Nutrition Examination Survey, 1988–1991. *J Am Diet Assoc*. 1999;99(7):813–20. [https://doi.org/10.1016/S0022-8223\(99\)00193-5](https://doi.org/10.1016/S0022-8223(99)00193-5).
10. Qin Y, Roberts JD, Grimm SA, et al. An obesity-associated gut microbiome reprograms the intestinal epigenome and leads to altered colonic gene expression. *Genome Biol*. 2018;19(1):7. <https://doi.org/10.1186/s13059-018-1389-1>.
11. Augustus GJ, Ellis NA. Colorectal cancer disparity in African Americans: risk factors and carcinogenic mechanisms. *Am J Pathol*. 2018;188(2):291–303. <https://doi.org/10.1016/j.ajpath.2017.07.023>.
12. Ridlon JM, Wolf PG, Gaskins HR. Taurocholic acid metabolism by gut microbes and colon cancer. *Gut Microbes*. 2016;7(3):201–15. <https://doi.org/10.1080/19490976.2016.1150414>.
13. Ridlon JM, Kang DJ, Hylemon PB, Bajaj JS. Bile acids and the gut microbiome. *Curr Opin Gastroenterol*. 2014;30(3):332–8. <https://doi.org/10.1097/MOG.0000000000000057>.
14. Haeusler RA, Camasta S, Nannipieri M, et al. Increased bile acid synthesis and impaired bile acid transport in human obesity. *J Clin Endocrinol Metab*. 2016;101(5):1935–44. <https://doi.org/10.1210/jc.2015-2583>.
15. Reddy BS. Diet and excretion of bile acids. *Cancer Res*. 1981;41(9 Pt 2):3766–8.
16. Babidge W, Millard S, Roediger W. Sulfides impair short chain fatty acid  $\beta$ -oxidation at acyl-CoA dehydrogenase level in colonocytes: implications for ulcerative colitis. *Mol Cell Biochem*. 1998;181(1-2):117–24. <https://doi.org/10.1023/A:1006838231432>.
17. Chen J, Zhao K-N, Vitetta L. Effects of intestinal microbial–elaborated butyrate on oncogenic signaling pathways. *Nutrients*. 2019;11(5):1026. <https://doi.org/10.3390/nu11051026>.
18. Aleman JO, Bokulich NA, Swann JR, et al. Fecal microbiota and bile acid interactions with systemic and adipose tissue metabolism in diet-induced weight loss of obese postmenopausal women. *J Transl Med*. 2018;16(1):244. <https://doi.org/10.1186/s12967-018-1619-z>.

19. Biemann R, Penner M, Borucki K, et al. Serum bile acids and GLP-1 decrease following telemetric induced weight loss: results of a randomized controlled trial. *Sci Rep.* 2016;6(1):30173. <https://doi.org/10.1038/srep30173>.

20. Shively CA, Register TC, Appt SE, et al. Consumption of Mediterranean versus Western diet leads to distinct mammary gland microbiome populations. *Cell Rep.* 2018;25(1):47–56.e3. <https://doi.org/10.1016/j.celrep.2018.08.078>.

21. Ghosh TS, Rampelli S, Jeffery IB, et al. Mediterranean diet intervention alters the gut microbiome in older people reducing frailty and improving health status: the NU-AGE 1-year dietary intervention across five European countries. *Gut.* 2020;69(7):1218 LP–1228. <https://doi.org/10.1136/gutjnl-2019-319654>.

22. Jones P, Cade JE, Evans CEL, Hancock N, Greenwood DC. The Mediterranean diet and risk of colorectal cancer in the UK Women's Cohort Study. *Int J Epidemiol.* 2017;46(6):1786–96. <https://doi.org/10.1093/ije/dyx155>.

23. Blumenthal JA, Babyak MA, Hinderliter A, et al. Effects of the DASH diet alone and in combination with exercise and weight loss on blood pressure and cardiovascular biomarkers in men and women with high blood pressure: the ENCORE study. *Arch Intern Med.* 2010;170(2):126–35. <https://doi.org/10.1001/archinternmed.2009.470>.

24. Franz MJ, VanWormer JJ, Crain AL, et al. Weight-loss outcomes: a systematic review and meta-analysis of weight-loss clinical trials with a minimum 1-year follow-up. *J Am Diet Assoc.* 2007;107(10):1755–67. <https://doi.org/10.1016/j.jada.2007.07.017>.

