



Association of dietary fibre intake and gut microbiota in adults

Daniel Lin¹, Brandilyn A. Peters², Charles Friedlander^{3,4}, Hal J. Freiman⁴, James J. Goedert⁵, Rashmi Sinha⁵, George Miller^{1,6}, Mitchell A. Bernstein⁶, Richard B. Hayes^{1,2} and Jiyoung Ahn^{1,2*}

¹Perlmutter Cancer Center, New York University Langone Medical Center, New York, NY 10016, USA

²Department of Population Health, New York University School of Medicine, New York, NY 10016, USA

³Concorde Medical Group, New York, NY 10016, USA

⁴Department of Medicine, New York University School of Medicine, New York, NY 10016, USA

⁵Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20850, USA

⁶Department of Surgery, New York University School of Medicine, New York, NY 10016, USA

(Submitted 3 January 2018 – Final revision received 26 July 2018 – Accepted 7 August 2018)

Abstract

Increasing evidence indicates that gut microbiota may influence colorectal cancer risk. Diet, particularly fibre intake, may modify gut microbiota composition, which may affect cancer risk. We investigated the relationship between dietary fibre intake and gut microbiota in adults. Using 16S rRNA gene sequencing, we assessed gut microbiota in faecal samples from 151 adults in two independent study populations: National Cancer Institute (NCI), n 75, and New York University (NYU), n 76. We calculated energy-adjusted fibre intake based on FFQ. For each study population with adjustment for age, sex, race, BMI and smoking, we evaluated the relationship between fibre intake and gut microbiota community composition and taxon abundance. Total fibre intake was significantly associated with overall microbial community composition in NYU ($P=0.008$) but not in NCI ($P=0.81$). In a meta-analysis of both study populations, higher fibre intake tended to be associated with genera of class Clostridia, including higher abundance of *SMB53* (fold change (FC) = 1.04, $P=0.04$), *Lachnospira* (FC = 1.03, $P=0.05$) and *Faecalibacterium* (FC = 1.03, $P=0.06$), and lower abundance of *Actinomyces* (FC = 0.95, $P=0.002$), *Odoribacter* (FC = 0.95, $P=0.03$) and *Oscillospira* (FC = 0.96, $P=0.06$). A species-level meta-analysis showed that higher fibre intake was marginally associated with greater abundance of *Faecalibacterium prausnitzii* (FC = 1.03, $P=0.07$) and lower abundance of *Eubacterium dolichum* (FC = 0.96, $P=0.04$) and *Bacteroides uniformis* (FC = 0.97, $P=0.05$). Thus, dietary fibre intake may impact gut microbiota composition, particularly class Clostridia, and may favour putatively beneficial bacteria such as *F. prausnitzii*. These findings warrant further understanding of diet-microbiota relationships for future development of colorectal cancer prevention strategies.

Key words: Gut microbiome; Dietary fibre intake; Cross-sectional studies; Epidemiology

The human gastrointestinal tract hosts an estimated 100 trillion bacteria, which play a role in key physiological activities, including gastrointestinal immune system maturation and fermentation of nutrients into beneficial metabolites⁽¹⁾. Disruption of this symbiotic relationship between human host and gut microbiota has been implicated in the development of the metabolic syndrome, as well as intestinal pathology, including inflammatory bowel disease and colorectal cancer^(2–4). A growing number of epidemiological studies have provided increasing evidence that perturbation of microbial community composition in the gut exists in colorectal cancer, with alterations of microbial taxa abundance in cancer cases compared with healthy controls^(5–8).

Colorectal cancer risk has been attributed to dietary habits, with Western-style diets – low in fibre and high in red meat and

fat – associated with higher risk for colorectal cancer^(9,10). Fibre intake, in particular, has remained an appealing modifiable dietary factor, given its postulated beneficial biological effects. Fibre speeds colonic transit and may decrease exposure of colonic epithelial cells to ingested carcinogens. In addition, fibre undergoes fermentation by the microbiota to yield SCFA end products, such as butyrate, which is not only essential for colon energy metabolism and epithelial proliferation, but in mouse models also exhibits tumour-suppressive activity through histone deacetylase inhibition⁽¹¹⁾.

Consequently, there has been growing interest to understand the impact of dietary fibre on gut microbiota composition, which may ultimately reduce cancer risk. A diet low in fibre – the ‘fibre gap’ – has been associated with a depletion in human gut microbiota diversity, and potentially beneficial

Abbreviations: FC, fold change; NCI, National Cancer Institute; NYU, New York University.

* **Corresponding author:** J. Ahn, email Jiyoung.Ahn@nyumc.org



metabolites⁽⁴⁾. Furthermore, although short-term dietary intervention trials have demonstrated that different amounts of fibre intake can significantly alter microbiota composition in a span of a few weeks^(12–14), there remain fewer studies evaluating the effect of long-term dietary habits of fibre intake on the gut microbiota in humans⁽¹⁵⁾.

We investigated the association between long-term dietary fibre intake habits and gut microbiota composition in faecal samples of generally healthy adults from two independent study populations: controls from a hospital-based case-control study of colorectal cancer and gut microbiome⁽⁶⁾ and polyp-free adults from a cross-sectional colonoscopy study⁽¹⁶⁾. We sought to examine the relationship between higher dietary fibre intake and overall gut microbiota composition, as well as specific taxa abundance.

Methods

Study population

We assessed faecal samples of adults from two independent study populations: control subjects from a National Cancer Institute (NCI) hospital-based case-control study, hereafter referred to as NCI⁽⁶⁾, and polyp-free adults from a cross-sectional colonoscopy study at New York University (NYU), called the NYU Human Microbiome and Colorectal Tumor study, hereafter referred to as NYU⁽¹⁶⁾.

NCI enrolled participants from three Washington, DC area hospitals from 1985 to 1989^(17,18). We included seventy-five control subjects who were awaiting elective surgery for non-oncological, non-gastrointestinal conditions, and reported no antibiotic intake during the year before recruitment. Participants provided 2-d faecal samples that were freeze-dried, and samples with at least 100 mg of lyophilised faecal material available were included for analyses.

NYU enrolled participants from Kips Bay Endoscopy Center in New York City from 2012 to 2014. We included seventy-six polyp-free participants from this study. We excluded subjects with missing colonoscopy reports, history of inflammatory bowel disease, prior surgical anastomosis, prior history of colorectal cancer, history of familial adenomatous polyposis, and with most recent colonoscopy report >3 years before stool sample collection.

