UCSF
UC San Francisco Previously Published Works

Title

Variety of Fruit and Vegetables and Alcohol Intake are Associated with Gut Microbial Species and Gene Abundance in Colorectal Cancer Survivors

Permalink

https://escholarship.org/uc/item/6vc348d2

Journal

American Journal of Clinical Nutrition, 118(3)

ISSN

0002-9165

Authors

Kyaw, Than S
Upadhyay, Vaibhav
Tolstykh, Irina
et al.

Publication Date

2023-09-01

DOI

10.1016/j.ajcnut.2023.07.011

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

The American Journal of Clinical Nutrition 118 (2023) 518–529

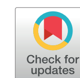
American Society for Nutrition
Excellence in Nutrition Research and Practice

The American Journal of
CLINICAL NUTRITION

journal homepage: https://ajcn.nutrition.org/

Original Research Article

Variety of Fruit and Vegetables and Alcohol Intake are Associated with Gut Microbial Species and Gene Abundance in Colorectal Cancer Survivors



Than S. Kyaw [1], Vaibhav Upadhyay [1,2], Irina Tolstykh [3], Katherine Van Loon [2,4], Angela Laffan [4], Dalila Stanfield [4], Daryll Gempis [1], Stacey A. Kenfield [3,5], June M. Chan [3,5], Sorbarikor Piawah [2,4], Chloe E. Atreya [2,4], Kimmie Ng [6], Alan Venook [2,4], Wesley Kidder [2,4], Peter J. Turnbaugh [1,7,*,†], Erin L. Van Blarigan [3,5,**,†]

[1] *Department of Microbiology and Immunology, University of California, San Francisco, CA, United States;* [2] *Department of Medicine, University of California, San Francisco, CA, United States;* [3] *Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, United States;* [4] *University of California San Francisco Helen Diller Family Comprehensive Cancer Center, San Francisco, CA, United States;* [5] *Department of Urology, University of California, San Francisco, United States;* [6] *Dana Farber Cancer Institute, Boston, MA, United States;* [7] *Chan Zuckerberg Biohub - San Francisco, San Francisco, CA, United States*

A B S T R A C T

Background: Adherence to the American Cancer Society (ACS) guidelines of avoiding obesity, maintaining physical activity, and consuming a diet rich in fruits, vegetables, and whole grains is associated with longer survival in colorectal cancer (CRC) survivors. Dietary components of the ACS guidelines may act in part by changing the microbiome, which is implicated in CRC outcomes.
Objectives: We conducted a pilot cross-sectional study to explore associations between ACS guidelines and the gut microbiome.
Methods: Stool samples and questionnaires were collected from 28 CRC survivors at the University of California, San Francisco from 2019 to 2020. ACS scores were calculated based on validated questionnaires. Gut microbial community structure from 16S amplicons and gene/pathway abundances from metagenomics were tested for associations with the ACS score and its components using ANOVA and general linear models.
Results: The overall ACS score was not significantly associated with variations in the fecal microbiota. However, fruit and vegetable intake and alcohol intake accounted for 19% ($P = 0.005$) and 13% ($P = 0.01$) of variation in the microbiota, respectively. Fruit/vegetable consumption was associated with increased microbial diversity, increased Firmicutes, decreased Bacteroidota, and changes to multiple genes and metabolic pathways, including enriched pathways for amino acid and short-chain fatty acid biosynthesis and plant-associated sugar degradation. In contrast, alcohol consumption was positively associated with overall microbial diversity, negatively associated with Bacteroidota abundance, and associated with changes to multiple genes and metabolic pathways. The other components of the ACS score were not statistically significantly associated with the fecal microbiota in our sample.
Conclusions: These results guide future studies examining the impact of changes in the intake of fruits, vegetables, and alcoholic drinks on the gut microbiome of CRC survivors.

Keywords: colorectal cancer, survivorship, nutrition, human gut microbiome, fruits and vegetables, alcohol

Introduction

Colorectal cancer (CRC) is the second-leading cause of cancer death in the United States, with 52,580 deaths in 2022 [1]. Health behaviors are associated with risk of CRC mortality [2,3]. For example, the American Cancer Society (ACS) Nutrition and Physical Activity Guidelines for Cancer Survivors (hereafter referred to as the ACS guidelines) recommend avoiding obesity, being physically active, and consuming a healthy diet rich in fruits, vegetables, and whole grains [2]. Our team previously reported that adherence to these guidelines

Abbreviations: ACS, American Cancer Society; ANOVA, analysis of variance; BMI, body mass index; CLR, centered-log-ratio; CRC, colorectal cancer; FFQ, food frequency questionnaire; GLM, general linear model; IQR, interquartile range; LOGIC, Lifestyle and Outcomes after Gastro-Intestinal Cancer; MET-h/wk, metabolic equivalent task-hours per week; PERMANOVA, permutational multivariate analysis of variance; rRNA, ribosomal RNA; SCFA, short-chain fatty acid; UCSF, University of California San Francisco.

* Corresponding author.
** Corresponding author.
E-mail addresses: peter.turnbaugh@ucsf.edu (P.J. Turnbaugh), erin.vanblarigan@ucsf.edu (E.L. Van Blarigan).
† PJT and ELVB are shared last authors.

https://doi.org/10.1016/j.ajcnut.2023.07.011
Received 11 December 2022; Received in revised form 28 June 2023; Accepted 11 July 2023; Available online 18 July 2023
0002-9165/© 2023 The Author(s). Published by Elsevier Inc. on behalf of American Society for Nutrition. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

was associated with a 42% lower risk of death and longer 5-y survival among stage III colon cancer patients [3]. However, the mechanisms through which health behaviors may alter CRC survival are poorly understood. Motivated by the rapidly expanding literature on the role of the microbiome in cancer [4–7] and the key role diet plays in shaping the gut microbiome [8–12], we hypothesized that dietary components of the ACS guidelines act in part by changing the gut microbial communities (microbiota) or their aggregate genomes and metabolic activities (microbiome).

There is now a large and rigorous literature providing support for a causal role of the microbiome in the etiology and treatment of CRC. Multiple procarcinogenic bacteria have been identified, including members of the *Fusobacterium nucleatum* [13], *Escherichia coli* [4], and *Bacteroides fragilis* [4,5] species. Broadly speaking, the mechanisms responsible include microbial effects on the host immune system [6] and the production of genotoxic metabolites [14,15]. Furthermore, the tremendous metabolic diversity found within the human gut and tumor-associated microbial communities can alter the metabolism and absorption of multiple anti-cancer drugs [16–19].

Diet is perhaps the most important modifiable factor that influences the gut microbiome [8,9]. Short-term dietary interventions in healthy individuals lead to significant shifts in gut bacterial abundance and gene expression [10]. Dietary interventions in obese subjects have revealed marked changes in the gut microbiome in response to caloric restriction [11] and the consumption of a high-fat, low-carbohydrate ketogenic diet [12]. As noted above, we and others have reported that diets rich in whole grains and fruits and vegetables were associated with a lower risk of recurrence and death in CRC survivors [3,20,21]. However, data on the links between diet and the gut microbiomes of CRC patients remain lacking, especially in the context of survivorship.

Here, we report the results of a pilot study of 28 CRC survivors at University of California, San Francisco (UCSF) to investigate whether adherence to the ACS guidelines was associated with interindividual variations in the gut microbiomes of CRC survivors.

Methods

Study participants

Participants for this pilot study were recruited from a single-center open cohort of gastrointestinal cancer survivors, Lifestyle and Outcomes after Gastrointestinal Cancer (LOGIC). This study was approved by the UCSF institutional review board and conducted accordingly. All participants provided informed consent.

The goal of the LOGIC study is to examine health behaviors, quality of life, and clinical outcomes among gastrointestinal cancer survivors. Adults (aged ≥18 y) who have been diagnosed with any gastrointestinal cancer, can complete online surveys in English, and have been seen at UCSF are eligible to enroll. Participants were recruited through the Gastrointestinal Oncology Survivorship Clinic as well as through MyChart invitations sent to UCSF patients with a diagnosis of any gastrointestinal cancer in their medical record. For this feasibility study focused on the gut microbiome, we sent 1 e-mail invitation in October 2019 to 152 active participants in LOGIC who had a diagnosis of colon or rectal cancer and had consented to be contacted about future research opportunities. The study staff also invited 4 participants who enrolled in the LOGIC in October 2019 after the initial e-mail invitation. In February 2020, we sent a second invitation to 56 people who had been invited in October and had not replied, as well as 5 new participants, so the total number invited was 161 (Figure 1A). The stool collection sub-study was closed in March 2020 due to the onset of the COVID-19 pandemic, so no additional invitations were sent out. Response rates are provided in the results section.

Survey data

Participants in LOGIC are asked to complete surveys online using Research Electronic Data Capture [22] at enrollment and every 6 mo thereafter for 5 y. The survey sent at enrollment includes sociodemographics (for example, self-reported gender, race, ethnicity, education), medical history (for example, height, weight), a validated food frequency questionnaire (FFQ) [23–25], and a validated physical activity questionnaire [26]. Self-reported medical history (including weight) is updated every 12 mo; the FFQ is repeated once at 18 mo; and the baseline physical activity questionnaire is repeated at 24 and 48 mo.

Participants who consented to the optional stool sub-study were asked to complete an additional survey at the time of stool collection that asked about antibiotic, prebiotic, and probiotic use. They were also asked to complete an FFQ and a physical activity survey at the time of stool collection if they had not completed one within the past 12 mo as part of the parent study. Most recent survey data were used for all analyses reported in this manuscript.

