Long-term association between diet quality and characteristics of the gut microbiome in the multiethnic cohort study

Erica Ma1, Gertraud Maskarinec1*, Unhee Lim1, Carol J. Boushey1, Lynne R. Wilkens1, V. Wendy Setiawan2, Loïc Le Marchand3, Timothy W. Randolph3, Isaac C. Jenkins3, Keith R. Curtis3, Johanna W. Lampe3 and Meredith A.J. Hullar3

1University of Hawai’i Cancer Center, Honolulu, HI, USA
2Department of Preventive Medicine, University of Southern California, Los Angeles, CA, USA
3Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Abstract
As past usual diet quality may affect gut microbiome (GM) composition, we examined the association of the Healthy Eating Index (HEI)-2015 assessed 21 and 9 years before stool collection with measures of fecal microbial composition in a subset of the Multiethnic Cohort. A total of 5936 participants completed a validated quantitative FFQ (QFFQ) at cohort entry (Q1, 1993–1996), 5280 at follow-up (Q3, 2003–2008) and 1685 also at a second follow-up (Adiposity Phenotype Study (APS), 2013–2016). All participants provided a stool sample in 2013–2016. Fecal microbial composition was obtained from 16S rRNA gene sequencing (V1–V3 regions). HEI-2015 scores were computed based on each QFFQ. Using linear regression adjusted for relevant covariates, we calculated associations of HEI-2015 scores with gut microbial diversity and the relative abundance of the phylogenetic α- and β-diversity assessed by the Shannon Index was significantly higher with increasing tertiles of HEI-2015. Of the 152 bacterial genera tested, seven (Anaerostipes, Coprococcus_2, Eubacterium eligens, Lachnospira, Lachnospiraceae_UCG-013, Ruminococcaceae_UCG-013 and Ruminococcus) were positively and five (Collinsella, Parabacteroides, Ruminiclostridium_5, Ruminococcus gnavus and Tyzzerella) were inversely associated with HEI-2015 assessed in Q1, Q3 and APS. The estimates of change per unit of the HEI-2015 score associated with the abundance of these twelve genera were consistent across the three questionnaires. The quality of past diet, assessed as far as 20 years before stool collection, is equally predictive of GM composition as concurrently assessed diet, indicative of the long-term consistency of this relation.

Key words: Gut microbiome; Diet; Healthy Eating Index; 16S rRNA gene sequencing

Over the past decade, interest in gut microbiome (GM) composition in relation to health status has increased tremendously due to technical advances in genomic analyses(1,2). Many determinants of microbiome patterns, e.g. ageing, use of antibiotics, other medications, environmental exposures and genetics, have been identified(2,3). Nutritional factors influence the presence and activity of specific gut bacteria and are linked to conditions such as the metabolic syndrome, colorectal cancer, inflammatory bowel disease and CVD have emerged(2,4). Diet quality indices are commonly utilised to analyse the impact of overall diet composition and GM composition have also been detected in several investigations(4,5). Lower HEI-2005 scores were associated with significantly higher levels of total SCFA in fecal samples and linked to a significant reduction in overall mortality and with greater α diversity and the relative abundance of the phylum Actinobacteria and twenty-one genera(6,7), including seven members of the Lachnospiraceae family from the phylum Firmicutes. Similarly, in a Swedish study, higher scores for a “health-conscious” food pattern were related to the abundance of six genera of the phylum Firmicutes(8), loading positively in relation to Roseburia, Lachnospira and unclassified in order RF39 and negatively in relation to Blautia, Eubacterium and Anaerotruncus. Lower HEI-2005 scores were associated with a lower relative abundance of Roseburia in a recent report(9). Beneficial effects of the Mediterranean Diet on the diversity and GM composition have also been detected in several investigations(8,11–13). For example, adherence to the MD was associated with significantly higher levels of total SCFA in fecal samples and linked to a significant reduction in overall mortality and