In both studies, participants provided written informed consent, reported no long-term antibiotic treatment and completed diet and demographic questionnaires. We excluded subjects with <1000 microbial sequence reads, missing or extreme energetic intake (≤ 2092 or $> 16\ 736$ kJ/d) and with a history of other cancers, for a final sample size of 151 ($n = 75$ in NCI, $n = 76$ in NYU). The NYU study was approved by the institutional review board (IRB) of NYU School of Medicine, and the NCI study was approved by the IRB of NYU School of Medicine and the NCI.

Demographic information and dietary fibre assessment

Information on age, sex, height, weight, race and cigarette smoking status was collected by means of a questionnaire at stool collection. BMI was calculated by dividing weight in kg by squared height in m, and was then categorised as underweight or

normal weight ($< 25 \text{ kg/m}^2$), overweight ($25 \leq \text{BMI} < 30 \text{ kg/m}^2$) or obese ($\text{BMI} \geq 30 \text{ kg/m}^2$), based on WHO definition⁽¹⁹⁾. Cigarette smoking status was defined as never, current or former smoker.

Usual dietary intake was calculated from self-administered FFQ, which queried intake frequency and portion size of food types. Nutrient values per portion were multiplied by daily frequency of intake and summed across all relevant food items. Nutrient data were standardised by total energy intake⁽²⁰⁾. Study-specific quartiles of total fibre intake were used (NCI: <11·3, 11·3–14·3, 14·4–16·6, $\geq 16\cdot7 \text{ g/d}$; NYU: <19·8, 19·8–23·7, 23·8–31·6, $\geq 31\cdot7 \text{ g/d}$). In addition, study-specific quartiles of fibre from different dietary sources were assessed: fibre from fruits and vegetables (NCI: <6·5, 6·5–8·7, 8·8–11·0, $\geq 11\cdot1 \text{ g/d}$; NYU: <11·1, 11·1–14·9, 15·0–21·2, $\geq 21\cdot3 \text{ g/d}$), fibre from grains (NCI: <2·7, 2·7–3·2, 3·3–4·8, $\geq 4\cdot9 \text{ g/d}$; NYU: <9·1, 9·1–12·6, 12·7–16·6, $\geq 16\cdot7 \text{ g/d}$) and fibre from beans (NCI: <0·9, 0·9–1·3, 1·4–2·1, $\geq 2\cdot2 \text{ g/d}$; NYU: <1·8, 1·8–2·7, 2·8–4·4, $\geq 4\cdot5 \text{ g/d}$).

Faecal samples

In NCI, faecal samples were collected by participants at home over a 2-d period, before hospitalisation and treatment, and stored in a plastic container in a styrofoam chest containing dry ice. Faecal samples were shipped to a USDA laboratory, lyophilised and stored at a minimum of -40°C in sealed, air-tight containers. Sample aliquots were shipped to NYU for microbiome assay. In NYU, faecal samples were collected by participants onto two sections of Beckman Coulter Hemoccult II SENSA® cards (Beckman Coulter) at home. Samples were shipped to NYU and stored immediately at -80°C .

Microbiota assay

In both NCI and NYU, DNA was extracted from faecal samples using the Mo Bio PowerSoil DNA Isolation Kit (Mo Bio) with bead-beating, as previously reported⁽¹⁶⁾. In NCI, 16S ribosomal RNA (rRNA) gene amplicons covering variable regions V3–V4 were generated using the 347F-5'GG AGGCAGCAGTRRGGAAT'-3' and 803R 5'-CTACCRGGGTATCTAATCC-3' primer pair^(6,21). Amplicons were sequenced with the 454 Roche FLX Titanium pyrosequencing system, according to the manufacturer's protocol. In NYU, 16S rRNA gene amplicons covering the V4 region were generated using the F515-5'GTGCCAGCMGCCGCGTAA'-3' and R806-5'GGACTACHVGGGTWTCTAAT-3' primer pair⁽²²⁾. Amplicons were sequenced with the Illumina MiSeq platform.

Sequence data processing

Because two different sequencing platforms were used, we processed the sequence data separately. Sequences were demultiplexed, and poor-quality sequences excluded, using the default parameters of Quantitative Insights into Microbial Ecology (QIIME) script *split_libraries.py* (for NCI) or *split_libraries_fastq.py* (for NYU)⁽²³⁾. Filtered sequence reads were clustered into *de novo* operational taxonomic units (OTU) at 97% identity, and representative sequence reads for each OTU were assigned taxonomy based on fully sequenced microbial genomes (Integrated Microbial Genomes (IMG)/Greengenes



(GG))⁽²³⁾. Chimeric sequences were removed with ChimeraSlayer⁽²⁴⁾. Blinded quality control specimens in all sequencing batches showed good reproducibility: high intraclass correlation coefficients for the Shannon diversity index and abundances of bacterial phyla and genera have been previously reported^(6,16,25).

Statistical analysis

α -Diversity. We evaluated the association between quartiles of fibre intake and within-subject microbial diversity (α -diversity) indices of Shannon diversity and evenness⁽²⁶⁾. In both studies, these indices were calculated in 100 iterations of rarefied OTU tables of 1000 sequence reads per sample. We modelled the Shannon index and evenness as outcomes in linear regression, adjusting for age, sex, race, categorical BMI and cigarette smoking status.

β -Diversity. We assessed the relationship of overall gut microbiota composition and quartiles of dietary fibre intake using weighted (quantitative) and unweighted (qualitative) phylogenetic UniFrac distance matrices⁽²⁷⁾. Permutational multivariate ANOVA (PERMANOVA) of both weighted and unweighted UniFrac distances was used to evaluate whether fibre intake is associated with overall microbial community composition, after adjusting first for age, sex, race, categorical BMI and cigarette smoking status (adonis function, ‘vegan’ package in R)⁽²⁸⁾. Principal coordinate analysis plots were generated using the first two principal coordinates, and labelled according to quartile of fibre intake.

Differential abundance testing. We assessed the relationship between higher quartiles of total fibre intake and specific taxa abundance using negative binomial generalised linear models, in the ‘DESeq2’ package in R⁽²⁵⁾. Models were adjusted for age, sex, race, categorical BMI and cigarette smoking status.