25. Neltje R, E R, Goldman J, E W, AR G, Tangney C. Acceptability and reliability of the paper and online versions of the MEPA-III screener. *J Hum Nutr* 2020;4. doi:<https://doi.org/10.36959/487/287>

26. Defining the Role of Authors and Contributors. Int Comm Med J Ed. <https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html>. Accessed 1 Nov 2022.

27. Bandura A. Human agency in social cognitive theory. *Am Psychol.* 1989;44(9):1175–84. <https://doi.org/10.1037/0003-066X.44.9.1175>.

28. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol.* 2000;55(1):68–78. <https://doi.org/10.1037/0003-066X.55.1.68>.

29. Djuric Z, Vanloon G, Radakovich K, DiLaura NM, Heilbrun LH, Sen A. Design of a Mediterranean exchange list diet implemented by telephone counseling. *J Am Diet Assoc.* 2008;108(12):2059–65. <https://doi.org/10.1016/j.jada.2008.09.006>.

30. Tussing-Humphreys L, Lamar M, Blumenthal JA, et al. Building research in diet and cognition: The BRIDGE randomized controlled trial. *Contemp Clin Trials.* 2017;59:87–97. <https://doi.org/10.1016/j.jcct.2017.06.003>.

31. National Diabetes Prevention Program. 2012. <https://www.cdc.gov/diabetes/prevention/resources/curriculum.html>. Accessed 1 Nov 2022.

32. Dahle JH, Ostendorf DM, Pan Z, et al. Weight and body composition changes affect resting energy expenditure predictive equations during a 12-month weight-loss intervention. *Obesity (Silver Spring).* 2021;29(10):1596–605. <https://doi.org/10.1002/oby.23234>.

33. Hughes SL, Tussing-Humphreys L, Schiffer L, et al. Fit & Strong! Plus trial outcomes for obese older adults with osteoarthritis. *Gerontologist.* 2020;60(3):558–70. <https://doi.org/10.1093/geront/gny146>.

34. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377–81. <https://doi.org/10.1016/j.jbi.2008.08.010>.

35. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and  $\beta$ -cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia.* 1985;28(7):412–9. <https://doi.org/10.1007/BF00280883>.

36. Kristal AR, Kolar AS, Fisher JL, et al. Evaluation of web-based, self-administered, graphical food frequency questionnaire. *J Acad Nutr Diet.* 2014;114(4):613–21. <https://doi.org/10.1016/j.jand.2013.11.017>.

37. Automated Self-Administered 24-Hour (ASA24<sup>®</sup>) Dietary Assessment Tool. <https://epi.grants.cancer.gov/asa24/>. Accessed 5 May 2022.

38. Panagiotakos DB, Pitsavos C, Arvaniti F, Stefanidis C. Adherence to the Mediterranean food pattern predicts the prevalence of hypertension, hypercholesterolemia, diabetes and obesity, among healthy adults; the accuracy of the MedDietScore. *Prev Med (Baltimore).* 2007;44(4):335–40. <https://doi.org/10.1016/j.ypmed.2006.12.009>.

39. Tangney CC, Kwasny MJ, Li H, Wilson RS, Evans DA, Morris MC. Adherence to a Mediterranean-type dietary pattern and cognitive decline in a community population. *Am J Clin Nutr.* 2011;93(3):601–7. <https://doi.org/10.3945/ajcn.110.007369>.

40. Naqib A, Poggi S, Wang W, Hyde M, Kunstman K, Green SJ. Making and sequencing heavily multiplexed, high-throughput 16S ribosomal RNA gene amplicon libraries using a flexible, two-stage PCR protocol. *Methods Mol Biol.* 2018;1783:149–69. [https://doi.org/10.1007/978-1-4939-7834-2\\_7](https://doi.org/10.1007/978-1-4939-7834-2_7).

41. Bybee SM, Bracken-Grissom H, Haynes BD, et al. Targeted amplicon sequencing (TAS): a scalable next-gen approach to multilocus, multitaxa phylogenetics. *Genome Biol Evol.* 2011;3:1312–23. <https://doi.org/10.1093/gbe/evr106>.

42. Barengolts E, Green SJ, Eisenberg Y, et al. Gut microbiota varies by opioid use, circulating leptin and oxytocin in African American men with diabetes and high burden of chronic disease. *PLoS One.* 2018;13(3):e0194171. <https://doi.org/10.1371/journal.pone.0194171>.

43. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods.* 2016;13(7):581–3. <https://doi.org/10.1038/nmeth.3869>.

44. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2013;41(Database issue):D590–6. <https://doi.org/10.1093/nar/gks1219>.

45. Douglas GM, Maffei VJ, Zaneveld JR, et al. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol.* 2020;38(6):685–8. <https://doi.org/10.1038/s41587-020-0548-6>.

46. Kurtz ZD, Müller CL, Miraldi ER, Littman DR, Blaser MJ, Bonneau RA. Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol.* 2015;11(11):e1004226. <https://doi.org/10.1371/journal.pcbi.1004226>.

47. Ou J, Carbonero F, Zoetendal EG, et al. Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *Am J Clin Nutr.* 2013;98(1):111–20. <https://doi.org/10.3945/ajcn.112.056689>.

48. Ridlon JM, Devendran S, Alves JM, et al. The “in vivo lifestyle” of bile acid 7 $\alpha$ -dehydroxylating bacteria: comparative genomics, metatranscriptomic, and bile acid metabolomics analysis of a defined microbial community in gnotobiotic mice. *Gut Microbes.* 2020;11(3):381–404. <https://doi.org/10.1080/19490976.2019.1618173>.

49. Lampe JW, Kim E, Levy L, et al. Colonic mucosal and exfoliate transcriptomic profiling and fecal microbiome response to a flaxseed lignan extract intervention in humans. *Am J Clin Nutr.* 2019;110(2):377–90. <https://doi.org/10.1093/ajcn/nqy325>.

50. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–9. <https://doi.org/10.1038/nmeth.1923>.

51. D'Antonio M, D'Onorio De Meo P, Pallocca M, et al. RAP: RNA-Seq Analysis Pipeline, a new cloud-based NGS web application. *BMC Genomics.* 2015;16(6):S3. <https://doi.org/10.1186/1471-2164-16-S6-S3>.

52. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015;31(2):166–9. <https://doi.org/10.1093/bioinformatics/btu638>.

53. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Questionnaire. Hyattsville: CDC; 2012.

54. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 1989;28(2):193–213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4).

55. Williams DR, Yu Y, Jackson JS, Anderson NB. Racial differences in physical and mental health: socio-economic status, stress and discrimination. *J Health Psychol.* 1997;2(3):335–51. <https://doi.org/10.1177/135910539700200305>.

56. Economic Research Service. U.S. Household Food Security Survey Module. <https://www.ers.usda.gov/media/8282/short2012.pdf>. Published 2012. Accessed 5 Dec 2022.

57. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J*

Affect Disord. 2009;114(1-3):163–73. <https://doi.org/10.1016/j.jad.2008.06.026>.

58. Cohen S, Williamson G. Perceived stress in a probability sample of the United States. *Soc Psychol Heal*. 1988;13:31–67 <http://doi.apa.org/psychinfo/1988-9883-002>.

59. Pennebaker JW, Susman JR. Disclosure of traumas and psychosomatic processes. *Soc Sci Med*. 1988;26(3):327–32. [https://doi.org/10.1016/0277-9536\(88\)90397-8](https://doi.org/10.1016/0277-9536(88)90397-8).

60. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092–7. <https://doi.org/10.1001/archinte.166.10.1092>.

61. McClain AC, Ayala GX, Sotres-Alvarez D, et al. Frequency of intake and type of away-from-home foods consumed are associated with diet quality in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). *J Nutr*. 2018;148(3):453–63. <https://doi.org/10.1093/jn/nxx067>.

62. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol*. 2014;14:135. <https://doi.org/10.1186/1471-2288-14-135>.

63. Julious SA, Patterson SD. Sample sizes for estimation in clinical research. *Pharm Stat*. 2004;3(3):213–5. <https://doi.org/10.1002/pst.125>.

64. Glaeck DH. Sample size calculations in clinical research 2nd edition by CHOW, S.-C., SHAO, J., and WANG, H. *Biometrics*. 2008;64(4):1307–8. [https://doi.org/10.1111/j.1541-0420.2008.01138\\_10.x](https://doi.org/10.1111/j.1541-0420.2008.01138_10.x).