Abbreviations: DXA, Dual-energy X-ray Absorptiometry; GM, gut microbiome; HEI, Healthy-Eating Index; mGWAS, Microbiome Genome-Wide Association Study; MEC, Multiethnic Cohort; MRI, Magnetic Resonance Imaging; QFFQ, quantitative FFQ.
* Corresponding author: Gertraud Maskarinec, email gertraud@cc.hawaii.edu
morbidity. Also, a greater presence of Bacteroidetes was associated with a lower animal protein intake. A recent report identified a panel of intestinal species that were associated with healthy dietary habits and overlapped with those associated with favourable cardiometabolic and postprandial markers. The current analysis examines the association of the HEI-2015 with gut microbial diversity and the structure of the gut microbial community, expanding our previous MEC cross-sectional analysis of 1735 participants and 103 individual bacteria to a larger study population of more than 5000 cohort members and an updated reference library with 152 genera. By analysing diet quality assessed at cohort entry (Q1) and follow-up (Q3), 21 and 9 years before stool collection, respectively, long-term exposure and changes in diet quality over time can be evaluated. Significant findings for past diet as assessed in Q1 and Q3 were confirmed in the data set that assessed diet quality (APS) and the GM at the same time.

Materials and methods

Study population

The MEC is an ongoing study of more than 215 000 participants aged 45–75 years from five ethnic groups (Japanese American, white, Latino, African American and Native Hawaiian) recruited in Hawaii and Los Angeles, California (online Supplemental Fig. S1). As previously published, all participants completed a self-administered questionnaire (Q1) at cohort entry (1993–1996), which collected information on anthropometrics, sociodemographic factors, medical history, physical activity and a quantitative FFQ (QFFQ). A follow-up questionnaire (Q3) during 2003–2008 was completed by 98 214 (46 %) participants. As part of a Microbiome Genome-Wide Association Study (mGWAS) study and the APS conducted in 2013–2016, 6094 participants were recruited from the MEC study using a uniform stool collection protocol. The study protocol was approved by the Biomedical Institutional Review Boards at the University of Hawaii (CHS#17200) and the University of Southern California (HS#12-00625); all participants signed an informed consent form. After excluding individuals with invalid dietary information as well as missing BMI and/or microbiome data (n = 158), the final data set consisted of 5936 mGWAS and APS participants. As 656 did not complete the follow-up questionnaire, 5280 cohort members were part of the Q3 data set.

Microbiome genome-wide association study

The mGWAS includes a MEC subset of 4502 participants aged 60–90 years who had been previously genotyped for different GWAS projects. Exclusion criteria due to their known effects on the gut microbiota included ileostomy/colectomy and dialysis. Participation was deferred for recent treatment with chemotherapy, radiation therapy, corticosteroid hormones, prescription weight-loss drugs, insulin or thyroid medications, antibiotics and flu shot or other vaccination; recent endoscopy or irrigation and/or cleansing of the large intestine and recent substantial weight change. Participants were asked to mail their stool sample with a completed questionnaire addressing exposures related to the fecal microbiome, such as recent antibiotics use. The overall participation rate was 43 % among eligible individuals.

Adiposity phenotype study

The APS recruited a subset of 1861 MEC participants aiming to study body fat distribution and its determinants. Participants aged 60–77 years were stratified by sex, ethnicity and BMI categories (18–5–21.9, 22–24.9, 25–26.9, 27–29.9, 30–34.9 and 35–40 kg/m²). Exclusion criteria included current BMI outside 18.5–40 kg/m², current/recent (<2 years) smoking, soft or metal implants or serious health conditions, with similar deferment criteria as in the mGWAS for recent relevant exposures. The overall participation rate was 25 % of eligible individuals. During a clinic visit, blood and stool samples, anthropometric measures, questionnaire data and Magnetic Resonance Imaging (MRI) and Dual-energy X-ray Absorptiometry (DXA) scans were obtained.

Dietary assessment and healthy eating index-2015

A detailed description of the QFFQ development and calibration has been reported previously. Briefly, the QFFQ was designed to include food groups that contributed >85 % (for each ethnic group) of the intake of nutrients of interest during the past year. The frequency of over 180 food items was assessed using eight categories for food groups and nine for beverages and the amount of each item was recorded in three serving sizes. Although the substance of the QFFQ remained the same, the original QFFQ was updated in 2003 for Q3 and the APS to modify the food lists, amounts and examples or names given for the food items without substantial change.