Table 1. Participant demographic characteristics
(Quartiles (Q) and medians)

	NCI (n 75)					NYU (n 76)				
	Q1*	Q2*	Q3*	Q4*	P†	Q1‡	Q2‡	Q3‡	Q4‡	P‡
n	19	19	18	19		19	19	19	19	
Age (years)					0.11					0.34
Median	50	55	61.5	62		56	57	62	61	
Sex (%)					0.08					0.16
Female	31.6	21.1	50.0	57.9		36.8	36.8	52.6	68.4	
Male	68.4	78.9	50.0	42.1		63.2	63.2	47.4	31.6	
Race (%)					0.46					0.10
White	89.5	89.5	77.8	73.7		94.7	89.5	89.5	68.4	
Non-white	10.5	10.5	22.2	26.3		5.3	10.5	10.5	31.6	
BMI (kg/m ²) (%)					0.19					
<25	57.9	68.4	33.3	68.4		36.8	63.2	57.9	63.2	0.29
25–30	36.8	15.8	38.9	21.1		42.1	31.6	26.3	36.8	
≥30	5.3	15.8	27.8	10.5		21.1	5.3	15.8	–	
Smoking history (%)					0.10					0.13
Never	26.3	36.8	61.1	57.9		63.2	84.2	47.4	63.2	
Former/current	73.7	63.2	38.9	42.1		36.8	15.8	52.6	36.8	

NCI, National Cancer Institute; NYU, New York University.

* NCI quartiles of fibre intake: Q1: <11.3 g; Q2: 11.3–14.3 g; Q3: 14.4–16.6 g; Q4: ≥16.7 g.

† P values are calculated from Kruskal–Wallis test (continuous variables) or χ^2 test (categorical variables).

‡ NYU quartiles of fibre intake: Q1: <19.8 g; Q2: 19.8–23.7 g; Q3: 23.8–31.6 g; Q4: ≥31.7 g.

Nominal P values and false discovery rate (FDR) adjusted q -values were calculated⁽²⁹⁾. DESeq2 default outlier replacement, independent filtering of low-count taxa and filtering of count outliers were turned off. Taxa models with maximum Cook’s distance >10 were removed before P value adjustment for the FDR⁽³⁰⁾.

To identify similar taxa associations in both NCI and NYU, we then performed a taxonomy-based meta-analysis to evaluate for genera and species with higher or lower abundance by fibre intake concomitantly in each of the two study populations. In addition, we performed sub-analyses to examine associations between taxon abundance and higher intake of fibre from specific sources, such as fruit and vegetables, grains and beans. We calculated nominal meta-analysis P values based on Z-score methods⁽³¹⁾.

All trends were tested using the median values of each quartile of fibre intake. All analyses were performed using R, version 3.3.2.

Results

Subject characteristics

A total of 151 adult subjects were included for analysis: n 75 in NCI and n 76 in NYU. Subject characteristics are reported in Table 1. Among these participants, 73.3% in NCI and 51.3% in NYU were male, and 82.7% in NCI and 85.5% in NYU were white. Median age in both study groups generally increased with higher fibre intake, although the trends were not statistically significant.

Global diversity

PERMANOVA analyses of between-sample UniFrac distances demonstrated that fibre intake was significantly associated with overall microbial community composition in NYU (weighted

