Long-term Diet Quality and Gut Microbiome Functionality: A Prospective, Shotgun Metagenomic Study among Urban Chinese Adults

Danxia Yu,[1] Yaohua Yang,[1] Jirong Long,[1] Wanghong Xu,[2] Qiuyin Cai,[1] Jie Wu,[1] Hui Cai,[1] Wei Zheng,[1] and Xiao-Ou Shu[1]

[1]Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA and [2]Department of Epidemiology, School of Public Health, Fudan University, Shanghai, China

ABSTRACT

Background: Diet is known to affect human gut microbiome composition; yet, how diet affects gut microbiome functionality remains unclear.

Objective: We compared the diversity and abundance/presence of fecal microbiome metabolic pathways among individuals according to their long-term diet quality.

Methods: In 2 longitudinal cohorts, we assessed participants' usual diets via repeated surveys during 1996–2011 and collected a stool sample in 2015–2018. Participants who maintained a healthy or unhealthy diet (i.e., stayed in the highest or lowest quintile of a healthy diet score throughout follow-up) were selected. Participants were excluded if they reported a history of cancer, cardiovascular disease, diabetes, or hypertension; had diarrhea or constipation in the last 7 d; or used antibiotics in the last 6 mo before stool collection. Functional profiling of shotgun metagenomics was performed using HUMAnN2. Associations of dietary variables and 420 microbial metabolic pathways were evaluated via multivariable-adjusted linear or logistic regression models.

Results: We included 144 adults (mean age = 64 y; 55% female); 66 had an unhealthy diet and 78 maintained a healthy diet. The healthy diet group had higher Shannon α-diversity indexes of microbial gene families and metabolic pathways (both $P < 0.02$), whereas β-diversity, as evaluated by Bray-Curtis distance, did not differ between groups (both $P > 0.50$). At $P < 0.01$ [false discovery rate (FDR) <0.15], the healthy diet group showed enriched pathways for vitamin and carrier biosynthesis (e.g., tetrahydrofolate, acetyl-CoA, and l-methionine) and tricarboxylic acid (TCA) cycle, and increased degradation (or reduced biosynthesis) of certain sugars [e.g., cytidine monophosphate (CMP)-legionaminate, deoxythymidine diphosphate (dTDP)-l-rhamnose, and sucrose], nucleotides, 4-aminobutanoate, methylglyoxal, sulfate, and aromatic compounds (e.g., catechol and toluene). Meanwhile, several food groups were associated with the CMP-legionaminate biosynthesis pathway at FDR <0.05.

Conclusions: In a small longitudinal study of generally healthy, older Chinese adults, we found long-term healthy eating was associated with increased α-diversity of microbial gene families and metabolic pathways and altered symbiotic functions relevant to human nutrition and health. *Curr Dev Nutr* 2021;5:nzab026.

Keywords: gut microbiota, shotgun metagenomics, nutrition, diet quality, dietary pattern, prospective cohort study, epidemiology

© The Author(s) 2021. Published by Oxford University Press on behalf of the American Society for Nutrition. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Manuscript received January 21, 2021. Initial review completed March 15, 2021. Revision accepted March 24, 2021. Published online April 2, 2021.

The Shanghai Women's Health Study was funded by UM1 CA182910 to WZ, and the Shanghai Men's Health Study was funded by UM1 CA173640 to X-OS from the National Cancer Institute at the National Institutes of Health. DY was supported by Faculty Research Scholars Program/KL2 from the Vanderbilt University Medical Center. The analysis used computing resources at the Advanced Computing Center for Research and Education of Vanderbilt University (NIH S10 Shared Instrumentation Grant 1S10OD023680–01).

Author disclosures: The authors report no conflicts of interest.

Supplemental Figures 1 and 2 and Supplemental Tables 1–3 are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/cdn/.

Address correspondence to DY (e-mail: danxia.yu@vanderbilt.edu).