Using an extensive food composition database and the same approach, estimated daily intakes were computed to calculate the HEI-2015, which is an updated version of the original HEI-2005 and HEI-2010 reflecting the Dietary Guidelines for American 2015–2020. The HEI-2015 is composed of thirteen components with a maximum of 100 points and includes nine adequacy components (total fruits, whole fruits, total vegetables, greens and beans, whole grains, dairy, total protein foods, seafood and plant proteins and fatty acids) as well as four moderation components (refined grains, Sodium, added sugars and saturated fats). Most components are scored as amounts per 1000 kcal of intake. The key differences between the 2010 and 2015 index include the allocation of legumes, which was modified to total vegetables, greens and beans, total protein foods and seafood/plant proteins and the empty energy component, which was replaced with saturated fat and added sugars.

Stool collection

During mGWAS and APS studies in 2013–2016, stool samples were collected in RNAlater (Ambion) and stored at −80 °C until bulk shipment on dry ice from Honolulu or Los Angeles to Seattle for processing. Stool sampling was delayed for at least 3 months for participants who first reported use of antibiotics. As previously described, DNA was extracted and amplified for the V1–V3 region of the 16S rRNA genes, and amplicons were sequenced on the MiSeq platform (Illumina, San Diego, CA). To classify bacterial taxonomy, sequences were processed using
Diet quality and the gut microbiome

QIIME v.1.8 (20) and SILVA 1.32, as previously described but updated in comparison with the previous study (29). The filtering strategy for operational taxonomic units included parameters in QIIME to exclude low abundant sequences, singletons and chimeras and final filtering at the genera level, in which we removed genera that appeared in <10% of the samples (27). Diversity indices were calculated on ComBat-adjusted (28) operational taxonomic unit counts and then exported for statistical analysis.

Statistical analysis

All statistical modelling was conducted with SAS version 9.4 software (SAS Institute Inc., Cary, NC). All phylum and genus variables had undergone ComBat-adjustment (28) to correct values across laboratory batches and centered log-ratio transformation to account for their compositional nature (29). The HEI-2015 scores were categorised into tertiles with tertile 1 indicating the lowest and tertile 3 the highest diet quality. To examine the association of diet quality with α diversity (Shannon, Chao1 and PD Whole Tree) and relative abundance of phyla and genera, general linear models were applied to estimate covariate-adjusted least square means of gut microbial measures by HEI-2015 tertiles; trend tests using the continuous variables assessed possible dose–response relations. The regression coefficients in these models indicate the change in relative abundance of a GM component per one unit of the HEI-2015 score. All models were adjusted for ethnicity, sex, age at stool collection, BMI, alcohol intake, physical activity (<0.5 v. ≥0.5 h moderate/vigorous activity), total energy intake (log-transformed), smoking status and antibiotic use within the last year, but not other medications as only information on use of antifungals (1%), laxatives (19%) and probiotics (8%) was requested at the time of stool collection.

Assuming a nominal type-I error of 0.05, Bonferroni corrections were applied to the analysis of the 10 phyla (0.05/10 = 0.005) and the 152 genera (0.05/152 = 0.00033) to lower the likelihood of chance associations. To analyse diet quality with α diversity assessed 20·8 (sd 1·2) and 9·3 (sd 1·4) years before stool sample collection, separate models were performed for nutritional information obtained at Q1 and Q3. In addition, the APS data set with 1685 participants who completed the QFFQ at the same time as stool sample collection were analysed. Participants who had completed Q1 and Q3 were assigned change variables (T1T1, T1T2, T1T3, T2T1, T2T2, T2T3, T3T1, T3T2 and T3T3). For instance, T1T2 indicates that the individual was in tertile 1 at Q1 and in tertile 2 at Q3. Because of the small sample size, change from Q1 to APS (n = 1685) was not computed. To evaluate the influence of change in diet quality between Q1 and Q3, these nine categories were modelled as independent variables using the same approach as above and represented in a heat map. The 13 HEI-2015 components were plotted in a spider plot to identify changes in diet quality responsible for a decrease or increase in tertiles from Q1 to Q3.