Dietary assessment

We used a semiquantitative FFQ available to external investigators from the Nutrition Department of the Harvard T.H. Chan School of Public Health. This FFQ is based on the extensively validated FFQs used in the Nurses' Health Study and Health Professionals' Follow-up Study [24,25]. The FFQ includes ~150 items. For each item, portion size is specified, and participants are asked how often, on average, they consumed that amount of each food over the past year. Nine response options range from never or less than once per month to 6 or more times per day. The FFQs were processed for nutrient analyses by the Harvard School of Public Health Nutrition Department. To calculate nutrient intake (for example, grams/day of fiber), the amount of a given nutrient in the specific serving size of each food item is determined based on data from the United States Department of Agriculture and other sources. These nutrient amounts are then multiplied by the frequency of intake and summed across all food items.

ACS guideline score

We used a standardized score to estimate adherence to the ACS guidelines that has been previously described (Figure 1B) [27,28]. Briefly, participants were assigned 0–2 points for body mass index (BMI), total physical activity metabolic equivalent task-hours per week (MET-h/wk), and dietary habits. For BMI, 0 points were assigned to BMI <18.5 kg/m^2 or ≥30 kg/m^2, 1 point for a BMI of 25.0–29.9 kg/m^2, and 2 points for a BMI of 18.5–24.9 kg/m^2. For physical activity, 0 points were assigned for <8.75 MET-h/wk, 1 point for 8.75 to <17.5 MET-h/wk, and 2 points for 17.5 MET-h/wk or more.

To assign the 0–2 points for diet, a diet sub-score was first calculated based on intake of fruits and vegetables (1 point for ≥5 servings/d, 0 points if <5 servings/d), the number of unique fruits and vegetables (0–2 points assigned to sex-specific tertiles; Figure 2A), percent of total grains that are whole grains (0–3 points assigned to sex-specific quartiles), and intake of red and processed meat (3–0 points assigned to sex-specific quartiles with the highest score given to the lowest quartile). Fruit and vegetable intake represents the combined points for quantity (0 or 1) and variety (0–2) for a total range of 0–3 (Figure 2A). The total diet score summed points from fruits and vegetables, grain, and meat (possible range of 0–9), and then was reweighted to 0–2 points as follows: 0–2 = 0 points, 3–6 = 1 point, and 7–9 = 2 points.

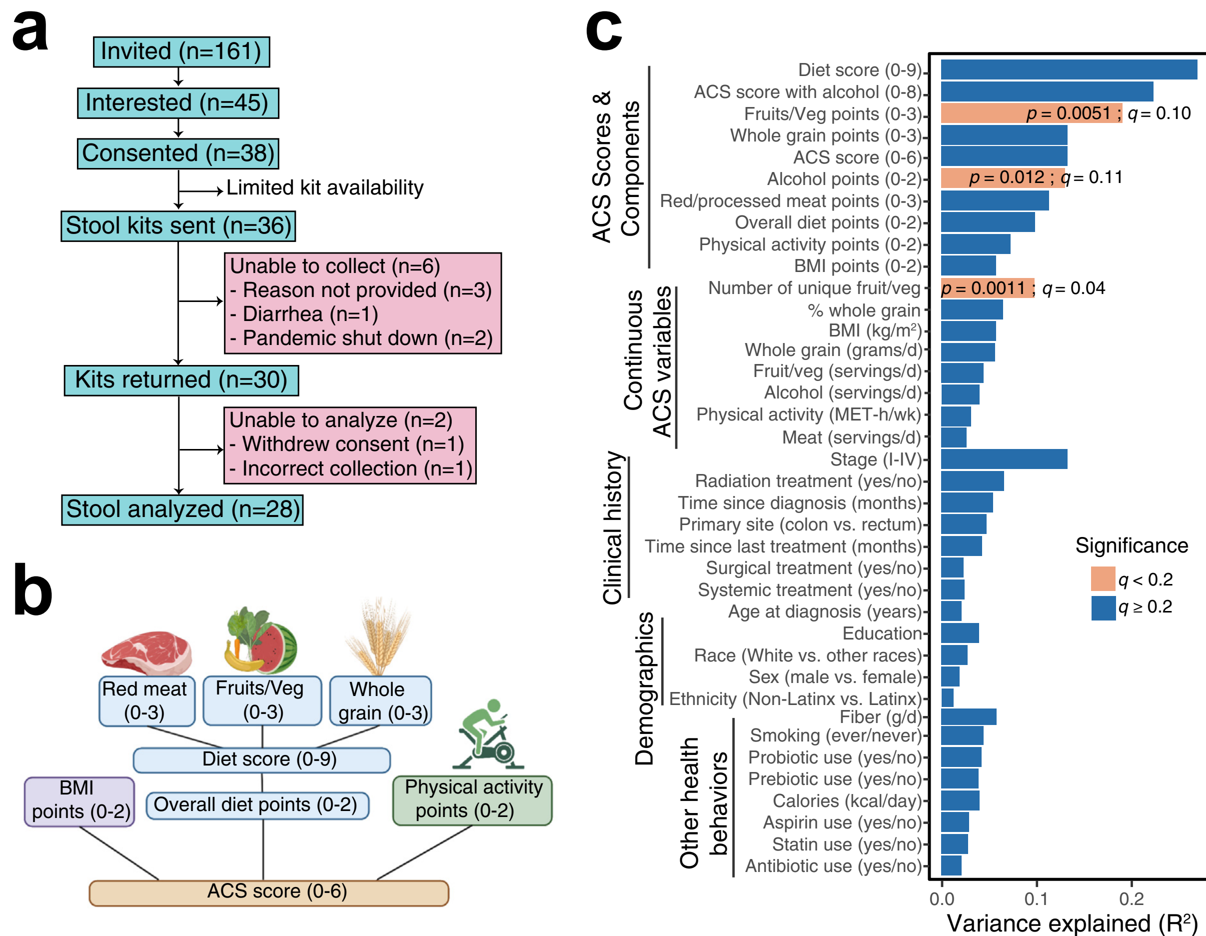


FIGURE 1. Components of American Cancer Society (ACS) scores significantly account for variations in the microbiota of colorectal cancer survivors. (A) Derivation of our study population (Table 1). (B) A schematic for calculating the ACS score. Fruit/veg points include the sum of points for amount (5+ servings/d = 1 point; <5 servings/d = 0 points) and tertiles of variety (0, 1, 2 points). When examining the ACS score with alcohol, the alcohol points (0–2) are added to the ACS score for a total range of 0–8. (C) Permutational multivariate analysis of variance (PERMANOVA) testing of ACS score and component points, clinical history, demographics, other health behavior variables, and continuous variables used to calculate ACS scores. Nominal P values were reported from the PERMANOVA test (ADONIS function in the Vegan package in R) using the weighted-UniFrac beta diversity metric of bacterial community composition, and false discovery rates q were calculated with Benjamini-Hochberg multiple-testing corrections (Supplementary Table 2). A PERMANOVA test with a different beta diversity metric (Bray-Curtis) is shown in Supplementary Fig. 2a. $N = 27$ participants. BMI, body mass index.

The points for BMI, physical activity, and diet were summed for a total score ranging from 0 to 6 points.

We also explored a secondary score that included alcohol. This score assigned 0 points to excessive drinkers (>1 drink/d for women, >2 drinks/d for men), 1 point to nondrinkers, and 2 points to low-to-moderate drinkers (1 drink/d or fewer for women, 2 drinks/d or fewer for men). We chose this scoring system because it was previously shown to be associated with colon cancer survival [3]. Points for alcohol were added to the standard ACS score, creating a score that ranged from 0 to 8 points.

For ACS score calculation, we included the most recent data and data collected ≤6 mo after stool collection because we considered the analysis cross-sectional, and the FFQ and physical activity survey both asked participants to recall their usual behaviors over the past year.

Other dietary variables

In addition to examining the ACS score and its components, we explored intake of calories, fiber [29], glycemic index [30], glycemic load [30], long-chain omega-3 fatty acids [31], fish [31], vitamin D [32], coffee [33,34], and tree nuts [35] in relation to the gut microbiome because these factors have been associated with CRC mortality [35] but are not included in the ACS guideline score.

Clinical data

Clinical data, including date of diagnosis, cancer type, stage at diagnosis, and treatment history, were abstracted from medical records by trained staff at enrollment in LOGIC. The medical records of LOGIC participants were reviewed annually to abstract information on any treatments received in the past year as well as any diagnoses of recurrence or metastasis.

Fecal sample and data collection

Participants were provided a commercially available OMNIgene-GUT kit (DNA Genotek Inc), which contained a stool collection tube with preservatives, user instructions, a spatula, 2 toilet accessories, a biospecimen bag, a 2-way mailer prepaid shipping box, and a custom label. Per user instruction, participants were asked to mix the fecal sample with preservatives immediately upon collection and ship it at room temperature to the Turnbaugh laboratory at UCSF. The samples were stored at −80°C until further analysis.