Abbreviations used: CMP, cytidine monophosphate; dTDP, deoxythymidine diphosphate; FDR, false discovery rate; GABA, 4-aminobutanoate; GLM, general linear model; HDS, healthy diet score; PERMANOVA, permutational multivariate ANOVA; SMHS, Shanghai Men's Health Study; SWHS, Shanghai Women's Health Study; TCA, tricarboxylic acid.

Introduction

Diet is one of the most important, modifiable determinants of human health. Emerging research in recent years has indicated that diet may affect human health, in part through modifying the gut microbiota (1, 2). The commensal gut microbes are usually harmless or beneficial to humans; they protect the host against pathogens, regulate immune responses, synthesize essential nutrients, and play a major role in metabolizing xenobiotics, including various dietary components, environmental chemicals, and pharmaceuticals (3, 4). On the other hand, a disrupted gut microbiota, characterized by loss of diversity, loss of commensals, gain of pathogens, and change in functionality, has been linked to multiple human diseases, from gastrointestinal diseases to cancer, cardiometabolic diseases, and neurobehavioral disorders (2, 5). An improved understanding of how diet influences the composition and functionality of the human gut microbiota will advance nutrition science and

inform potential diet- or microbiota-based interventions to improve human health.

An increasing number of population-based studies have collected fecal samples to examine habitual diets' influence on the gut microbiota (6–12). However, due to the high cost of shotgun metagenomics, most of the studies, including our prior research, conducted 16S rRNA sequencing and can only evaluate associations of dietary variables with microbial taxonomic composition and diversity. Meanwhile, although a few population studies conducted shotgun metagenomics to assess microbial functional diversity and pathways, few had comprehensive long-term dietary data to evaluate habitual food intakes and overall dietary patterns (10, 11, 13, 14). Given the functional redundancy of gut microbiota in which taxonomically unrelated species may perform similar functions, it is important to investigate microbial functionality beyond taxonomy (5, 15). Understanding how habitual diets affect microbial pathways may reveal novel biological mechanisms underlying the diet–gut microbiota–host interactions (11).

In the present study, we leveraged resources of 2 population-based, prospective cohorts—the Shanghai Women's Health Study (SWHS) and Shanghai Men's Health Study (SMHS)—including longitudinally collected dietary data and recently collected fecal samples. We compared the diversity, abundance, and presence of microbial metabolic pathways among generally healthy participants who had a long-term healthy or unhealthy diet, determined by a healthy diet score (HDS). The HDS incorporates 8 major food groups (fruits, vegetables, dairy, fish and seafood, nuts and legumes, refined grains, red meat, and processed meat) and is significantly associated with a reduced risk of type 2 diabetes and increased species-level α-diversity and abundance of multiple fiber-fermenting bacteria in the SWHS and SMHS (12, 16).

Methods

Study population

The SWHS and SMHS are population-based prospective cohort studies conducted in Shanghai, China, which enrolled 74,940 women during 1996–2000 (response rate = 93%) and 61,491 men during 2002–2006 (response rate = 74%), respectively (17, 18). In-person interviews were conducted at baseline and follow-up visits every 2–4 y (follow-up rates >92% in both cohorts) to collect and update participants' sociodemographic, disease history, diet, lifestyle, and anthropometric information. Biospecimens were also collected, including blood, urine, and oral rinse samples at baseline and stool samples at the latest follow-up (see details in the section below). Participants' death and disease outcomes were also updated via linkages to Shanghai Vital Statistics and Cancer Registries (completion rate >99% in both cohorts). The SWHS and SMHS were approved by Institutional Review Boards of the Shanghai Cancer Institute and Vanderbilt University Medical Center; informed consent was obtained from each participant.