65. Littell RC, Henry PR, Ammerman CB. Statistical analysis of repeated measures data using SAS procedures. *J Anim Sci*. 1998;76(4):1216–31. <https://doi.org/10.2527/1998.7641216x>.

66. Mallinckrod CH, Lane PW, Schnell D, Peng Y, Mancuso JP. Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. *Drug Inf J*. 2008;42(4):303–19. <https://doi.org/10.1177/09286150804200402>.

67. Brown H, Prescott R. Applied mixed models in medicine. 2nd ed. Hoboken: Wiley; 2006. <https://doi.org/10.1002/0470023589.mfatter>.

68. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods*. 2013;10(12):1200–2. <https://doi.org/10.1038/nmeth.2658>.

69. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol*. 2005;71(12):8228–35. <https://doi.org/10.1128/AEM.71.12.8228-8235.2005>.

70. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6):R60. <https://doi.org/10.1186/gb-2011-12-6-r60>.

71. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B*. 1995;57(1):289–300.

72. McMurdie PJ, Holmes S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 2013;8(4):e61217. <https://doi.org/10.1371/journal.pone.0061217>.

73. Dugas LR, Bernabé BP, Priyadarshini M, et al. Decreased microbial co-occurrence network stability and SCFA receptor level correlates with obesity in African-origin women. *Sci Rep*. 2018;8(1):17135. <https://doi.org/10.1038/s41598-018-35230-9>.

74. Schwartz S, Friedberg I, Ivanov IV, et al. A metagenomic study of diet-dependent interaction between gut microbiota and host in infants reveals differences in immune response. *Genome Biol*. 2012;13(4):r32. <https://doi.org/10.1186/gb-2012-13-4-r32>.

75. Dougherty E, Kim S, Chen Y. Coefficient of determination in nonlinear signal processing. *Signal Process*. 2000;80:2219–35. [https://doi.org/10.1016/S0165-1684\(00\)00079-7](https://doi.org/10.1016/S0165-1684(00)00079-7).

76. Sima C, Dougherty ER. What should be expected from feature selection in small-sample settings. *Bioinformatics*. 2006;22(19):2430–6. <https://doi.org/10.1093/bioinformatics/btl407>.

77. Sima C, Braga-Neto UM, Dougherty ER. High-dimensional bolstered error estimation. *Bioinformatics*. 2011;27(21):3056–64. <https://doi.org/10.1093/bioinformatics/btr518>.

78. Kim E, Ivanov I, Dougherty ER. Quantifying the notions of canalizing and master genes in a gene regulatory network-a Boolean network modeling perspective. *Bioinformatics*. 2019;35(4):643–9. <https://doi.org/10.1093/bioinformatics/bty665>.

79. Kim S, Dougherty ER, Shmulevich I, et al. Identification of combination gene sets for glioma classification. *Mol Cancer Ther*. 2002;1(13):1229–36.

80. Dougherty ER, Brun M, Trent JM, Bittner ML. Conditioning-based modeling of contextual genomic regulation. *IEEE/ACM Trans Comput Biol Bioinforma*. 2009;6(2):310–20. <https://doi.org/10.1109/TCBB.2007.70247>.

81. Chapkin RS, Zhao C, Ivanov I, et al. Noninvasive stool-based detection of infant gastrointestinal development using gene expression profiles from exfoliated epithelial cells. *Am J Physiol Gastrointest Liver Physiol*. 2010;298(5):G582–9. <https://doi.org/10.1152/ajpgl.00004.2010>.

82. Fornace AJJ, Alamo IJ, Hollander MC. DNA damage-inducible transcripts in mammalian cells. *Proc Natl Acad Sci U S A*. 1988;85(23):8800–4. <https://doi.org/10.1073/pnas.85.23.8800>.

83. Witten DM, Tibshirani RJ. Extensions of sparse canonical correlation analysis with applications to genomic data. *Stat Appl Genet Mol Biol*. 2009;8(1):Article28. <https://doi.org/10.2202/1544-6115.1470>.

84. Donovan SM, Wang M, Monaco MH, et al. Noninvasive molecular fingerprinting of host-microbiome interactions in neonates. *FEBS Lett*. 2014;588(22):4112–9. <https://doi.org/10.1016/j.febslet.2014.07.008>.