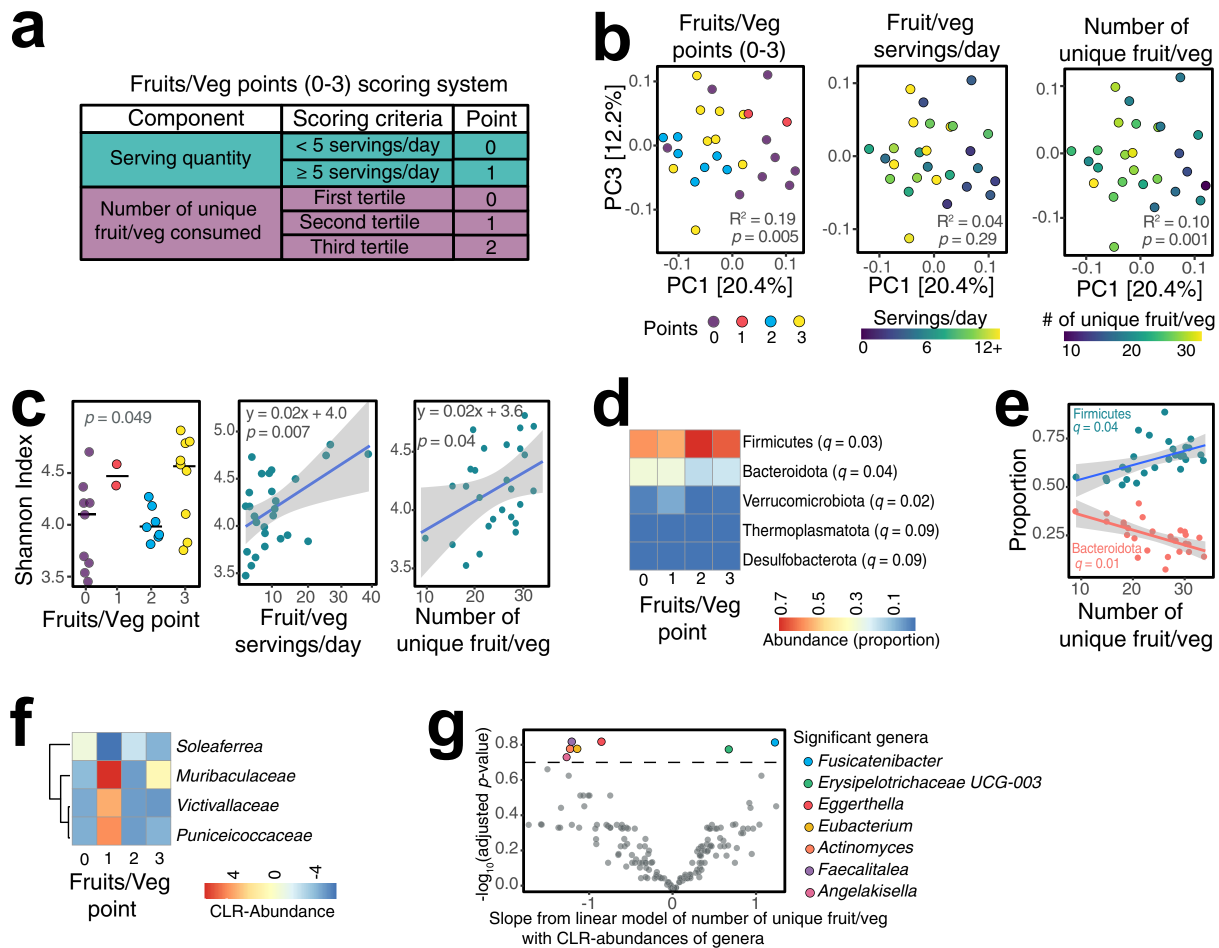


FIGURE 2. Fruit and vegetable intake is associated with gut microbial diversity and composition. (A) The Fruits and Vegetable score is calculated as 1 point for consuming 5 or more servings/d and 0–2 points for tertiles of the number of unique fruits and vegetables usually consumed in a year. (B) Principal coordinate (PC) analysis plots of weighted-UniFrac distance matrices of bacterial community composition colored based on fruits/veg points, servings, and the number of unique fruit/veg consumed. R^2 and P values were extracted from permutational multivariate analysis of variance (PERMANOVA) tests. (C) Alpha diversity analysis using Shannon diversity metric. (D,E) Proportions of phyla that were significantly ($q < 0.2$) correlated with (D) Fruit/veg points and (E) The number of unique fruits and vegetables consumed. (F, G) Centered-log-ratio (CLR) transformed abundances of genera significantly associated with (F) Fruit/veg points and (G) the number of unique fruits and vegetables consumed. Adjusted P values (q) were calculated using Benjamini-Hochberg multiple-testing correction. The line and gray ribbons represent 95% confidence intervals. $N = 27$ participants.

Amplicon sequencing and analysis

DNA was extracted from ethanol-preserved stool samples using the International Human Microbiome Standard operating procedure (Protocol Q) [36], and the 16S ribosomal RNA (rRNA) gene sequencing library was constructed using dual error-correcting barcodes. Briefly, a quantitative primary polymerase chain reaction (PCR) was performed using KAPA HiFi Hot Start Kit (KAPA KK2502) with V4 515F/806R Nextera universal bacterial primers. The amplified products were diluted 1:100 in UltraPure DNase/RNase-free water and were indexed using unique dual indexing primers. The products were quantified using a Quant-iT PicoGreen double-stranded DNA assay kit (Invitrogen P11496) and pooled at equimolar concentrations. The pooled library was quantified via quantitative PCR (qPCR) using the KAPA Library Quantification kit (KAPA KK4824), and its quantity was assessed on 1.5% agarose gel to check the predicted product size. Once quality checks were complete, the library was sequenced on the Illumina MiSeq platform at 270 x 12 x 12 x 270 cycles. Demultiplexed sequences were processed using our 16S rRNA gene analysis pipeline [37]. High-quality reads were analyzed using qiime2R [38] and phyloseq [39] packages in R. Alpha diversity was assessed through the Shannon diversity index using reads subsampled at 40,795 reads per sample (Supplementary Table 1). Analysis of variance (ANOVA) with permutation test (ADONIS function in Vegan package in R) was used to examine the associations between health behavior, sociodemographic and clinical variables, and microbial community structures [40]. Differentially abundant amplicon sequence variants were determined using ANOVA and linear regression models in R.

In our preliminary analysis, we discovered that 1 participant's sample was different from the rest of the samples (Supplementary Figure 1A). Specifically, this sample had lower alpha diversity (Supplementary Figure 1B), with 80% of its community dominated by facultative anaerobes, including *Escherichia-Shigella*, *Streptococcus*, and *Klebsiella* (Supplementary Figure 1C, D). This participant was the only one in the cohort who had an ileostomy (compared with colostomy or no ostomy). Due to its distinct microbiome profile, we excluded this outlier from the remaining analyses described in this manuscript.

qPCR quantification of *Fusobacterium nucleatum*

Extracted genomic DNA was used to quantify *Fusobacterium nucleatum* using a previously published and validated primer pair (F-CGGGTGAGTAACGCGTAAAG, R-GCCGTGTCTCAGTCCCCT) that targets the 16S rRNA gene [41]. 4 μL of DNA, 1 μL of 3μM primer stock, and 5 μL SYBR Select Master Mix for CFX (Thermo Fisher Scientific) were mixed and amplified in triplicate using the following cycling conditions: 50°C for 2 min, 95°C for 2 min, and 40 cycles of 95°C for 15 sec and 60°C for 1 min. DNA extracted from *F. nucleatum* pure culture was used to develop a standard curve. A primer pair targeting general 16S (F-TGGAGCATGTGGTTTAATTCGA, R-TGC GGGACTTAACCCAACA) was also included to calculate the total amount of bacteria present in the sample.

The quantity of DNA in nanogram from *F. nucleatum* culture used to develop a standard curve was converted to copy numbers using the following equation [42]: copy number = (DNA quantity in ng * 6.022 * 10^{23} molecules/mol)/(length of amplicon * 660 g/mole * 10^9 ng/g). The lengths of the amplicon for *Fusobacterium* primer and general 16S primer were 228 bp and 163 bp, respectively. The Cq values from each sample were converted to copy numbers using the semi-log equations generated using the standard curves. Copy number was divided by ng DNA input. Percent abundance was calculated using the following equation: (*Fusobacterium* primer copy number/general 16S primer copy number)*100. Each metric was then tested for associations between covariates using ANOVA for categorical variables and general linear model (GLM) for continuous variables.

Metagenomic sequencing and analysis

300 ng of normalized DNA was used in the Nextera DNA Flex library prep kit along with DNA UD Indexes Set A barcodes (Illumina) to assemble the metagenomic library. ZymoBIOMICS synthetic microbial community standard containing 8 bacteria and 2 yeasts (Zymo Research) was included as a positive control to assess bias and errors in the library preparation. The actual taxonomic composition was correlated with the theoretical composition ($r = 0.83$, $P = 0.003$, Pearson correlation). Each sample was quantified with PicoGreen (ThermoFisher) and checked for quality with TapeStation 4200 (Agilent). Once every sample passed the quality checks, the samples were pooled to the same concentration and sequenced using an S1 flow cell on NovaSeq 6000 system (Illumina) at the Chan Zuckerberg Biohub-San Francisco. The demultiplexed sequences were processed through our metagenomic pipeline, which trims low-quality reads and removes adapters with Fastp [43], removes human reads with Bowtie2 [44], determines genome equivalents with Microbecensus [45], and assigns microbial reads to taxonomic and functional profiles with HUMAnN3 [46].

Of 29.5 million raw sequencing reads per sample on average, 99.0% of reads passed the quality filtering step, yielding 29.2 million high-quality reads per sample (Supplementary Table 1). To account for uneven sequencing depth, reads were normalized to genome equivalents and converted to reads per kilobase per genome equivalent (RPKG). Unstratified gene families and pathways assigned from the HUMAnN3 package [46] were used to analyze the functional composition of the microbiome. Gene or pathway distances were calculated via the Vegan package [40] and visualized in a nonmetric multidimensional scaling of variations format using the Bray-Curtis distance metric. ADONIS permutation test was used to test for association between covariates and the functional composition of the microbiome. Genes or pathways detected in <3 samples or with the labels “unmapped” or “unintegrated” were removed. The abundances were converted to z-scores. The top 2.5% (13,533 gene families) of the most abundant gene families were analyzed for differential abundance using ANOVA and generalized linear models in R.