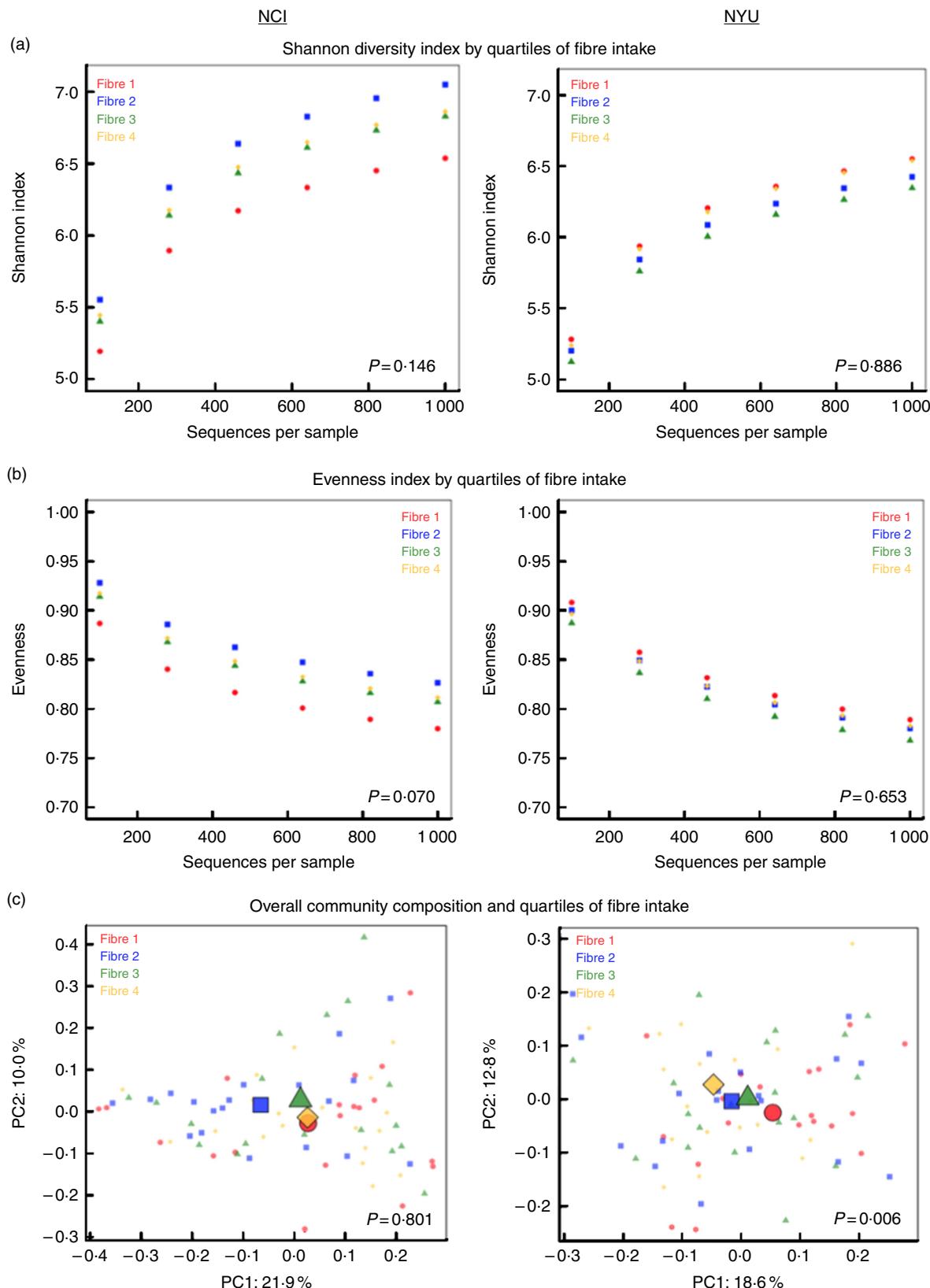


Fig. 1. α and β diversity in relation to quartiles of total fibre intake (a) Shannon diversity index and (b) evenness index by quartiles of total fibre intake are shown in n 151 healthy adult subjects from two independent study populations (National Cancer Institute (NCI) = 75, New York University (NYU) = 76). These indices were calculated in 100 iterations of rarefied operational taxonomic units tables of 1000 sequence reads per sample. Fibre 1, fibre 2, fibre 3 and fibre 4 represent increasing quartiles of fibre intake. Shannon index and evenness were modelled as outcomes in linear regression, adjusting for age, sex, race, categorical BMI and cigarette smoking status. P values of fibre variable in regression analysis are reported in the figure. (c) Principal coordinate analysis (PCoA) plots, based on weighted UniFrac phylogenetic distances, showed a difference between lowest and highest fibre intake in NYU. This relationship was not observed in NCI. PCoA plots were generated using the first two principal coordinates (PC): PC1 and PC2. P values reported in the figure are based on permutational multivariate ANOVA of weighted UniFrac distances evaluating the association between fibre intake and overall microbial community composition, after adjusting for age, sex, race, categorical BMI and cigarette smoking status.

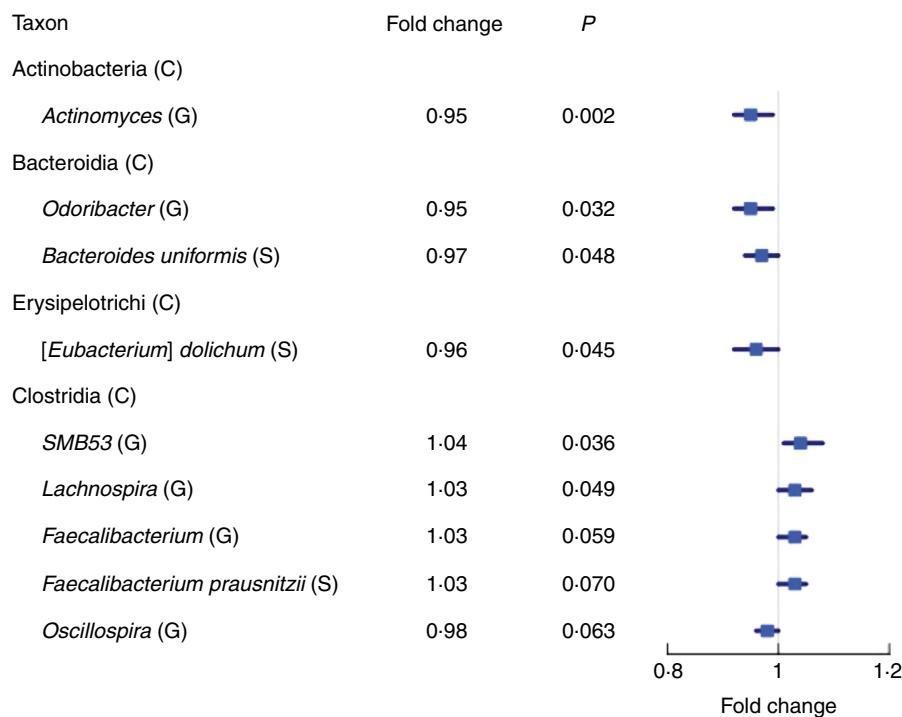


Fig. 2. Forest plot of fold change of select genera (G) and species (S) in both the New York University and National Cancer Institute study populations with a significant or marginally significant association with higher total fibre intake, based on meta-analysis of the two study populations. Meta-analysis nominal *P* values were calculated based on *Z*-score methods. C, class.

UniFrac $P=0.008$; unweighted UniFrac $P=0.01$) but not NCI (weighted UniFrac $P=0.81$; unweighted UniFrac $P=0.75$) after adjusting for covariates of age, sex, race, BMI and cigarette smoking (Fig. 1(c)). In contrast, in both NCI and NYU, total fibre intake was not significantly associated with microbial community diversity as measured by the Shannon diversity index or evenness (Fig. 1(a)).

Taxon abundance

We found fourteen genera with the same direction of association with fibre intake in NCI and NYU, out of twenty-nine total genera that overlapped (i.e. were detected) in both studies. Higher total fibre intake was associated with lower abundance of genera *Actinomyces* (fold change (FC)=0.95, $P=0.002$) of class Actinobacteria and *Odoribacter* (FC=0.95, $P=0.03$) of class Bacteroidia, and was marginally associated with lower abundance of *Oscillospira* (FC=0.96, $P=0.06$) of class Clostridia. Higher total fibre intake also tended to be associated with higher abundance of selected genera of class Clostridia: *SMB53* (FC=1.04, $P=0.04$), *Lachnospira* (FC=1.03, $P=0.05$) and *Faecalibacterium* (FC=1.03, $P=0.06$) (Fig. 2 and Table 2).

At the species level, we found eight species with the same direction of association with fibre intake in the two study populations, out of seventeen total species observed overlapping in both studies. A meta-analysis at this taxonomic level showed a marginal association between higher total fibre intake and higher abundance of *Faecalibacterium prausnitzii* (FC=1.03, $P=0.07$), and lower abundance of *Eubacterium*

dolichum (FC=0.96, $P=0.04$) and *Bacteroides uniformis* (FC=0.97, $P=0.05$) (Fig. 2 and Table 2).