Results

The 5936 participants at Q1 (Table 1) were equally distributed between women (52·7%) and men (48·3%). A total of 5280 participants completed Q3 and 1685 at APS. The respective mean ages at Q1, Q3 and APS were 54, 65 and 69 years, while the mean time periods to stool collection were 21, 9 and 0 years. Mean HEI-2015 scores increased from Q1 (67 (sd 10)) to Q3 (71 (sd 11)) and APS (72 (sd 10)). Japanese Americans comprised the majority of the study population. A larger proportion of women and normal weight individuals were classified as HEI-2015 tertile 3, whereas men, as well as overweight and obese individuals, were more likely to belong to tertile 1. The Spearman correlation coefficient between Q1 and Q3 was 0·49 (P < 0·0001). Mean total energy intake and alcohol consumption were lower in participants in tertile 3 than 1.

Alpha diversity

Alpha diversity as assessed by the Shannon Index (Table 2) was significantly greater across tertiles for diet assessed in Q1 (P < 0·0001), Q3 (P < 0·0001) and APS (P < 0·02). The difference between tertiles 1 and 3 was approximately 1% for both Q1 and Q3. This trend was also seen for PD Whole Tree, but was only significant for diet assessed at Q1 (P = 0·002). There were no significant associations of the Chao1 index with diet quality at any point in time.

Phyla

Of the 10 phyla, Firmicutes had the highest mean relative abundance (54%), followed by Bacteroidetes (40%) and Proteobacteria (3%) (data not shown). Although only the HEI-2015 assessed at the time of stool collection showed a significant inverse relation with Actinobacteria, all regression coefficients describing the difference in relative abundance per one unit of the HEI-2015 score were in the same direction (Table 3). Diet quality at Q1 was positively associated with the presence of Firmicutes but not with diet assessed at stool collection. A significant inverse association of the HEI-2015 with Proteobacteria was only seen for diet at stool collection.

Genera

Of the 152 genera, the following had a relative abundance of at least 1%: Bacteroides (28%), Faecalibacterium (9·4%), Parabacteroides (1·5%), Ruminococcus 1 (1·5%), Lachnoclostridium (1·1%) and Escherichia-Shigella (1·0%). The genera Eubacterium xylanophilum and Tyzzerella had the highest zero prevalence (26·7% and 17·8%). Among the 152 bacterial genera, twenty eight were significantly associated with diet as assessed in at least one of the questionnaires (online Supplemental Table S1). At the phylum (family) level, they were twenty-three families in Firmicutes (12 Lachnospiraceae, 9 Ruminococcaceae, 1 Christensenellaceae and 1 Erysipelotrichaceae), three in Bacteroidetes (Bacteroidaceae, Marinillicaceae and Tannerellaceae), one in Actinobacteria (Coriobacteriaceae) and one in Proteobacteria (Enterobacteriaceae).