Statistical analysis

We used descriptive statistics to describe the sociodemographic and clinical characteristics of the study sample, including the median (interquartile range, IQR) for continuous variables and N (%) for categorical variables. To examine the ACS score, its components, and other dietary factors in relation to the microbiome, we modeled the ACS score as a continuous variable; components of the ACS score were examined categorically using the point cut-offs defined above as well as continuously; and the other dietary factors were examined continuously.

Permutational multivariate analysis of variance (PERMANOVA) testing was used to evaluate the association between the health behavior, sociodemographic, and clinical variables and microbial beta diversity in R using the Vegan package [40]. A general linear univariate model and 1-factor ANOVA were used to test univariate associations between continuous variables (for example, number of unique fruit/veg, % whole grain, BMI, fruit/veg servings/d, alcohol g/d, physical activity MET-h/wk, and meat servings/d) and categorical variables (for example, ACS score, ACS score with alcohol, diet sub-score, fruit/veg points, whole grain points, alcohol points, meat points, overall diet points, physical activity points, and BMI points) with microbiome features, respectively. The Benjamini-Hochberg test was used to account for multiple testing [47].

Results

Study population

We contacted 161 CRC survivors by e-mail between 2019 and 2020. Of 45 (28%) individuals that expressed interest in participating in our study, we collected stool samples from 28 patients (Figure 1A; Table 1). One outlier sample was excluded, as described above (see *Methods* and Supplementary Figure S1). The 28 participants had a mean age of 52.7 ± 10.9 y; 57% were male, 79% identified as White, and 14% identified as Hispanic or Latino ethnicity. Nearly half of the participants ($n = 13$; 48%) had an ACS score of 3, with the other participants having ACS scores ranging from 4 to 6; no participants had ACS scores from 0 to 2 (Table 1). There were no significant differences in the assessed demographic, clinical, or dietary factors between the sampled population and the population that was originally invited to participate in this study (Table 1).

For the 28 people who provided stool samples, the median time from diagnosis to completion of the FFQ at stool collection was 4 y (IQR: 3.4–6.5 y). The median time from the FFQ used for analyses to stool collection was 0 mo (IQR: 2 mo prior to 1 mo after stool collection).

Dietary components of the ACS score predict variability in the gut microbiota

16S rRNA sequencing was used to profile microbial community compositions of stool from 27 CRC survivors after excluding 1 outlier sample (Supplementary Figure 1). The overall ACS score was not significantly associated with the gut microbial community structure (Figure 1C). However, 2 components of the ACS score were statistically significant. This included (*i*) fruits and vegetables and (*ii*) alcohol, accounting for 19% and 13% of the variation seen in the fecal microbiota, respectively (Figure 1C). The number of unique fruits and vegetables consumed was also significantly associated with the microbiota when modeled as a continuous variable (Figure 1C). Similar

TABLE 1
Characteristics of invited and sampled CRC survivors in a pilot study of the microbiome, overall and by the ACS Nutrition and Physical Activity Guideline score

	Invited population	Sampled population	Invited vs. sampled population (*P* value)[1]	Sample population by ACS score	
				3–4	5–6
Clinical data					
Number of participants (%)	161 (100)	28 (17)		18 (64)	10 (36)
Age at diagnosis, y; median (IQR)	53 (47, 61)	53 (46, 60)	0.97	52 (45, 60)	56 (48, 58)
Time since diagnosis, y; median (IQR)[2]	3.9 (2.6, 5.6)	4.1 (3.5, 6.5)	0.43	4.3 (3.1, 7.1)	4.1 (3.8, 5.8)
Sex, *n* (%)			0.31		
Male	78 (48)	16 (57)		10 (56)	6 (60)
Female	83 (52)	12 (43)		8 (44)	4 (40)
Cancer site, *n* (%)			0.97		
Colon	103 (64)	18 (64)		13 (72)	5 (50)
Rectum	58 (36)	10 (36)		5 (28)	5 (50)
Stage at diagnosis, *n* (%)			0.80		
Stage I	17 (11)	4 (14)		4 (22)	0
Stage II	42 (26)	5 (18)		2 (11)	3 (30)
Stage III	91 (57)	17 (61)		10 (56)	7 (70)
Stage IV	7 (4)	1 (4)		1 (6)	0
Unknown	4 (2)	1 (4)		1 (6)	0
Treatment history, *n* (%)					
Surgery	152 (94)	26 (93)	0.69	17 (94)	9 (90)
Chemotherapy	39 (24)	8 (29)	0.55	5 (28)	3 (30)
Radiation	112 (70)	21 (75)	0.49	13 (72)	8 (80)
Time since last treatment, y; median (IQR)[3]	3.1 (1.9, 4.8)	3.2 (2.5, 4.4)	0.43	2.9 (1.7, 4.3)	3.4 (3.1, 4.5)
Stoma bag present, *n* (%)[2,3]			0.49		
No	112 (70)	24 (86)		15 (83)	9 (90)
Yes	17 (13)	4 (14)		3 (17)	1 (10)
Missing	5 (4)	0		0	0
Self-reported sociodemographic characteristics					
Number of participants who completed surveys querying demographic factors[4]	134	28		18	10
Race, *n* (%)[3]			0.17		
African American or Black	2 (1)	1 (4)		1 (6)	0
American Indian or Alaskan Native	3 (2)	1 (4)		1 (6)	0
Asian	18 (13)	0		0	0
>1 race	8 (6)	3 (11)		2 (11)	1 (10)
White	96 (72)	22 (79)		13 (72)	9 (90)
Not reported	7 (5)	1 (4)		1 (6)	0
Ethnicity, *n* (%)[3]			0.37		
Hispanic or Latino	11 (8)	4 (14)		3 (17)	1 (10)
Not Hispanic or Latino	122 (91)	24 (86)		15 (83)	9 (90)
Not reported	1 (1)	0		0	0
Education, *n* (%)[3]			0.41		
Grade school	2 (1)	0		0	0
High school or equivalent	9 (7)	2 (7)		1 (6)	1 (10)
Some college	2 (1)	1 (4)		0	1 (10)
Associate's degree	12 (9)	0		0	0
Bachelor's degree	46 (34)	11 (39)		9 (50)	2 (20)
Graduate/professional degree	63 (47)	14 (50)		8 (44)	6 (60)
Smoking status, *n* (%)[3]			0.76		
Never	90 (67)	19 (68)		13 (72)	6 (60)
Past	42 (31)	8 (29)		4 (22)	4 (40)
Current	2 (1)	1 (4)		1 (6)	0
ACS Nutrition and Physical Activity Guideline Score components					
Number of participants who completed surveys for ACS score[4]	124	28		18	10
ACS score, median (IQR)	4 (3, 5)	3.5 (3, 5)	0.30	3 (3, 3)	5 (5, 6)
Body mass index, kg/m^2; median (IQR)	24.8 (21.8, 27.4)	24.4 (23.0, 27.9)	0.67	26.8 (24.0, 29.6)	23.0 (21.5, 23.8)
Total physical activity, MET-h/wk; median (IQR)	30.2 (12.6, 71.3)	27.4 (13.3, 50.2)	0.39	20.0 (6.3, 34.2)	44.7 (26.5, 75.3)
Fruits and vegetables, servings/d; median (IQR)	7.2 (4.8, 10.5)	7.6 (4.0, 11.0)	0.39	7.2 (3.7, 10.0)	9.4 (6.4, 11.6)
Number of unique fruits and vegetables typically consumed in past year, median (IQR)	25 (21, 29)	25 (20, 30)	1.0	24 (18, 28)	28 (21, 30)
Percent of grains consumed that are whole grains, median (IQR)	59 (40, 79)	66 (50, 83)	0.20	64 (47, 86)	68 (56, 79)

(*continued on next page*)

TABLE 1 (*continued*)

	Invited population	Sampled population	Invited vs. sampled population (*P* value)[1]	Sample population by ACS score	
				3–4	5–6
Red and processed meat intake, servings/wk; median (IQR)	6.2 (3.4, 8.9)	6.4 (3.7, 9.7)	0.67	6.5 (5.0, 9.0)	4.4 (1.0, 10.4)
Alcohol, drinks/d; median (IQR)	0.4 (0.1, 1.0)	0.5 (0, 1.3)	0.66	0.1 (0, 1.3)	0.7 (0.4, 1.2)
Fiber, g/d; median (IQR)[5]	27.7 (19.0, 42.4)	31.5 (21.7, 45.2)	0.36	28.1 (21.5, 43.2)	39.3 (21.9, 49.2)

ACS, American Cancer Society; CRC, colorectal cancer; IQR, interquartile range; MET-h/wk, metabolic equivalent task-hours per week.

[1] Chi-square test for categorical measure or nonparametric median comparison test for continuous measure, as appropriate.

[2] For comparability between the invited and sample populations, ostomy status, time since diagnosis, and time since last treatment were measured at the start of the microbiome sub-study (October 2019).

[3] Data on self-reported ostomy status, race, ethnicity, and smoking status were available for 134 (83%) of the invited participants. $N = 27$ (17%) of the invited participants consented but did not complete enrollment in the parent study.

[4] Data on ACS guideline score components are available for 124 (77%) of the invited participants. $N = 10$ of the enrolled participants did not complete components to calculate ACS.

[5] Fiber is not a component in the ACS guideline score but was included here because of its potential impact on the gut microbiome and colorectal cancer.

results were obtained using a different beta diversity metric (Bray-Curtis), suggesting that these findings were robust and independent of the distance metrics used (Supplementary Figure 2A and Supplementary Table 2). These variables were the largest predictors of variability in the microbiome in our sample—more than clinical, demographic, or other behavioral variables. Other components of the ACS score, such as physical activity, BMI, red/processed meat intake, and whole grains consumption, were not associated with the microbiota (Figure 1C). Surprisingly, total fiber intake was not associated with microbial diversity (Figure 1C; Supplementary Figure 3A-B) or phylum-level abundance (Supplementary Figure 3C). Total fiber intake was associated with a differential abundance of 9/171 (5.3%) genera (Supplementary Figure 3D).