In a sub-analysis to identify associations between taxon abundance and fibre intake from specific dietary sources, we found that higher fibre intake from fruit and vegetables was associated with lower abundance of genera *Actinomyces* (FC=0.97, $P=0.007$), *Odoribacter* (FC=0.96, $P=0.04$) and *Oscillospira* (FC=0.99, $P=0.06$). Higher abundance of genus *Faecalibacterium* was most significant specifically with higher fibre intake from beans (FC=1.11, $P=0.01$). At the species level, higher fibre intake from beans was associated with higher abundance of *F. prausnitzii* (FC=1.11, $P=0.01$) and lower abundance of *B. uniformis* (FC=0.87, $P=0.08$). A heat map representing colour-coded FC of select genera and species, by total fibre and fibre from specific sources (fruits/vegetables, grains and beans), is presented in Fig. 3. Among fibre from different food sources, there was a minimal association between fibre from grains and these specific genera and species.

Although we observed similar associations with taxon abundance and fibre intake in both the NCI and NYU studies, we also noted some inconsistent associations (online Supplementary Table S1). For example, at the phylum level, higher total fibre intake was associated with higher abundance of Proteobacteria in NCI ($P=0.03$) but not in NYU ($P=0.53$). Higher total fibre intake was also marginally associated with higher abundance of phylum Bacteroidetes in NYU ($P=0.07$), but not in NCI ($P=0.47$).

Full results of the meta-analyses examining the association between taxon abundance at each level and dietary fibre intake (total fibre and fibre from different food sources) are reported in online Supplementary Tables S1–S4.

Table 2. Meta-analysis of the association between total fibre intake and genera and species in National Cancer Institute (NCI) and New York University (NYU) study populations*
 (Mean values and 95 % confidence intervals)

Phylum; class; order; family; genus; species	NCI			NYU			Meta FC	Meta 95 % CI	Meta P	Meta Q
	Base mean	FC	P	Base mean	FC	P				
Genus level										
Actinobacteria; Actinobacteria; Actinomycetales; <i>Actinomycetaceae; Actinomyces</i>	2.16	0.884	0.003	0.75	0.969	0.130	0.951	0.917, 0.986	0.002	0.039
Actinobacteria; Actinobacteria; Bifidobacteriales; Bifidobacteriaceae; <i>Bifidobacterium</i>	84.01	1.001	0.973	90.68	1.004	0.850	1.003	0.967, 1.041	0.874	0.967
Actinobacteria; Coriobacteriia; Coriobacteriales; Coriobacteriaceae; <i>Eggerthella</i>	6.18	0.985	NA	2.09	0.990	0.631	0.989	0.951, 1.027	NA	NA
Bacteroidetes; Bacteroidia; Bacteroidales; [Odoribacteraceae]; <i>Odoribacter</i>	0.68	0.995	0.916	10.88	0.946	0.004	0.953	0.920, 0.986	0.032	0.273
Firmicutes; Bacilli; Lactobacillales; Streptococcaceae; <i>Streptococcus</i>	55.61	0.957	0.299	57.84	0.985	0.443	0.980	0.946, 1.014	0.202	0.657
Firmicutes; Clostridia; Clostridiales; Clostridiaceae; <i>SMB53</i>	0.62	1.035	0.443	0.88	1.045	0.029	1.044	1.006, 1.081	0.036	0.273
Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; <i>Lachnospira</i>	31.63	1.042	0.279	163.91	1.026	0.091	1.028	1.000, 1.057	0.050	0.273
Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; <i>Faecalibacterium</i>	230.42	1.030	0.367	549.73	1.025	0.078	1.025	1.000, 1.051	0.059	0.273
Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; <i>Oscillospira</i>	22.58	0.973	0.386	85.60	0.983	0.079	0.982	0.964, 1.000	0.063	0.273
Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; <i>Anaerostipes</i>	5.92	1.007	0.859	10.66	1.018	0.287	1.016	0.985, 1.047	0.377	0.709
Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; <i>Dorea</i>	104.47	1.022	0.418	65.24	1.004	0.666	1.006	0.988, 1.025	0.382	0.709
Firmicutes; Clostridia; Clostridiales; Veillonellaceae; <i>Phascolarctobacterium</i>	18.58	1.009	0.832	34.43	1.012	0.558	1.012	0.975, 1.049	0.571	0.921
Firmicutes; Clostridia; Clostridiales; Veillonellaceae; <i>Dialister</i>	16.99	0.984	0.713	25.15	0.994	0.780	0.992	0.954, 1.030	0.648	0.921
Proteobacteria; Betaproteobacteria; Burkholderiales; <i>Alcaligenaceae; Sutterella</i>	4.32	1.009	0.849	85.39	1.008	0.653	1.008	0.976, 1.040	0.650	0.921
Species level										
Actinobacteria; Actinobacteria; Bifidobacteriales; <i>Bifidobacteriaceae; Bifidobacterium; adolescentis</i>	38.42	1.007	0.898	75.56	1.002	0.932	1.003	0.959, 1.048	0.880	0.880
Actinobacteria; Coriobacteriia; Coriobacteriales; Coriobacteriaceae; <i>Eggerthella; lenta</i>	6.14	0.984	NA	1.99	0.986	0.613	0.986	0.938, 1.034	NA	NA
Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; <i>Bacteroides; uniformis</i>	32.05	0.948	0.202	305.56	0.972	0.128	0.968	0.936, 1.000	0.048	0.256
Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; <i>Bacteroides; ovatus</i>	5.59	0.987	0.790	12.37	0.982	0.309	0.982	0.949, 1.015	0.363	0.570
Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; <i>Faecalibacterium; prausnitzii</i>	230.05	1.032	0.354	540.38	1.024	0.103	1.026	0.998, 1.053	0.070	0.256
Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; <i>Ruminococcus; bromii</i>	24.12	0.966	0.489	1.65	0.969	0.161	0.968	0.930, 1.008	0.138	0.379
Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; <i>Dorea; formicigenerans</i>	1.92	1.033	0.447	20.11	1.018	0.323	1.020	0.987, 1.053	0.216	0.476
Firmicutes; Erysipelotrichi; Erysipelotrichales; Erysipelotrichaceae; <i>[Eubacterium]; dolichum</i>	7.70	0.933	0.198	3.34	0.963	0.122	0.958	0.917, 1.000	0.045	0.256

FC, fold change; Q, false discovery rate-adjusted q-value; NA, non-applicable.

* Taxa shown were associated with fibre intake in the same direction in both NCI and NYU. Relationship between higher quartiles of total fibre intake and differential taxon abundance was evaluated using negative binomial generalised linear models in the DESeq2 package in R. Models were adjusted for age, sex, race, categorical BMI and cigarette smoking status. Meta-analysis was conducted using Z-score methods.