Stool sample collection

During the cohort follow-up in 2015–2018, we collected a stool sample using the 95% ethanol method from ∼10,000 living and willing cohort participants, as previously described (12). Briefly, participants were provided with stool sample collection kits, each of which included disposable collection tools, a tube containing 5-mL 95% ethanol and glass beads, an instruction sheet, and a sample collection form. They were asked to collect a peanut-sized stool into the tube and shake the tube until the sample was well mixed with ethanol. The collection form asked the date and time of stool collection, use of antibiotics and other medications in the last 7 d and 6 mo, bowel movement frequency, and any diet or body weight changes in the last 7 d before stool collection. The samples were stored at room temperature and transported to a research laboratory within 24 h, and then placed into aliquots and stored at −80°C.

Assessment of long-term diet

Semiquantitative FFQs were used to estimate participants' usual food intake over the past 12 mo. The FFQs used in the SWHS and SMHS were very similar, and contained 77 and 81 common food items, respectively. Both FFQs were validated against 24-h dietary recalls that were administered monthly over 12 mo; correlation coefficients for major food groups were 0.41–0.66 in SWHS and 0.42–0.72 in SMHS, suggesting good validity of the FFQs (19, 20). The FFQ was administered 3 times in the SWHS, capturing food intakes during 1996–2011, and twice in the SMHS, capturing intakes during 2002–2011. We generated an HDS to evaluate overall diet quality, incorporating 8 food groups with equal weights: fruits, vegetables (excluding potatoes), dairy, fish and seafood, nuts and legumes, refined grains, red meat, and processed meat (16). Based on energy-adjusted, sex-specific intake quintiles, the first 5 groups were assigned by ascending values (1 to 5) and the last 3 groups were assigned by descending values (5 to 1). Their sum represented the HDS, ranging from 8 to 40; a higher score reflects a higher diet quality. We calculated the HDS for each participant at each FFQ and generated a cumulative average HDS reflecting long-term diet quality. Participants who stayed in the highest quintile of HDS at each FFQ were defined as having a long-term healthy diet, whereas participants who stayed in the lowest quintile of HDS at each FFQ were defined as having a long-term unhealthy diet. We have previously reported in SWHS and SMHS that a high HDS, especially if maintained for a long term, was associated with a 15–25% reduced risk of type 2 diabetes (16). Furthermore, in a recent 16S rRNA metagenomics study, we reported that a high HDS was associated with an increased species-level Shannon index and abundances of fiber-fermenting bacteria (e.g., genera *Coprococcus*, *Faecalibacterium*, and *Bifidobacterium*) (12).

Participant selection

Participants were selected based on their availability of fecal microbial DNA as well as their disease history, recent use of medications, recent bowel movements, and long-term diet quality. Among 3194 participants whose fecal DNA was extracted in the aforementioned 16S study, we excluded 692 participants who did not have long-term dietary data (those who only had 1 dietary survey or extreme calorie intake, i.e., <500 or >3500 kcal/d for women and <800 or >4200 kcal/d for men). We also excluded participants who reported a history of cancer, cardiovascular diseases, diabetes, or hypertension ($n = 1026$); who were aged >80 y ($n = 141$); who had antibiotic use in the last 6 mo ($n = 68$); who had constipation or diarrhea ($n = 139$; i.e., ≤2 or ≥21 bowel movements in the last 7 d); or who had use of antidiabetic, antihypertensive, or antidyslipidemic medications or substantial changes in body weight or diet in the last 7 d ($n = 125$) at the time of stool collection. After those exclusions, we selected participants who had a healthy or unhealthy long-term diet using repeated HDS data, which yielded

TABLE 1 Participant characteristics by long-term diet quality[1]