Fruit and vegetable consumption is associated with the gut microbiome

Microbial community structure was significantly associated with the total fruit and vegetable points (fruits/veg points) and variety of fruits and vegetables consumed, but not the total quantity (servings/d) (Figure 2B). Microbial diversity was positively correlated with all 3 metrics of fruit and vegetable consumption (fruits/veg points, serving size, and variety; Figure 2C). At the phylum-level, the proportion of Firmicutes increased with higher fruit and vegetable intake, whereas Bacteroidota decreased with more intake (Figure 2D,E). Differences in the abundance of 3 additional phyla were observed with respect to the combined fruits/veg points (Figure 2D). At the genus level, the abundances of 4 genera had a nonlinear association with fruits/veg points (Figure 2F). A GLM testing the association between the variety of fruits and vegetables consumed and the abundances of individual genera revealed 2 positively correlated genera (*Fusicatenibacter* and *Erysipelotrichaceae*) and 5 negatively correlated genera (*Eggerthella*, *Eubacterium*, *Actinomyces*, *Faecalitalea*, *Angelakisella*; Figure 2G).

To test if fruit and vegetable intake also accounted for the observed variations in the functional composition of the gut microbiome, PERMANOVA tests were used with the unstratified gene and pathway abundances. Similar to the taxonomic composition, fruit and vegetable intake accounted for 15% ($P = 0.018$; Figure 3A) and 26% ($P = 0.011$; Figure 3B) of the variations seen in gene and pathway abundances, respectively. Abundances of 21 gene families ($q < 0.2$; Figure 3C and Supplementary Table 3) and 33/375 pathways ($q < 0.2$; Figure 3D and Supplementary Table 4) were significantly associated with fruit and vegetable consumption. The dTDP-N-acetylviosamine biosynthesis pathway was the only pathway with a clear linear relationship with the fruit and vegetable points (Figure 3D).

Overall, fruit and vegetable consumption was associated with pathways for the biosynthesis of certain amino acids, nucleic acids, and short-chain fatty acids (SCFAs) and the degradation of sugars (Figure 3D). Seven out of 33 (21%) of the significant pathways were involved in amino acid biosynthesis, such as arginine, ornithine, glutamine, putrescine, and chorismate, a precursor of aromatic amino acids. Four of the 33 (12%) pathways involved pyrimidine and adenosine biosynthesis. Bifidobacterium shunt, a pathway that breaks down hexose sugar to produce acetate and lactate [48], was significant. However, other pathways involved in the production of SCFAs via fermentation of L-lysine, pyruvate, or acetyl-CoA had no detectable association with fruit and vegetable intake. Degradation of galactose and stachyose, sugars found in avocados and numerous vegetables [49], was also significantly associated with fruit and vegetable consumption.

Variety of fruit and vegetable consumption is associated with lower *Fusobacterium nucleatum*

We examined the abundance of 12 CRC-associated bacterial species [4,6,7,50,51]. Most (10/12) of these species, including members of the *Fusobacterium* genus, were only detected in a minority of metagenomic data sets (0–3, Supplementary Figure 4A). *Escherichia coli* and *Bacteroides fragilis* were more common but were not associated with our clinical metadata (*data not shown*). Next, we sought to increase our sensitivity to detect *F. nucleatum* by leveraging a published qPCR assay [41]. *F. nucleatum* was detected by qPCR in every sample but varied in relative abundance (Supplementary Figure 4B). *F. nucleatum* levels were positively correlated with time since the diagnosis (Supplementary Figure 4C) and time since the last treatment (Supplementary Figure 4D). Interestingly, *F. nucleatum* abundance was negatively correlated with the variety of unique fruits and vegetables consumed (Supplementary Figure 4E).

Alcohol intake is associated with gut microbial species and gene abundance

Intake of alcoholic drinks was associated with interindividual variations in gut microbial community structure ($P = 0.01$; Figure 1C, Figure 4A) and higher microbial diversity ($P = 0.03$; Figure 4B). Alcohol intake negatively correlated with the abundance of the Bacteroidota phylum (Figure 4C). Two other phyla (*Thermoplasmatota* and *Proteobacteria*) and 1 species-level taxon (*Clostridium innocuum*) were also

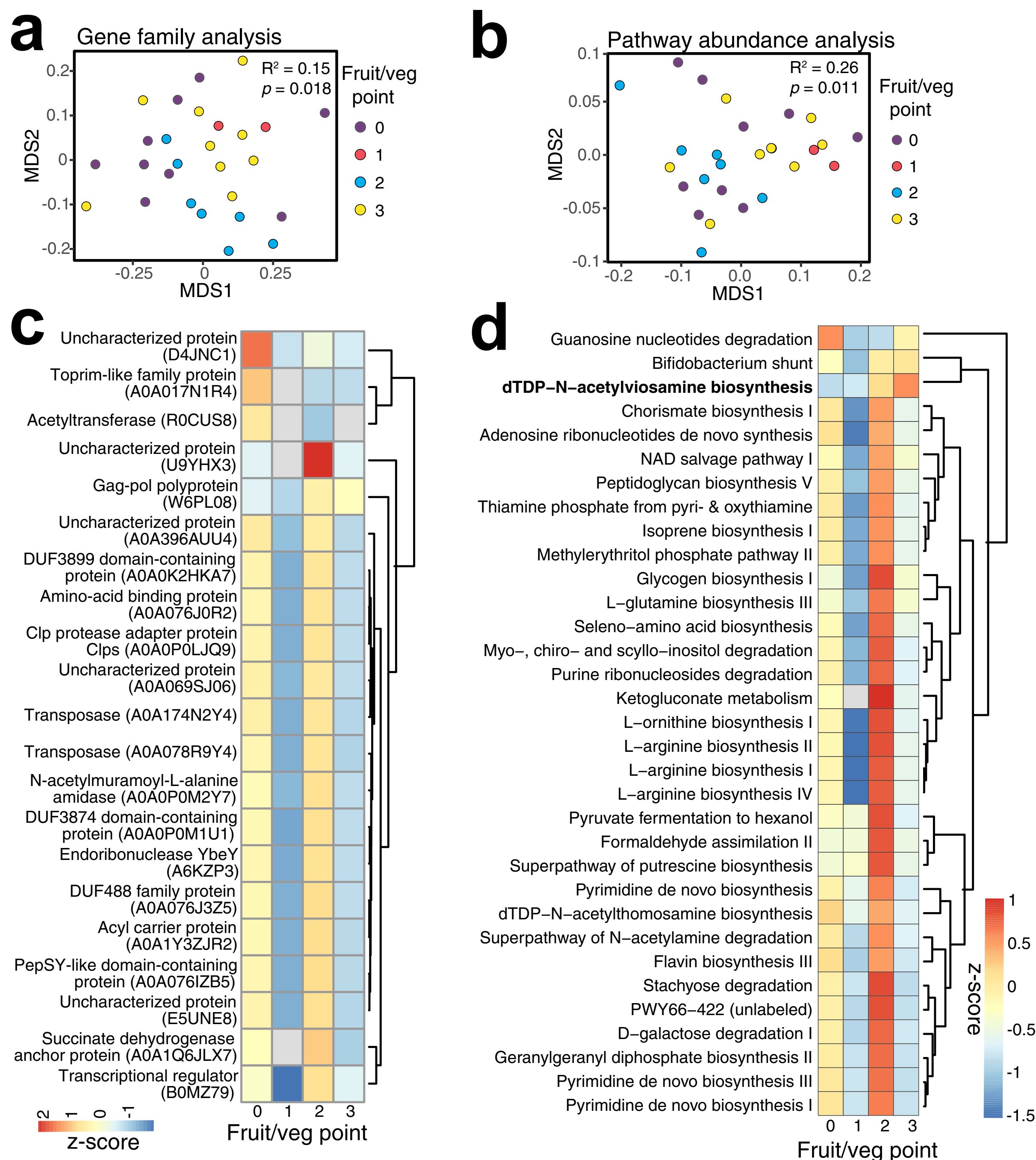


FIGURE 3. Fruit and vegetable intake is associated with microbial gene and pathway abundances. (A, B) Nonmetric multidimensional scaling (MDS) of variations in (A) gene family-level abundances and (B) pathway abundances colored by fruit/vegetable points. R^2 and P values were extracted from permutational multivariate analysis of variance (PERMANOVA) tests. (C,D) Significant (false-discovery-rate $q < 0.2$; ANOVA test with Benjamini-Hochberg correction) (C) Gene families (Supplementary Table 3) and (D) pathways associated with fruit/veg points (Supplementary Table 4). Bolded pathways are significant by general linear model (GLM). Z-score was calculated from average normalized abundances. The identifier in the parentheses refers to gene family entries in the UniRef90 database. $N = 27$ participants.

significantly associated with alcohol use (Figure 4C, D). Notably, alcohol use was not correlated to either the quantity or variety of fruit and vegetable consumption (Supplementary Figure 2B), suggesting they represent independent factors associated with variations in the microbiota.