Discussion

In the present study, we examined the association of usual dietary fibre intakes with gut microbiota composition in adults from two independent study populations. In a taxonomy-based meta-analysis, we found that higher total fibre intake is associated with specific taxon abundances, including higher abundance of select genera of Clostridia class. Some of these fibre and taxon abundance associations were consistent with specific fibre food sources, such as fruit and vegetables, grains and beans.

The usual dietary intake of participants measured in our study provides additional insight into the potential effect of longer-term dietary patterns on gut microbiota composition, compared with controlled dietary intervention studies^(12,13).

Although Wu *et al.*⁽¹⁵⁾ demonstrated in a controlled feeding study that certain dietary modifications altered microbiome composition in humans, they did not significantly alter the overall types of microbial communities characterising individuals; the overall structure may be better correlated with long-term diet. Sonnenburg *et al.*⁽³²⁾ also reported that chronic low-fibre diets in mice may drive changes in microbial diversity. Thus, our study associations derived from long-term dietary habits are more likely to represent microbial composition in an uncontrolled, real-world setting.

We found that higher fibre intake was associated with higher abundance of select genera of Clostridia class. This finding is notable given the particular role of *Clostridium* spp. in colonocyte metabolism through production of SCFA, including acetate, propionate and butyrate, via fermentation⁽³³⁾. Although

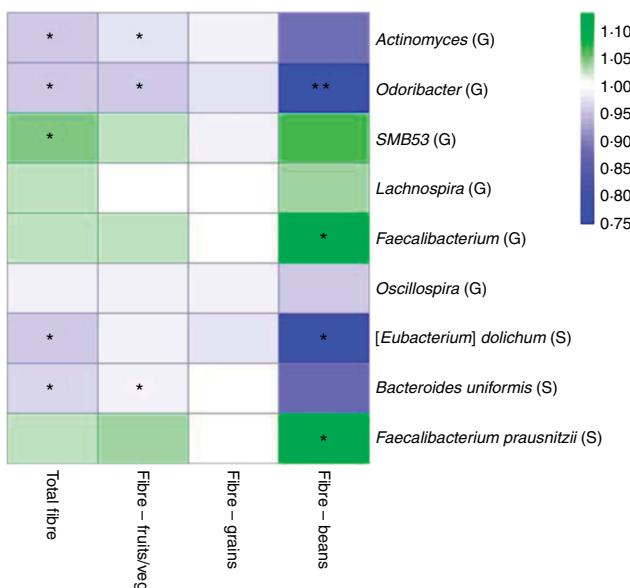


Fig. 3. Heat map representing colour-coded fold changes of select genera (G) and species (S), by total fibre and fibre from specific sources (fruits/vegetables, grains, and beans). * $P < 0.05$ and ** $P < 0.001$ in the meta-analysis of the National Cancer Institute and New York University study populations.

butyrate is the least abundant SCFA produced, it serves as the preferred energy source for colonocytes and has garnered attention for its involvement in cell signalling, maintenance of intestinal barrier function and modulation of immune cell chemotaxis and adhesion^(33–37). Butyrate exerts an anti-inflammatory effect through inhibition of inflammatory cytokines, particularly NF- κ B⁽³⁸⁾. Furthermore, mouse models have shown that butyrate inhibits histone deacetylases and consequently inhibits gene expression and causes tumour suppression^(11,39), carrying implications for colorectal cancer treatment.

Within the Clostridia class, we noted a marginal association between higher fibre intake and higher abundance of species *F. prausnitzii*. The fibre and *F. prausnitzii* relationship has also been observed in a cross-sectional cohort of middle-age and older adults in Spain, in which greater adherence to a Mediterranean diet, rich in higher fibre content foods such as fruits, vegetables and whole grains, correlated with higher levels of *F. prausnitzii*, as well as *Clostridium* cluster XVIa⁽⁴⁰⁾. *Faecalibacterium prausnitzii*, one of the most abundant species found in the gut and a key producer of butyrate, has been associated with anti-inflammatory activity^(41–43). Sokol *et al.*⁽⁴³⁾ suggested that it exerts its anti-inflammatory effects on colitis mouse models, partly owing to associated metabolites blocking production of the pro-inflammatory molecules, NF- κ B and IL-8. Moreover, a reduction in *F. prausnitzii* was associated with a higher risk of recurrence of Crohn's disease in patients post resection. In addition to inflammatory bowel disease, Lopez-Siles *et al.*⁽⁴⁴⁾ reported that lower levels of *F. prausnitzii* were also found in patients with colorectal cancer compared with healthy controls. Thus, these results suggest a potential therapeutic and preventive role of *F. prausnitzii* in countering microbial dysbiosis in human disease. Our finding of higher abundance of *F. prausnitzii* particularly with higher fibre intake

from beans lends support to future investigation of specific diet modifications that could affect disease states.

We also observed that higher fibre intake was associated with lower abundance of genera *Odoribacter*, *Actinomyces* and *Oscillospira*, and higher abundance of genus *Lachnospira*. In mouse models, Zackular *et al.*⁽⁴⁵⁾ reported significant microbial shifts found within stool samples of mice with colon tumours, specifically with enrichment of taxa affiliated with members of *Odoribacter*. Along similar lines, Thomas *et al.*⁽⁴⁶⁾ examined human tissue samples collected during colonoscopy from rectal cancer cases and non-cancer controls, and noted higher abundance of *Odoribacter* in the cancer cases. Kasai *et al.*⁽⁴⁷⁾ noted higher proportions of several genera, including *Actinomyces*, in faecal samples from human subjects with colorectal carcinoma. Furthermore, higher abundance of *Actinomyces* has been reported in both colorectal adenoma and carcinoma cases^(16,47). Whether or not enrichment or depletion of these specific microbes reflects cause or consequence of colon tumour development remains uncertain. Nonetheless, the opposite relationship observed between abundance of these microbes and higher fibre intake in our tumour-free adult population compared with cancer cases in prior studies is notable, and merits further examination of diet, microbiota and disease associations.

Similarly, we previously reported an association between lower abundance of Clostridia and both colorectal carcinoma and adenoma^(6,16). In a case-control study of colorectal cancer cases and controls, stool samples from colorectal cancer cases were characterised by depletion of phylum Firmicutes, predominantly of class Clostridia, relative to controls⁽⁶⁾. In addition, we noted a depleted abundance of members of Clostridia class in stool samples of colorectal adenoma cases, compared with controls⁽¹⁶⁾. These findings suggest a potential relationship between this shift in Clostridia composition and both pre-cancerous and cancerous events. Thus, our finding that higher dietary fibre intake is associated with a higher abundance of select genera of Clostridia class encourages future investigation of specific dietary strategies and their potential protective role against colorectal cancer risk.