Specifically, twelve genera were significantly associated with diet quality assessed at all three time points as shown by the regression coefficients for relative abundance per one unit of the HEI-2015 score (Table 4), while six genera were significantly
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All T1 †</td>
<td>T2</td>
<td>T3</td>
</tr>
<tr>
<td>Range of index scores</td>
<td>31.7–62.3</td>
<td>62.3–72.0</td>
<td>72.0–96.6</td>
</tr>
<tr>
<td>N</td>
<td>1964</td>
<td>1975</td>
<td>1997</td>
</tr>
<tr>
<td>Age at Qx, y</td>
<td>5936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>52.2</td>
<td>53.7</td>
<td>55.3</td>
</tr>
<tr>
<td>sd</td>
<td>6.2</td>
<td>6.8</td>
<td>7.1</td>
</tr>
<tr>
<td>Age at stool sample, y</td>
<td>5936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>73.0</td>
<td>74.5</td>
<td>76.2</td>
</tr>
<tr>
<td>sd</td>
<td>6.3</td>
<td>6.8</td>
<td>7.1</td>
</tr>
<tr>
<td>Time to stool collection, y</td>
<td>5936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>20.8</td>
<td>20.8</td>
<td>20.9</td>
</tr>
<tr>
<td>Sex, %</td>
<td>5936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2866</td>
<td>39</td>
<td>34</td>
</tr>
<tr>
<td>F</td>
<td>3070</td>
<td>27</td>
<td>33</td>
</tr>
<tr>
<td>Ethnicity, %</td>
<td>1004</td>
<td>23</td>
<td>33</td>
</tr>
<tr>
<td>Whites</td>
<td>935</td>
<td>23</td>
<td>32</td>
</tr>
<tr>
<td>African American</td>
<td>793</td>
<td>37</td>
<td>32</td>
</tr>
<tr>
<td>Native Hawaiian</td>
<td>2087</td>
<td>39</td>
<td>32</td>
</tr>
<tr>
<td>Japanese American</td>
<td>1117</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>Latino</td>
<td>Normal (18.5–24.9)</td>
<td>2601</td>
<td>29</td>
</tr>
<tr>
<td>Overweight (25–29.9)</td>
<td>2350</td>
<td>36</td>
<td>33</td>
</tr>
<tr>
<td>Obese (30+)</td>
<td>985</td>
<td>37</td>
<td>36</td>
</tr>
<tr>
<td>Moderate/vigorous active, %</td>
<td>&lt;0.5 h/d</td>
<td>1626</td>
<td>37</td>
</tr>
<tr>
<td>0.5– h/d</td>
<td>4140</td>
<td>31</td>
<td>33</td>
</tr>
<tr>
<td>Smoking Status, %</td>
<td>Never</td>
<td>2845</td>
<td>28</td>
</tr>
<tr>
<td>Ever‡</td>
<td>3045</td>
<td>38</td>
<td>32</td>
</tr>
<tr>
<td>Antibiotic use, %</td>
<td>1119</td>
<td>34</td>
<td>33</td>
</tr>
<tr>
<td>Total energy, kcal/d</td>
<td>Mean</td>
<td>2321</td>
<td>2251</td>
</tr>
<tr>
<td>sd</td>
<td>1086</td>
<td>1017</td>
<td>919</td>
</tr>
<tr>
<td>Alcohol intake, g/d</td>
<td>Mean</td>
<td>7.9</td>
<td>8.0</td>
</tr>
<tr>
<td>sd</td>
<td>19.1</td>
<td>1017</td>
<td>19.0</td>
</tr>
</tbody>
</table>

† T1 = lowest or first tertile, T2 = intermediate or second tertile, T3 = highest or third tertile.
‡ Individuals who were either past-smokers or current smokers.
Table 2. Mean (95% CI) for diversity measures by tertiles of HEI-2015 over time*
(Mean values and 95% confidence intervals)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean 95% CI</td>
<td>Mean 95% CI</td>
<td>Mean 95% CI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P&lt;sub&gt;mean&lt;/sub&gt;</td>
<td>P&lt;sub&gt;mean&lt;/sub&gt;</td>
<td>P&lt;sub&gt;mean&lt;/sub&gt;</td>
</tr>
<tr>
<td>Shannon Index</td>
<td>T1</td>
<td>6.44</td>
<td>6.27, 6.61</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>6.51</td>
<td>6.34, 6.68</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>6.53</td>
<td>6.36, 6.71</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>T1</td>
<td>1061</td>
<td>1027, 1095</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>1072</td>
<td>1038, 1106</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>1068*</td>
<td>1034, 1103</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>T1</td>
<td>24.7</td>
<td>23.8, 25.6</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>24.9</td>
<td>24.0, 25.7</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>24.9</td>
<td>24.0, 25.8</td>
<td>0.002</td>
</tr>
<tr>
<td>PD Whole Tree</td>
<td>T1</td>
<td>24.7</td>
<td>23.8, 25.6</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>24.9</td>
<td>24.0, 25.7</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>24.9</td>
<td>24.0, 25.8</td>
<td>0.002</td>
</tr>
</tbody>
</table>

* Healthy Eating Index-2015 (HEI-2015) as independent variable, adjusted mean obtained by linear regression for diversity indices; values shown are as means (95% CI) adjusted for sex, age at stool collection, ethnicity, BMI (kg/m²), total energy intake (log-transformed), physical activity (hour/day), smoking status (never/ever), antibiotic use in past year (yes/no) and alcohol intake
† T1 = lowest or first tertile, T2 = intermediate or second tertile, T3 = highest or third tertile
‡ P<sub>mean</sub> obtained with HEI-2015 as continuous variable and adjusted for sex, age at stool collection, ethnicity, BMI, total energy intake (log-transformed), physical activity, smoking status, antibiotic use in past year and alcohol intake