To test whether alcohol consumption was linked to the observed variations in the functional composition of the gut microbiome, we first assessed microbial gene abundances. We used 2 statistical models: *(i)* ANOVA testing for significant differences in gene abundances among nondrinkers, light-to-moderate drinkers, and heavy drinkers, and *(ii)* a GLM that tests for a linear association between alcohol consumption and gene abundances. Of 13,533 gene families, 1,278 (9.5%) were significant by ANOVA ($q_{ANOVA} < 0.2$), 2,812 (21%) were significant by GLM ($q_{GLM} < 0.2$), and 1,061 (7.8%) were significant by both models (Figure 4E, Supplementary Table 5). Notably, >97% (1,031/1,061) of the significant gene families had negative correlations with alcohol consumption.

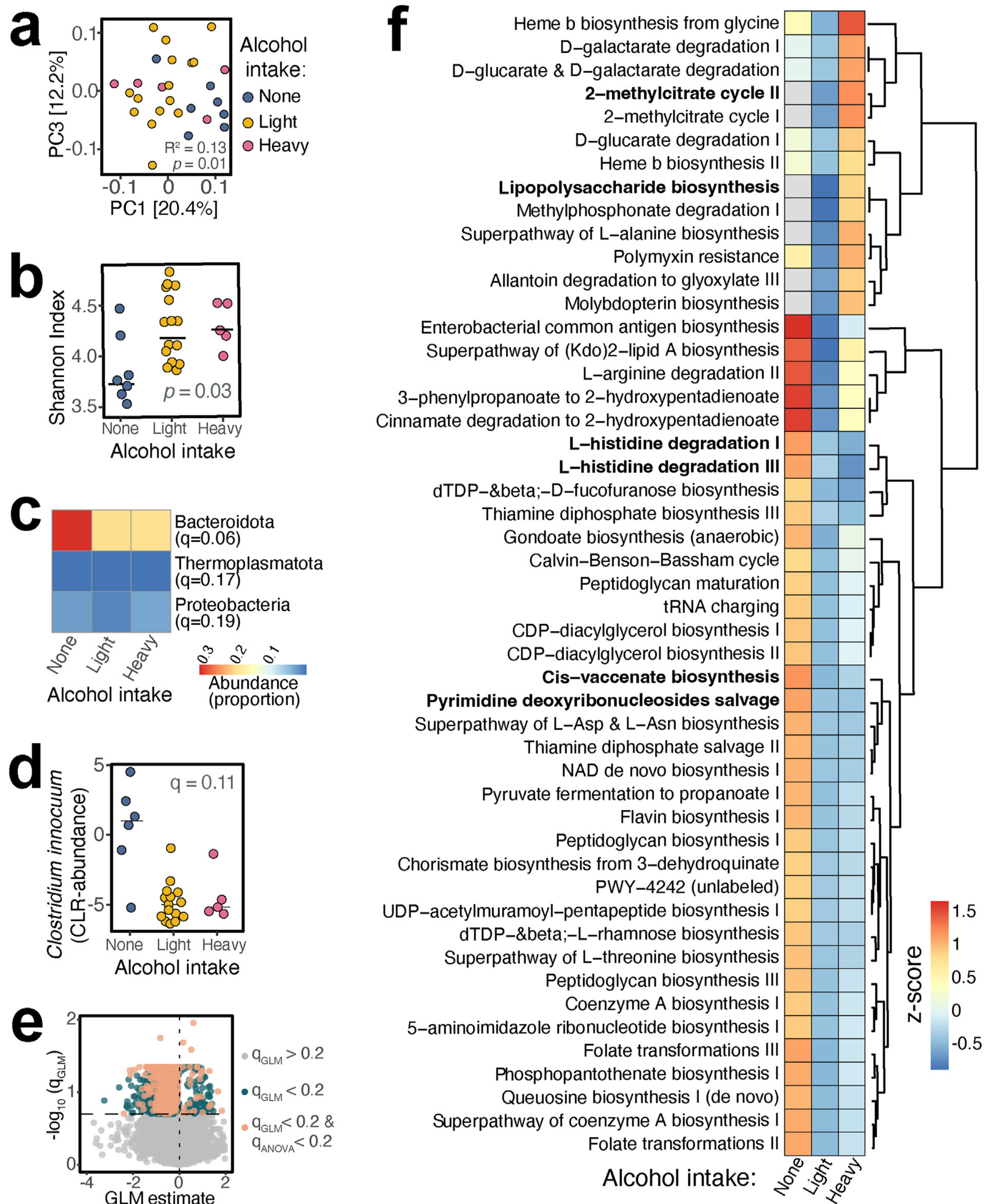


FIGURE 4. Alcohol consumption is associated with the gut microbiome of cancer survivors. (A) Principal coordinate (PC) analysis of weighted-UniFrac distance matrices of bacterial community composition from 16S ribosomal RNA gene sequencing colored based on alcohol intake. Light drinking—defined as >0 and ≤1 drink/d for women, >0 and ≤2 drinks/d for men; amounts above these cut-offs are labeled as heavy drinking. R^2 and P values were extracted from permutational multivariate analysis of variance (PERMANOVA) tests. (B) Alpha diversity analysis using Shannon diversity index of community composition for alcohol intake. (C) The proportion of phylum-level abundances associated with alcohol intake. (D) Centered-log-ratio (CLR)-transformed abundance of a species significantly associated with alcohol intake. (E) A volcano plot of gene families tested for association with alcohol intake using 2 statistical models - a general linear model (GLM) and ANOVA (Supplementary Table 5). GLM tests for the linear associations between normalized gene family-level abundances and alcohol intake. (F) Significant pathways associated with alcohol intake (Supplementary Table 6). Bolded pathways have a significant relationship based on GLM. Significance is defined as false-discovery-rate $q < 0.2$ using Benjamini-Hochberg multiple-testing correction. $N = 27$ participants.

Similar analyses were conducted with the pathway abundances. Of 375 total pathways, 49 (13%) were significant by ANOVA ($q_{ANOVA} < 0.2$), of which 6/49 (12%) had significant associations by GLM ($q_{ANOVA} < 0.2$; $q_{GLM} < 0.2$; Figure 4F and Supplementary Table 6). Two pathways, LPS biosynthesis, and 2-methylcitrate cycle, involved in degrading a common food and drink additive propionate, were positively associated with alcohol intake. In contrast, the abundances of 4 pathways involved in histidine degradation, cis-vacenate fatty acid biosynthesis, and pyrimidine salvage were negatively associated with alcohol use. Taken together, these results showed marked differences in taxonomic and functional compositions of the gut microbiome among CRC survivors linked to alcohol intake.

Discussion

In this pilot study, we explored whether the ACS score or its components were associated with variations in the gut microbiome of 27 CRC survivors. Whereas the overall ACS score was not associated with the microbiome in our sample, we found that fruit and vegetable intake and alcohol consumption were significantly associated with interindividual variations in fecal microbial diversity, community composition, and gene and pathway abundances. In fact, these 2 dietary components were the only factors among other variables, such as clinical history, demographics, and other health behaviors, that reached statistical significance in our data set.

This study supports the feasibility of studying the gut microbiomes of CRC survivors. Despite exclusively recruiting via e-mail a few months before the global COVID-19 pandemic began, we were able to enroll and obtain stool samples remotely from 28 CRC survivors. Importantly, these individuals did not differ from our invited population of CRC survivors in terms of demographic or clinical factors (Table 1). However, incorporating other strategies such as reminder invitations, phone calls, and compensation in future studies may increase the recruitment rate.

All samples, except 1, were primarily composed of Firmicutes and Bacteroidota (Supplementary Figure 1C), which are typically the 2 most abundant phyla found in human gut microbiota [52]. One outlier sample exhibiting a bloom of Proteobacteria and no Bacteroidota was collected from the only patient who had an ileostomy, which is a surgically created opening in the abdominal wall through which the ileum is attached (Supplementary Figure 1C). Based on this data, combined with prior data indicating that ileostomies markedly perturb the gut microbiota [53], we opted to exclude this sample from subsequent analyses.

We observed strong and significant associations between fruit and vegetable consumption and multiple aspects of the gut microbiomes of CRC survivors. Consistent with our previous intervention studies in healthy individuals [10], we found a decreased abundance of Bacteroidota phylum in individuals that consume high concentrations of fruits and vegetables. Surprisingly, it was the number of unique fruits and vegetables consumed, not the quantity or fiber content of the diet, that was linked to variations in the gut microbiota. Fruits and vegetable intake was also linked to a lower abundance of *Fusobacterium nucleatum*, a well-characterized organism with a strong CRC association. These results emphasize the need to broadly consider components of fruits and vegetables, in addition to dietary fiber [54], that may play a role in shaping the gut microbiome [55]. Eating a variety of fruits and vegetables provides diverse nutritional substrates that could influence the growth of many microorganisms; for example, dietary lignans [56] or other polyphenols [57].

The fruit and vegetable score, which factored in the quantity and variety of fruits and vegetables consumed, was associated with amino acids, nucleic acids, cofactor, and SCFA biosynthesis and sugar degradations. Supported by the observation that herbivorous fecal microbiomes have enriched biosynthetic pathways for amino acids compared with their carnivorous counterparts [58], this may suggest that the gut microbes may not be able to extract sufficient amino acids and cofactors from fruits and vegetables. Enriched abundance of pathways for sugar degradations suggests the microbes can utilize galactose and stachyose found in fruits and vegetables. Surprisingly, microbial pathway abundances were nonlinear with respect to fruit and vegetable points, which may suggest a certain combination of quantity and the number of unique fruits and vegetables triggered the expansion of biosynthetic pathways. Future intervention studies would be helpful to determine the reproducibility and mechanistic basis for such nonlinear effects on the gut microbiome, as we have recently demonstrated for ketogenic diets [12].