Our results are consistent with our previous report of dietary fibre intake and gut microbiota relationships in the NCI study⁽⁴⁸⁾. Our prior analysis of these data showed higher relative abundance of Clostridia class with higher fibre from fruit and vegetable intake, as well as higher relative abundance of phylum Actinobacteria with higher fibre from bean intake. In the present meta-analysis, we confirmed a higher abundance of select genera of Clostridia class with fibre from fruit and vegetables, and a higher abundance of phylum Actinobacteria with fibre from bean intake (online Supplementary Tables S2 and S4). Differences in reported findings between these two studies may be attributed to different statistical methods. Whereas our prior study categorised taxon abundance as a binary variable, in the present study we analysed taxon abundance as a continuous variable in the DESeq2 model, which may increase statistical power.

We note the use of different sequencing platforms for the NCI and NYU studies. As described previously, amplicons in NCI were sequenced using the 454 Roche FLX Titanium



pyrosequencing system, whereas the Illumina MiSeq platform was used in NYU. Certain differences in read lengths and error rates have been reported among these two sequencing platforms⁽⁴⁹⁾. Nonetheless, comparison of these two platforms in gut microbiome samples by Allali *et al.*⁽⁵⁰⁾ has demonstrated that, although there may be some differences in abundance of specific taxa, similar microbiome compositional profiles were still obtained, and similar biological conclusions could therefore be drawn across both platforms. In addition, we maintained a similar system of sequence data processing using the QIIME pipeline across both of our studies, to ensure consistency in analysis.

Several potential limitations also need to be considered. Measurement error is an inherent limitation in self-reported dietary assessment. Given that the NCI and NYU studies were independent investigations conducted during different time periods, dietary questionnaires utilised differed between studies, which probably contributed to noted differences in quantity of fibre intake in the two cohorts. For this reason, we used study-specific quartile trends of higher fibre intake to evaluate relationships with global microbial diversity and specific taxon abundance. Subjects in both study cohorts were mostly white, and thus our findings may not be generalisable to more racially diverse populations. The cross-sectional design of both studies limits assessment of the temporality of the diet–microbiota relationship. Furthermore, we lacked data in our study cohorts regarding faecal metabolites. Nonetheless, strengths of this study include the relatively large sample size, excellent quality control of microbiome assays in both studies and adjustment for potential confounders. Furthermore, our study hypothesis was tested in two independent study populations, and our meta-analysis findings identified similar taxon abundance associations across both populations.

In summary, we found that fibre intake was associated with gut microbiota composition in generally healthy adults. Fibre from specific food sources may also yield differential effects on gut microbiota. In our analysis, greater fibre intake, particularly from beans, tended to be associated with greater abundance of putatively beneficial bacteria, such as *F. prausnitzii*, which has demonstrated anti-inflammatory activity. Fibre intake has generally been acknowledged as beneficial in relation to human diseases, such as cancer, and recommendations for sufficient intake have been promoted by the United States Department of Agriculture⁽⁵¹⁾. However, there remains a substantial gap in understanding the complex interplay between specific dietary fibre sources, gut microbiota and disease state. Consequently, further investigation is warranted to study microbiota-targeted dietary interventions in both a therapeutic and preventive disease setting. Given the mounting evidence that microbial dysbiosis may affect human health and contribute to the development of colorectal cancer, it is imperative to better elucidate the association of diet–microbiota inter-relationships and their potential impact on colorectal cancer risk. A deeper understanding of diet and the gut microbiome may lay the groundwork needed for future study of diet–microbiota-based colorectal cancer prevention strategies.

Acknowledgements

Research reported in this publication was supported in part by the US National Cancer Institute under award numbers U01CA182370, R01CA164964, R03 159414 and P30CA016087. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the US Department of Veterans Affairs or the United States Government. The National Cancer Institute had no role in the design, analysis or writing of this article.

D. L., B. A. P., R. B. H. and J. A. designed the research, conducted the study, analysed the data, performed statistical analyses and wrote the first draft of the manuscript. C. F., H. J. F., M. A. B., R. S., J. J. G., R. B. H. and J. A. provided reagents or essential materials. J. A. had primary responsibility for final content. All authors critically revised and approved the final manuscript.

The authors declare that there are no conflicts of interest.

Supplementary material

For supplementary material/s referred to in this article, please visit <https://doi.org/10.1017/S0007114518002465>

References

- Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214.
- Hold GL, Smith M, Grange C, *et al.* (2014) Role of the gut microbiota in inflammatory bowel disease pathogenesis: what have we learnt in the past 10 years? *World J Gastroenterol* **20**, 1192–1210.
- Marchesi JR, Dutilh BE, Hall N, *et al.* (2011) Towards the human colorectal cancer microbiome. *PLoS ONE* **6**, e20447.
- Han M, Wang C, Liu P, *et al.* (2017) Dietary fiber gap and host gut microbiota. *Protein Pept Lett* **24**, 388–396.
- Sobhani I, Tap J, Roudot-Thoraval F, *et al.* (2011) Microbial dysbiosis in colorectal cancer (CRC) patients. *PLoS ONE* **6**, e16393.
- Ahn J, Sinha R, Pei Z, *et al.* (2013) Human gut microbiome and risk for colorectal cancer. *J Natl Cancer Inst* **105**, 1907–1911.
- Chen W, Liu F, Ling Z, *et al.* (2012) Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PLOS ONE* **7**, e39743.
- Weir TL, Manter DK, Sheflin AM, *et al.* (2013) Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults. *PLOS ONE* **8**, e70803.
- Bouvard V, Loomis D, Guyton KZ, *et al.* (2015) Carcinogenicity of consumption of red and processed meat. *Lancet Oncol* **16**, 1599–1600.
- Aune D, Chan DS, Lau R, *et al.* (2011) Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. *BMJ* **343**, d6617.
- Donohoe DR, Holley D, Collins LB, *et al.* (2014) A gnotobiotic mouse model demonstrates that dietary fiber protects against colorectal tumorigenesis in a microbiota- and butyrate-dependent manner. *Cancer Discov* **4**, 1387–1397.
- O'Keefe SJ, Li JV, Lahti L, *et al.* (2015) Fat, fibre and cancer risk in African Americans and rural Africans. *Nat Commun* **6**, 6342.