Discussion

Of the 5280 participants who completed Q1 and Q3 (74.9%), 3589 participants were included in the analysis. The absolute frequency of changes over the HEI-2015 tertiles was very similar across ethnic groups with 53% of cohort members remaining in the same tertile of diet quality, 1225 (23%) shifted to the lower tertile, and 886 (16%) moved to the upper tertile. Of the 5280 participants who completed Q1 and Q3, 2849 (54%) reached the lower values at Q1 and intermediate estimates at Q3. The regression coefficients were greatest at APS with 12% lower and 13% (25%) to a higher tertile. The association between diet quality and gut microbiota was stronger compared with the cross-sectional association of changes across tertiles, participants who moved to a higher tertile from Q1 to Q3 showed stronger adherence to the whole grains recommendations but also improved their scores for refined grains recommendations. This was consistent across all ethnic groups.

Table S1. Change in diet quality and the gut microbiota

Change in diet quality

The lower values at Q1 and intermediate estimates at Q3. The association became stronger compared with the cross-sectional association of 30% for diet as assessed at stool collection, and there was only significant for diet as assessed at stool collection. The exception was at Q1 and Q3 (p<0.005), a significant increase in the absolute differences across tertiles of the HEI-2015 were greater than Q1 (Fig. 1), the associations across tertiles for participants with lower diet quality at Q1.

As to the HEI-2015 components responsible for a change in the gut microbiota, we observed strong positive associations for Eubacterium eligens, Lachnospira, Parabacteroides, Ruminoclostridium 5, and five (Collinsella, Lachnospira, Parabacteroides, Ruminococcaceae UCG-008, and Tyzzerella). Of the twelve significant genera, the regression coefficients were greatest at APS with 12% lower and 13% (25%) to a higher tertile. The association became stronger compared with the cross-sectional association of changes across tertiles, participants who moved to a higher tertile from Q1 to Q3 showed stronger adherence to the whole grains recommendations but also improved their scores for refined grains recommendations. This was consistent across all ethnic groups.

Table S1. Change in diet quality and the gut microbiota

Change in diet quality

The lower values at Q1 and intermediate estimates at Q3. The association became stronger compared with the cross-sectional association of changes across tertiles, participants who moved to a higher tertile from Q1 to Q3 showed stronger adherence to the whole grains recommendations but also improved their scores for refined grains recommendations. This was consistent across all ethnic groups.
of twenty-eight genera was significantly associated with diet assessed in at least one questionnaire; the majority belonged to the Lachnospiraceae or Ruminococcaceae families of the Clostridiales order (online Supplemental Table S1). Of these, twelve genera were associated with diet quality as assessed at all three time points (Table 4): Anaerostipes, Coprococcus 2, Eubacterium eligens group, Lachnospira, Lachnospiraceae, Ruminococcus, Collinsella, Parabacteroides, Ruminiclostridium 5, Ruminococcus gnavus group and Tyzzerella. The fact the strength of the association as assessed by the regression coefficient for ten of the significant twelve genera were remained
quite stable supports the importance of long-term dietary exposure in shaping gut microbial composition although diet close to stool collection shows a stronger association.

Combined diet quality as assessed in Q1 and Q3 (Fig. 1) indicated trends, positive for some and inverse for other bacteria, in the abundance of the twelve genera for participants reporting higher diet quality at both time points and in particular when closer to stool collection as compared with those with low HEI-2015 scores at Q1 and Q3. The availability of repeated dietary information and the analysis of change in diet can be considered novel in comparison with previous publications reporting on diet quality and the gut microbiome.

The current findings are consistent with a previous MEC report that found \( \alpha \) diversity to be significantly associated with better diet quality. However, the importance of the very small change or lack of change in \( \alpha \) diversity seen here and in previous reports is not clear. A review of dietary intervention studies also detected a limited effect on microbial diversity across...
This report, a higher abundance of Tyzzerella, which has been associated with cardiovascular disease (44), was seen for lower scores of the HEI (10).