Characteristics	Unhealthy diet group (n = 66)	Healthy diet group (n = 78)	P
Healthy diet score	17.0 ± 1.8	32.3 ± 2.0	<0.0001
Age at stool collection, y	63.8 ± 6.3	64.5 ± 6.9	0.53
Female, %	56.1	55.1	0.91
High income, %	3.0	18.0	0.0003
Current smoking cigarette, %	37.8	23.1	0.04
Current alcohol drinking, %	21.2	21.8	0.93
Leisure-time exercise, %	40.9	53.9	0.02
BMI, kg/m^2	24.1 ± 3.2	23.7 ± 3.2	0.52
Dietary intakes			
Total energy, kcal/d	1844 ± 412	1786 ± 388	0.38
Fruit, g/d	76.8 ± 57.6	294 ± 151	<0.0001
Vegetables, g/d	221 ± 84.3	478 ± 166	<0.0001
Dairy products, g/d	39.3 ± 72.7	201 ± 236	<0.0001
Fish and seafood, g/d	22.6 ± 13.1	85.6 ± 53.4	<0.0001
Legumes and nuts, g/d	15.7 ± 9.3	31.5 ± 12.2	<0.0001
Refined grains, g/d	372 ± 59.9	277 ± 63.2	<0.0001
Red meat, g/d	63.3 ± 28.9	42.5 ± 22.4	<0.0001
Processed meat, time/wk	0.62 ± 0.79	0.25 ± 0.44	<0.0001
Shannon index of gene families	5.44 ± 0.09	5.48 ± 0.13	0.03
Shannon index of metabolic pathways	5.10 ± 0.21	5.18 ± 0.23	0.03

[1]Values are means ± SDs unless otherwise indicated. t tests for continuous variables and chi-square tests for categorical variables were used. All characteristics were updated during cohort follow-up, and the latest available data are presented, except for dietary variables, which were the cumulative averages using data from repeated FFQs during follow-ups.

49 men and 51 women in the healthy diet group and 36 men and 39 women in the unhealthy diet group. Among those 166 samples sent for shotgun metagenomics, 21 were further excluded due to low microbial DNA amount and 1 failed to pass sequencing quality control. Finally, 144 participants were included in our analysis. A participant selection flowchart is provided in **Supplemental Figure 1**.

Shotgun metagenomics

DNA was extracted from stool samples using Qiagen's DNeasy PowerSoil kit. DNA libraries were constructed using the Illumina Nextera XT kit. Sequencing was performed on the Illumina HiSeq X TEN platform using a 150 paired-end strategy, generating a median of 27.3 million raw reads per sample (range: 20.1 to 47.0 million reads). Trimmomatic (v0.39) was utilized to trim low-quality bases in raw reads, then reads with <70% of original read length were discarded (21). Bowtie2 (v2.3.5) was used to remove reads that could be mapped onto the human genome (version GRCh38) (22). After quality-trimming and removal of human reads, on average, 27.1 million reads were retained for each sample (range: 20.1 to 46.9 million reads). These reads were then used for functional profiling using HUMAnN2 (v2.8.1) to estimate the relative abundance of microbial gene families and MetaCyC pathways within each sample. For both gene families and metabolic pathways, α-diversity was evaluated by Shannon index and β-diversity was evaluated by Bray-Curtis distance (23). Statistical analyses were performed for α- and β-diversity indexes of gene families and metabolic pathways and relative abundance of microbial metabolic pathways.

Statistical analysis

Participant characteristics between healthy and unhealthy diet groups were compared by using t tests for continuous variables and chi-square tests for categorical variables. The associations between diet groups and α-diversity indexes were evaluated using general linear models (GLMs), adjusting for sequencing depth, sociodemographics (age at stool sample collection, sex, and income), lifestyle (cigarette smoking, alcohol drinking, and leisure-time exercise), BMI, and total calorie intake. We used the latest available information on income, lifestyle, and BMI, and average calorie intake from repeated FFQs. The associations between diet groups and Bray-Curtis distance were evaluated using permutational multivariate ANOVA (PERMANOVA) with 999 permutations, adjusting for sequencing depth and the same covariates. A pathway was considered present if its relative abundance was $\geq 1 \times 10^{-5}$ (0.001%). We defined pathways as common if present in >50% of samples and rare if present in 10–50% of samples; pathways present in <10% of samples were excluded from analysis as the sample size was too small. For common pathways, centered log-ratio transformation was applied to normalize relative abundance with zeros imputed by adding 1 pseudo count to the total reads. Then GLM was used to evaluate their associations with dietary variables, controlling for the covariates listed above. For rare pathways, logistic regression models were used to compare their presence between diet groups with adjustment for the covariates. Based on crude P values of both common and rare pathways, Benjamini and Hochberg false discovery rates (FDRs) were calculated; an FDR <0.10 was considered significant given the small sample size of the current study. We assessed potential effect modifications of participants' age, sex, lifestyle, and obesity status on diet–pathway associations by adding an interaction term of these covariates with diet to the regression model. Statistical analyses were performed using SAS (version 9.4; SAS Institute) and R (version 3.6.3; R Foundation for Statistical Computing).