13. Tap J, Furet JP, Bensaada M, *et al.* (2015) Gut microbiota richness promotes its stability upon increased dietary fibre intake in healthy adults. *Environ Microbiol* **17**, 4954–4964.
14. Liu P, Zhao J, Guo P, *et al.* (2017) Dietary corn bran fermented by *Bacillus subtilis* MA139 decreased gut cellulolytic bacteria and microbiota diversity in finishing pigs. *Front Cell Infect Microbiol* **7**, 526.
15. Wu GD, Chen J, Hoffmann C, *et al.* (2011) Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108.
16. Peters BA, Dominianni C, Shapiro JA, *et al.* (2016) The gut microbiota in conventional and serrated precursors of colorectal cancer. *Microbiome* **4**, 69.
17. Schiffman MH, Van Tassell RL, Robinson A, *et al.* (1989) Case-control study of colorectal cancer and fecapentaene excretion. *Cancer Res* **49**, 1322–1326.
18. Schiffman MH, Andrews AW, Van Tassell RL, *et al.* (1989) Case-control study of colorectal cancer and fecal mutagenicity. *Cancer Res* **49**, 3420–3424.
19. World Health Organization (2006) *BMI Classification. Global Database on Body Mass Index*. Geneva: WHO.
20. Willett WC, Howe GR & Kushi LH (1997) Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* **65**, Suppl. 4, 1220S–1228S (discussion 9S–31S).
21. Nossa CW, Oberdorf WE, Yang L, *et al.* (2010) Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome. *World J Gastroenterol* **16**, 4135–4144.
22. Caporaso JG, Lauber CL, Walters WA, *et al.* (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* **108**, Suppl. 1, 4516–4522.
23. Caporaso JG, Kuczynski J, Stombaugh J, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335–336.
24. Haas BJ, Gevers D, Earl AM, *et al.* (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* **21**, 494–504.
25. Love MI, Huber W & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.
26. Shannon CE (1997) The mathematical theory of communication. 1963. *MD Comput* **14**, 306–317.
27. Lozupone CA, Hamady M, Kelley ST, *et al.* (2007) Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Appl Environ Microbiol* **73**, 1576–1585.
28. McArdle BH & Anderson MJ (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology* **82**, 290–297.
29. Benjamini Y & Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Methodol* **57**, 289–300.
30. Benjamini Y, Drai D, Elmer G, *et al.* (2001) Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* **125**, 279–284.
31. Evangelou E & Ioannidis JP (2013) Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet* **14**, 379–389.
32. Sonnenburg ED, Smits SA, Tikhonov M, *et al.* (2016) Diet-induced extinctions in the gut microbiota compound over generations. *Nature* **529**, 212–215.
33. Liu H, Wang J, He T, *et al.* (2018) Butyrate: a double-edged sword for health? *Adv Nutr* **9**, 21–29.
34. Pryde SE, Duncan SH, Hold GL, *et al.* (2002) The microbiology of butyrate formation in the human colon. *FEMS Microbiol Lett* **217**, 133–139.
35. Donohoe DR, Garge N, Zhang X, *et al.* (2011) The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* **13**, 517–526.
36. Chen J, Li Y, Tian Y, *et al.* (2015) Interaction between microbes and host intestinal health: modulation by dietary nutrients and gut-brain-endocrine-immune axis. *Curr Protein Pept Sci* **16**, 592–603.
37. Meijer K, de Vos P & Priebe MG (2010) Butyrate and other short-chain fatty acids as modulators of immunity: what relevance for health? *Curr Opin Clin Nutr Metab Care* **13**, 715–721.
38. Aguilar EC, Leonel AJ, Teixeira LG, *et al.* (2014) Butyrate impairs atherosclerosis by reducing plaque inflammation and vulnerability and decreasing NFκB activation. *Nutr Metab Cardiovasc Dis* **24**, 606–613.
39. Donohoe DR, Collins LB, Wali A, *et al.* (2012) The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Mol Cell* **48**, 612–626.
40. Gutierrez-Diaz I, Fernandez-Navarro T, Salazar N, *et al.* (2017) Adherence to a Mediterranean diet influences the fecal metabolic profile of microbial-derived phenolics in a Spanish cohort of middle-age and older people. *J Agric Food Chem* **65**, 586–595.
41. Arumugam M, Raes J, Pelletier E, *et al.* (2011) Enterotypes of the human gut microbiome. *Nature* **473**, 174–180.
42. Louis P & Flint HJ (2009) Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol Lett* **294**, 1–8.
43. Sokol H, Pigneur B, Watterlot L, *et al.* (2008) *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci U S A* **105**, 16731–16736.
44. Lopez-Siles M, Martinez-Medina M, Suris-Valls R, *et al.* (2016) Changes in the abundance of *Faecalibacterium prausnitzii* phylotypes I and II in the intestinal mucosa of inflammatory bowel disease and patients with colorectal cancer. *Inflamm Bowel Dis* **22**, 28–41.
45. Zackular JP, Baxter NT, Iverson KD, *et al.* (2013) The gut microbiome modulates colon tumorigenesis. *mBio* **4**, e00692-e00713.
46. Thomas AM, Jesus EC, Lopes A, *et al.* (2016) Tissue-associated bacterial alterations in rectal carcinoma patients revealed by 16S rRNA community profiling. *Front Cell Infect Microbiol* **6**, 179.
47. Kasai C, Sugimoto K, Moritani I, *et al.* (2016) Comparison of human gut microbiota in control subjects and patients with colorectal carcinoma in adenoma: terminal restriction fragment length polymorphism and next-generation sequencing analyses. *Oncol Rep* **35**, 325–333.
48. Dominianni C, Sinha R, Goedert JJ, *et al.* (2015) Sex, body mass index, and dietary fiber intake influence the human gut microbiome. *PLOS ONE* **10**, e0124599.
49. Liu L, Li Y, Li S, *et al.* (2012) Comparison of next-generation sequencing systems. *J Biomed Biotechnol* **2012**, 251364.
50. Allali I, Arnold JW, Roach J, *et al.* (2017) A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiol* **17**, 194.
51. Fungwe TV, Bente L & Hiza H (2007) *The Food Supply and Dietary Fiber: Its Availability and Effect on Health: Nutrition Insight* 36. Alexandria, VA: USDA Center for Nutrition Policy and Promotion.