Heavy alcohol consumption is associated with a higher risk of several cancers, including incident CRC, but the effect of moderate drinking has conflicting reports [28]. Some studies considered moderate drinking to be optimal compared to nondrinking or heavy drinking in terms of CRC mortality, whereas other studies considered nondrinking to be optimal [28]. In this exploratory study, we designated low-to-moderate intake as the reference level to be consistent with our prior work. This scoring approach is also supported by a meta-analysis that concluded low-to-moderate alcohol intake was associated with longer CRC survival [59]. In our study, we found that alcohol drinking was associated with multiple aspects of the gut microbiome. The effect size of the observed associations is surprising, given that ethanol is rapidly absorbed in the stomach and small intestine [60]. We propose that the other components of alcohol-containing drinks (for example, polyphenolic compounds in wine or hops in beer) are a more likely source of the observed differences. However, it is also possible that more indirect mechanisms due to changes in host physiology in response to alcohol somehow impact the microbiome. Future studies in mice aimed at understanding these various factors, coupled with studies of the downstream health effects of alcohol-associated differences in the gut microbiome in humans, are needed.

Our study has multiple limitations. We invited 161 individuals to participate and obtained stool samples from 28 (17%). A strength of our study is that the sampled individuals did not differ in terms of demographic or clinical factors from those who did not provide stool samples. However, the invited and sampled population had limited racial, ethnic, and socioeconomic diversity. Additionally, none of our participants had low ACS scores (0–2 points), which limits the generalizability and may have prevented us from detecting differences in the microbiome between CRC survivors who engage in few health behaviors compared with people who are more adherent to the guidelines. In addition, this was an observational, cross-sectional analysis, and therefore, we cannot rule out the potential for confounding. Future studies will include repeated stool sample collections and measures of health behaviors in more diverse populations of CRC survivors. Due to our small sample size in this exploratory pilot study, we lacked sufficient power to pursue multivariable approaches. Of note, confounding by demographic factors is not likely to explain our observed associations with the variety of fruits and vegetables or alcohol intake given the homogeneity in our population (for example, 89% had a 4-y college degree) and lack of univariate association between these variables (education, race, ethnicity) and the structure of the fecal microbiome in our sample. Nevertheless, as stated above, future paired studies in mice and/or other model organisms will be

critical to identify the causal relationships between the variables assessed here, their mechanistic basis, and their potential relevance to host health and disease.

Despite these limitations, we were able to detect significant associations between components of the ACS score and interindividual variations in human gut microbiome, which is linked to cancer pathogenesis and progression. Continued progress in this area could help to tailor dietary guidelines to an individual's microbiomes or to develop microbiome-based interventions that could assist in the posttreatment recovery and survival of CRC and other cancer patients.

Acknowledgments

We are indebted to the other members of the Turnbaugh laboratory for their helpful suggestions during the preparation of this manuscript. We thank Cecilia Noecker for her advice on computational methods, protocol project manager Paige Steiding for their help with the LOGIC study, and Katherine Pollard for her advice on research design.

Author contributions

The authors' responsibilities were as follows—PJT and ELVB: designed research; SAK, JMC, and KN: provided advice on research design; TSK and VU: conducted research; DS, SP, CEA, KVL, AL, AV, and WK: recruited participants to provide essential materials; ELVB, IT, DS, DG, SAK, and JMC: managed biospecimen and survey data collection and storage; IT and TSK: analyzed data; TSK: wrote the original draft of the manuscript; PJT and ELVB: revised the manuscript and had primary responsibility for the final content; and all authors: read and approved the final manuscript.

Conflict of interest

PJT is a Chan Zuckerberg Biohub–San Francisco Investigator who holds an "Investigators in the Pathogenesis of Infectious Disease" Award from the Burroughs Wellcome Fund and was a Nadia's Gift Foundation Innovator who was supported in part by the Damon Runyon Cancer Research Foundation (DRR-42-16). PJT is on the scientific advisory boards for Pendulum, Seed, and SNIPRbiome. EVB is on the Medical Advisory Board for Fight Colorectal Cancer. CEA served on the scientific advisory board for Pionyr Immunotherapeutics and has received research funding (institution) from Bristol Meyer Squibb, Erasca, Guardant Health, Kura Oncology, Merck, and Novartis. There is no direct overlap with the current study. All other authors have no relevant declarations.

Funding

Funding was provided by NIH (R01HL122593, R01AT011117, R01DK114034, R01AR074500, and R21CA227232 to PJT; F30CA257378 to TSK; R37CA248774-02S1 to ELVB and SP) and the American Cancer Society (IRG-17-180-19 to ELVB).

Data availability

The amplicon sequences and metagenomic sequences generated in this study are publicly available in Genbank at BioProject PRJNA909793. General analysis pipelines are available at https://github.com/turnbaughlab.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajcnut.2023.07.011.

References

[1] Cancer of the Colon and Rectum – Cancer Stat Facts [Internet], 2022. SEER. [cited 2022 Jun 13]. Available from: https://seer.cancer.gov/statfacts/html/colorect.html.

[2] C.L. Rock, C. Thomson, T. Gansler, S.M. Gapstur, M.L. McCullough, A.V. Patel, et al., American Cancer Society guideline for diet and physical activity for cancer prevention, CA, Cancer J. Clin. 70 (2020) 245–271.

[3] E.L. Van Blarigan, C.S. Fuchs, D. Niedzwiecki, S. Zhang, L.B. Saltz, R.J. Mayer, et al., Association of survival with adherence to the American Cancer Society Nutrition and Physical Activity Guidelines for cancer survivors after colon cancer diagnosis: the CALGB 89803/alliance trial, JAMA Oncol 4 (2018) 783–790.

[4] C.M. Dejea, P. Fathi, J.M. Craig, A. Boleij, R. Taddese, A.L. Geis, et al., Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria, Science 359 (2018) 592–597.

[5] L. Chung, E.T. Orberg, A.L. Geis, J.L. Chan, K. Fu, C.E. DeStefano Shields, et al., Bacteroides fragilis toxin coordinates a pro-carcinogenic inflammatory cascade via targeting of colonic epithelial cells, Cell Host Microbe 23 (2018) 421.

[6] A.D. Kostic, E. Chun, L. Robertson, J.N. Glickman, C.A. Gallini, M. Michaud, et al., Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment, Cell Host Microbe 14 (2013) 207–215.

[7] A.M. Thomas, P. Manghi, F. Asnicar, E. Pasolli, F. Armanini, M. Zolfo, et al., Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation, Nat. Med. 25 (2019) 667–678.

[8] D. Rothschild, O. Weissbrod, E. Barkan, A. Kurilshikov, T. Korem, D. Zeevi, et al., Environment dominates over host genetics in shaping human gut microbiota, Nature 555 (2018) 210–215.

[9] D. Zeevi, T. Korem, N. Zmora, D. Israeli, D. Rothschild, A. Weinberger, et al., Personalized nutrition by prediction of glycemic responses, Cell 163 (2015) 1079–1094.

[10] L.A. David, C.F. Maurice, R.N. Carmody, D.B. Gootenberg, J.E. Button, B.E. Wolfe, et al., Diet rapidly and reproducibly alters the human gut microbiome, Nature 505 (2014) 559–563.

[11] R.J. von Schwartzenberg, J.E. Bisanz, S. Lyalina, P. Spanogiannopoulos, Q.Y. Ang, J. Cai, et al., Caloric restriction disrupts the microbiota and colonization resistance, Nature 595 (2021) 272–277.

[12] Q.Y. Ang, M. Alexander, J.C. Newman, Y. Tian, J. Cai, V. Upadhyay, et al., Ketogenic diets alter the gut microbiome resulting in decreased intestinal Th17 cells, Cell 181 (2020) 1263–1275.e16.

[13] A.D. Kostic, D. Gevers, C.S. Pedamallu, M. Michaud, F. Duke, A.M. Earl, et al., Genomic analysis identifies association of Fusobacterium with colorectal carcinoma, Genome Res 22 (2012) 292–298.

[14] Y. Cao, J. Oh, M. Xue, W.J. Huh, J. Wang, J.A. Gonzalez-Hernandez, et al., Commensal microbiota from patients with inflammatory bowel disease produce genotoxic metabolites, Science 378 (2022) eabm3233.

15 M.R. Wilson, Y. Jiang, P.W. Villalta, A. Stornetta, P.D. Boudreau, A. Carrá, et al., The human gut bacterial genotoxin colibactin alkylates DNA, Science 363 (2019) eaar7785.

[16] P. Spanogiannopoulos, T.S. Kyaw, B.G.H. Guthrie, P.H. Bradley, J.V. Lee, J. Melamed, et al., Host and gut bacteria share metabolic pathways for anti-cancer drug metabolism, Nat. Microbiol. 7 (2022) 1605–1620.

[17] T.S. Kyaw, M. Sandy, K. Trepka, J.J.N. Goh, K. Yu, V. Dimassa, et al., Human gut actinobacteria boost drug absorption by secreting P-glycoprotein ATPase inhibitors. bioRXiv. Preprint. https://doi.org/10.1101/2022.10.13.512142.

[18] L.T. Geller, M. Barzily-Rokni, T. Danino, O.H. Jonas, N. Shental, D. Nejman, et al., Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine, Science 357 (2017) 1156–1160.

[19] B.D. Wallace, H. Wang, K.T. Lane, J.E. Scott, J. Orans, J.S. Koo, et al., Alleviating cancer drug toxicity by inhibiting a bacterial enzyme, Science 330 (2010) 831–835.

[20] A. Balhareth, M.Y. Aldossary, D. McNamara, Impact of physical activity and diet on colorectal cancer survivors' quality of life: a systematic review, World J. Surg. Oncol. 17 (2019) 153.

[21] S. Piawah, E.J. Walker, E.L. Van Blarigan, C.E. Atreya, The gut microbiome in colorectal cancer, Hematol. Oncol. Clin. North Am. 36 (2022) 491–506.