85. Davidson LA, Lupton JR, Miskovsky E, Fields AP, Chapkin RS. Quantification of human intestinal gene expression profiles using exfoliated colonocytes: a pilot study. *Biomarkers Biochem Indic Expo response, susceptibility to Chem*. 2003;8(1):51–61. <https://doi.org/10.1080/1354750021000042268>.

86. Shih IM, Wang TL, Traverso G, et al. Top-down morphogenesis of colorectal tumors. *Proc Natl Acad Sci U S A*. 2001;98(5):2640–5. <https://doi.org/10.1073/pnas.051629398>.

87. Huels DJ, Sansom OJ. Stem vs non-stem cell origin of colorectal cancer. *Br J Cancer*. 2015;113(1):1–5. <https://doi.org/10.1038/bjc.2015.214>.

88. Schwitalla S, Fingerle AA, Cammareri P, et al. Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell*. 2013;152(1-2):25–38. <https://doi.org/10.1016/j.cell.2012.12.012>.

89. Mullens DA, Ivanov I, Hullar MAJ, Randolph TW, Lampe JW, Chapkin RS. Personalized nutrition using microbial metabolite phenotype to stratify participants and non-invasive host exfoliomics reveal the effects of flaxseed lignan supplementation in a placebo-controlled crossover trial. *Nutrients*. 2022;14(12). <https://doi.org/10.3390/nu14122377>.

90. He K, Donovan SM, Ivanov IV, Goldsby JS, Davidson LA, Chapkin RS. Assessing the multivariate relationship between the human infant intestinal exfoliated cell transcriptome (exfoliome) and microbiome in response to diet. *Microorganisms*. 2020;8(12). <https://doi.org/10.3390/microorganisms8122032>.

91. Magkos F, Sullivan S, Fitch M, et al. Effect of weight gain and weight loss on in vivo colonocyte proliferation rate in people with obesity. *Obesity (Silver Spring)*. 2017;25(Suppl 2):S81–6. <https://doi.org/10.1002/oby.21983>.

92. Afshar S, Kelly SB, Seymour K, Lara J, Woodcock S, Mathers JC. The effects of bariatric surgery on colorectal cancer risk: systematic review and meta-analysis. *Obes Surg*. 2014;24(10):1793–9. <https://doi.org/10.1007/s11695-014-1359-y>.

93. Adams TD, Hunt SC. Cancer and obesity: effect of bariatric surgery. *World J Surg*. 2009;33(10):2028–33. <https://doi.org/10.1007/s00268-009-0169-1>.

94. Ostlund MP, Lu Y, Lagergren J. Risk of obesity-related cancer after obesity surgery in a population-based cohort study. *Ann Surg*. 2010;252(6):972–6. <https://doi.org/10.1097/SLA.0b013e3181e33778>.

95. Derogar M, Hull MA, Kant P, Östlund M, Lu Y, Lagergren J. Increased risk of colorectal cancer after obesity surgery. *Ann Surg*. 2013;258(6):983–8. <https://doi.org/10.1097/SLA.0b013e318288463a>.

96. Beeken RJ, Croker H, Heinrich M, et al. The impact of diet-induced weight loss on biomarkers for colorectal cancer: an exploratory study (INTERCEPT). *Obesity*. 2017;25(Suppl 2):S95–S101. <https://doi.org/10.1002/oby.21984>.

97. Pendyala S, Neff LM, Suárez-Fariñas M, Holt PR. Diet-induced weight loss reduces colorectal inflammation: implications for colorectal carcinogenesis. *Am J Clin Nutr*. 2011;93(2):234–42. <https://doi.org/10.3945/ajcn.110.002683>.

98. Alemán JO, Iyengar NM, Walker JM, et al. Effects of rapid weight loss on systemic and adipose tissue inflammation and metabolism in obese

postmenopausal women. *J Endocr Soc.* 2017;1(6):625–37. <https://doi.org/10.1210/js.2017-00020>.

99. Madigan CD, Daley AJ, Kabir E, Aveyard P, Brown W. Cluster analysis of behavioural weight management strategies and associations with weight change in young women: a longitudinal analysis. *Int J Obes (Lond).* 2015;39(11):1601–6. <https://doi.org/10.1038/ijo.2015.116>.

100. Valdes AM, Walter J, Segal E, Spector TD. Role of the gut microbiota in nutrition and health. *BMJ.* 2018;361:36–44. <https://doi.org/10.1136/bmj.k2179>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](http://biomedcentral.com/submissions)