The only significant member of the phylum Actinobacteria, Collinsella, was inversely associated with diet across time periods in the current analysis and with higher scores of four indices (HEI-2010, AHEI, DASH and aMED) among APS participants (40). A prudent pattern, characterised by high consumption of fruits, vegetables, nuts, fish and chicken and turkey, among older community-dwelling men showed similar findings (34). Collinsella is known to be associated with low fibre diets (38) and result in inflammatory status (42) , fatty liver (45) and related diseases (46). Escherichia-Shigella was associated with diet quality at the time of stool collection only (online Supplemental Table S1). This may be due to the higher mean age of participants who participated in the APS as it is well known that ageing increases the likelihood to acquire these bacteria due to lower immunity or medication use (47–49).

Parabacteroides, the only significant Bacteroidetes, is involved in bile acid metabolism, fat degradation and protein digestion and showed contradictory findings. Our inverse relation agrees with a report that Parabacteroides were part of a cluster that increased on a diet rich in animal foods (36) and an inverse association with a vegetarian diet in Thailand (49). However, their abundance was greater with a higher HEI-2005 score in colonic mucosa-associated GM among healthy individuals (10). Bacteroides are known to be associated with consumption of animal protein, a variety of amino acids and saturated fat, which represent a more Westernised diet (5).

An increasing number of studies have indicated that dietary components may influence or shift the composition of the GM with the potential of beneficial health effects, such as prevention of CVD and obesity, potentially through alteration in host immunity and metabolic activity (53). As shown in the current analysis, a high-quality diet, as assessed by the HEI-2015, maintained over many years predicted a GM composition that included several bacteria that have been associated with favourable health effects. As diet quality closer to the time of stool collection had a stronger association with the GM, changes in dietary quality even at an older age may be beneficial (50–51). However, the current results suggest that a consistently high-quality diet over time predicts the most favourable pattern of gut bacteria. Although the number of years between diet quality assessment and stool collection varied, the strength of the associations remained similar across the three questionnaires, supporting the long-term effects of diet on the characteristics of the GM and the stability of dietary habits over time. Alternatively, a certain diet quality at a younger age may be formative in the selection of bacteria that will remain part of a person’s GM (52,53). This idea is supported by reports that sudden changes in diet temporarily modify the composition and diversity of the gut microbiota before reverting to its original state (25).

Although the microbial composition changed in ten subjects randomly controlled for either a high-fat/low-fibre or low-fat/high-fibre diet for 10 days, the dominant genera remained stable over time indicating its strong association with long-term diet (3).

The novelty of the current report lies in the long-term observation with repeated assessments of diet quality and the ability to examine changes in diet quality. The larger sample size and the
updated reference libraries partially confirm the previous findings\(^{69}\) but also make comparisons between the two reports challenging as the different taxonomic classifications of individual bacteria are not analogous. The choice of the HEI-2015 provides a valid and widely used measure of overall diet quality shown to be related to mortality\(^{75,50}\), type 2 diabetes\(^{56}\) and colorectal cancer risk\(^{59}\). However, several limitations have to be considered in the interpretation of the current findings. Foremost, the one-time collection of stool samples that may have been affected by concurrent disease\(^{58}\) or other events, the known measurement errors of FFQ\(^{59}\) and the different sample sizes at each point in time. As the important influence of different medications has been documented\(^{14,60}\), our lack of detailed use at stool collection may have confounded the current findings. Applying 16S sequencing instead of whole genome shotgun sequencing restricted our ability to identify bacteria at the species level. For a better understanding of metabolic mechanism involved in the relation between diet quality and the GM, the imputation of functional pathways with PICRUS\(^{61}\) or similar tools will provide an assessment of community function and insights into predictive metagenomics in future studies.

The most remarkable finding of the current analysis is the stability of the associations of diet quality as assessed many years before stool collection with characteristics of the GM. Participants who improved their diet quality scores between Q1 and Q5 showed microbial patterns that were comparable to individuals with high HEI-2015 scores at Q5. The quality of past diet, assessed as far as 21 years before stool collection, was equally predictive of GM composition as concurrently assessed diet, indicative of the long-term nature of this relation.

Acknowledgements

We thank the participants for their time to participate in the study.

This work was supported by US National Cancer Institute (NCI) grants P01 CA168530, U01 CA164973, P30 CA071789, P30 CA015704 and U1LTR000130.


All authors reviewed and approved the final version of the manuscript.

None of the authors reported a conflict of interest.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S0007114521002968

References