[22] P.A. Harris, R. Taylor, B.L. Minor, V. Elliott, M. Fernandez, L. O'Neal, et al., The REDCap consortium: building an international community of software platform partners, J. Biomed. Inform. 95 (2019) 103208.

[23] E.B. Rimm, E.L. Giovannucci, M.J. Stampfer, G.A. Colditz, L.B. Litin, W.C. Willett, Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals, Am. J. Epidemiol. 135 (1992) 1114–1126.

[24] C. Yuan, D. Spiegelman, E.B. Rimm, B.A. Rosner, M.J. Stampfer, J.B. Barnett, et al., Relative validity of nutrient intakes assessed by questionnaire, 24-hour recalls, and diet records as compared with urinary recovery and plasma concentration biomarkers: findings for women, Am. J. Epidemiol. 187 (2018) 1051–1063.

[25] L. Al-Shaar, C. Yuan, B. Rosner, S.B. Dean, K.L. Ivey, C.M. Clowry, et al., Reproducibility and validity of a semiquantitative food frequency questionnaire in men assessed by multiple methods, Am. J. Epidemiol. 190 (2021) 1122–1132.

[26] C.H. Pernar, A.K. Chomistek, J.B. Barnett, K. Ivey, L. Al-Shaar, S.B. Roberts, et al., Validity and relative validity of alternative methods of assessing physical activity in epidemiologic studies: findings from the men's lifestyle validation study, Am. J. Epidemiol. 191 (2022) 1307–1322.

[27] C.A. Thomson, M.L. McCullough, B.C. Wertheim, R.T. Chlebowski, M.E. Martinez, M.L. Stefanick, et al., Nutrition and physical activity cancer prevention guidelines, cancer risk, and mortality in the women's health initiative, Cancer Prev. Res. (Phila). 7 (2014) 42–53.

[28] M.L. McCullough, A.V. Patel, L.H. Kushi, R. Patel, W.C. Willett, C. Doyle, et al., Following cancer prevention guidelines reduces risk of cancer, cardiovascular disease, and all-cause mortality, Cancer Epidemiol. Biomarkers Prev. 20 (2011) 1089–1097.

[29] M. Song, K. Wu, J.A. Meyerhardt, S. Ogino, M. Wang, C.S. Fuchs, et al., Fiber intake and survival after colorectal cancer diagnosis, JAMA Oncol 4 (2018) 71–79.

[30] J.A. Meyerhardt, K. Sato, D. Niedzwiecki, C. Ye, L.B. Saltz, R.J. Mayer, et al., Dietary glycemic load and cancer recurrence and survival in patients with stage III colon cancer: findings from CALGB 89803, J. Natl. Cancer Inst. 104 (2012) 1702–1711.

[31] E.L. Van Blarigan, C.S. Fuchs, D. Niedzwiecki, X. Ye, S. Zhang, M. Song, et al., Marine ω-3 polyunsaturated fatty acid and fish intake after colon cancer diagnosis and survival: CALGB 89803 (Alliance), Cancer Epidemiol, Biomarkers Prev 27 (2018) 438–445.

[32] K. Ng, H.S. Nimeiri, N.J. McCleary, T.A. Abrams, M.B. Yurgelun, J.M. Cleary, et al., Effect of high-dose vs standard-dose vitamin D3 supplementation on progression-free survival among patients with advanced or metastatic colorectal cancer: the SUNSHINE randomized clinical trial, JAMA 321 (2019) 1370–1379.

[33] Y. Hu, M. Ding, C. Yuan, K. Wu, S.A. Smith-Warner, F.B. Hu, et al., association between coffee intake after diagnosis of colorectal cancer and reduced mortality, Gastroenterology 154 (2018) 916–926.e9.

[34] B.J. Guercio, K. Sato, D. Niedzwiecki, X. Ye, L.B. Saltz, R.J. Mayer, et al., coffee intake, recurrence, and mortality in stage III colon cancer: results from CALGB 89803 (Alliance), J. Clin. Oncol. 33 (2015) 3598–3607.

[35] T. Fadelu, S. Zhang, D. Niedzwiecki, X. Ye, L.B. Saltz, R.J. Mayer, et al., Nut consumption and survival in patients with stage III colon cancer: results from CALGB 89803 (Alliance), J. Clin. Oncol. 36 (2018) 1112–1120.

[36] P.I. Costea, G. Zeller, S. Sunagawa, E. Pelletier, A. Alberti, F. Levenez, et al., Towards standards for human fecal sample processing in metagenomic studies, Nat Biotechnol 35 (11) (2017) 1069–1076.

[37] J. Bisanz, AmpliconSeq [Internet], April 20, 2021. Github; [cited 2022 May 31]. Available from: https://github.com/jbisanz/AmpliconSeq.

[38] J. Bisanz, qiime2R: Import qiime2 artifacts to R [Internet], Nov 11, 2021. Github; [cited 2022 May 31]. Available from: https://github.com/jbisanz/qiime2R.

[39] P.J. McMurdie, S. Holmes, phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data, PLoS ONE 8 (4) (2013) e61217. http://dx.plos.org/10.1371/journal.pone.0061217.

[40] J. Oksanen, Vegan: community ecology package [Internet]. ci.nii.ac.jp, R package version 1.8-5 (2007) [cited 31 May 2022]. Available from: https://ci.nii.ac.jp/naid/10020010631/.

[41] K.S. Viljoen, A. Dakshinamurthy, P. Goldberg, J.M. Blackburn, Quantitative profiling of colorectal cancer-associated bacteria reveals associations between Fusobacterium spp., enterotoxigenic Bacteroides fragilis (ETBF) and clinicopathological features of colorectal cancer, PLOS ONE 10 (2015) e0119462.

[42] H. Ma, K.N. Bell, R.N. Loker, qPCR and qRT-PCR analysis: regulatory points to consider when conducting biodistribution and vector shedding studies, Mol. Ther. Methods Clin. Dev. 20 (2021) 152–168.

[43] S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor, Bioinformatics 34 (2018) i884–i890.

[44] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat. Methods. 9 (2012) 357–359.

[45] S. Nayfach, K.S. Pollard, Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome, Genome Biol 16 (2015) 51.

[46] E.A. Franzosa, L.J. McIver, G. Rahnavard, L.R. Thompson, M. Schirmer, G. Weingart, et al., Species-level functional profiling of metagenomes and metatranscriptomes, Nat. Methods. 15 (2018) 962–968.

[47] J.A. Ferreira, The Benjamini-Hochberg method in the case of discrete test statistics, Int. J. Biostat. 3 (2007) 11.

[48] W. de Vries, A.H. Stouthamer, Pathway of glucose fermentation in relation to the taxonomy of bifidobacteria, J. Bacteriol. 93 (1967) 574–576.

[49] J.G. Muir, R. Rose, O. Rosella, K. Liels, J.S. Barrett, S.J. Shepherd, et al., Measurement of short-chain carbohydrates in common Australian vegetables and fruits by high-performance liquid chromatography (HPLC), J. Agric. Food Chem. 57 (2009) 554–565.

[50] L. Chung, E. Thiele Orberg, A.L. Geis, J.L. Chan, K. Fu, C.E. DeStefano Shields, et al., Bacteroides fragilis toxin coordinates a pro-carcinogenic inflammatory cascade via targeting of colonic epithelial cells, Cell Host Microbe 23 (2018) 203–214.e5.

[51] D. Ternes, J. Karta, M. Tsenkova, P. Wilmes, S. Haan, E. Letellier, Microbiome in colorectal cancer: how to get from meta-omics to mechanism? Trends Microbiol 28 (2020) 401–423.

[52] The Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome, Nature 486 (2012) 207–214.

[53] A.L. Hartman, D.M. Lough, D.K. Barupal, O. Fiehn, T. Fishbein, M. Zasloff, et al., Human gut microbiome adopts an alternative state following small bowel transplantation, Proc. Natl. Acad Sci. U S A 106 (2009) 17187–17192.

[54] M.C.W. Myhrstad, H. Tunsjø, C. Charnock, V.H. Telle-Hansen, Dietary fiber, gut microbiota, and metabolic regulation-current status in human randomized trials, Nutrients 12 (2020) 859.

[55] R.N. Carmody, J.E. Bisanz, B.P. Bowen, C.F. Maurice, S. Lyalina, K.B. Louie, et al., Cooking shapes the structure and function of the gut microbiome, Nat. Microbiol. 4 (2019) 2052–2063.

[56] E.N. Bess, J.E. Bisanz, F. Yarza, A. Bustion, B.E. Rich, X. Li, et al., Genetic basis for the cooperative bioactivation of plant lignans by Eggerthella lenta and other human gut bacteria, Nat. Microbiol. 5 (2020) 56–66.

[57] L. Zhang, R.N. Carmody, H.M. Kalariya, R.M. Duran, K. Moskal, A. Poulev, et al., Grape proanthocyanidin-induced intestinal bloom of Akkermansia muciniphila is dependent on its baseline abundance and precedes activation of host genes related to metabolic health, J. Nutr. Biochem. 56 (2018) 142–151.

[58] B.D. Muegge, J. Kuczynski, D. Knights, J.C. Clemente, A. González, L. Fontana, et al., Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans, Science 332 (2011) 970–974.

[59] Y. Kim, Y. Je, E.L. Giovannucci, Association between alcohol consumption and survival in colorectal cancer: a meta-analysis, Cancer Epidemiol. Biomarkers Prev. 28 (2019) 1891–1901.

[60] K. Pohl, P. Moodley, A.D. Dhanda, Alcohol's impact on the gut and liver, Nutrients 13 (2021) 3170.