Results

Table 1 shows the characteristics of study participants with a long-term unhealthy or healthy diet. The HDS and intakes of all 8 food groups were

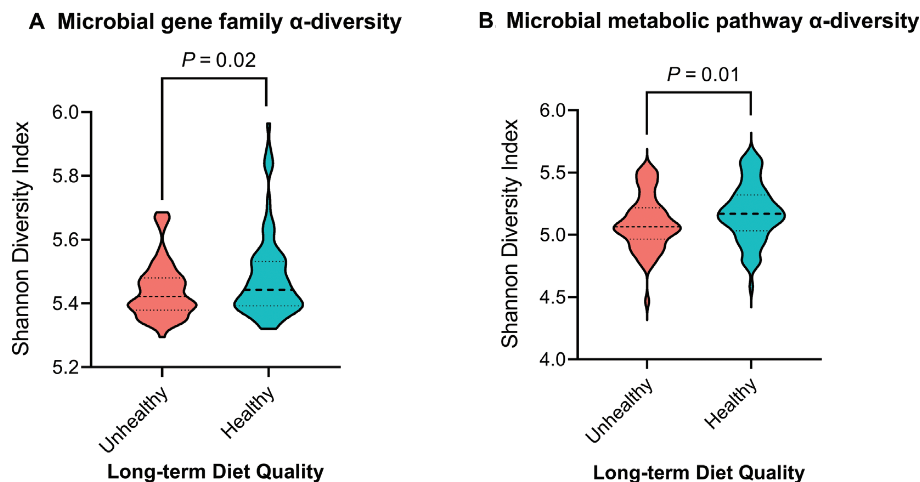

FIGURE 1 Fecal microbial gene family and metabolic pathway α-diversity by long-term diet quality. Violin plots of Shannon diversity indexes of microbial gene families (A) and metabolic pathways (B); P values were obtained from general linear models adjusting for sequencing depth, age at stool sample collection, sex, income, cigarette smoking, alcohol drinking, leisure-time exercise, BMI, and total calorie intake.

substantially different between groups, but total calorie intake was not. Two groups had similar distributions of age (mean: ~64 y), sex (~55% female), and BMI (kg/m^2; mean: ~24). However, the healthy diet group had a significantly higher income and a higher rate of leisure-time exercise and a lower rate of cigarette smoking than the unhealthy diet group.

Compared with the unhealthy diet group, the healthy diet group had small, but significant increases in Shannon indexes of both microbial gene families and metabolic pathways ($P = 0.02$ and 0.01, respectively; **Figure 1**). However, there were no significant differences between diet groups in Bray-Curtis distances of gene families or metabolic pathways (**Supplemental Figure 2**). The diet group explained 0.84% of variance in microbial gene families and 0.75% of variance in metabolic pathways (adjusted PERMANOVA, $P = 0.30$ and 0.29, respectively).

When examined individually, among a total of 366 common and 54 rare pathways, 20 common and 4 rare pathways were associated with diet group at $P < 0.01$ (**Table 2**), and an additional 44 common pathways and 1 rare pathway were associated with diet group at $0.01 < P < 0.05$ (**Supplemental Table 1**); all of these associations had an FDR between 0.10 to 0.15. The most significant results included associations of healthy diet with reduced pathway abundances for biosyntheses of nucleotide sugars (e.g., CMP-legionaminate and dTDP-L-rhamnose), nucleotides (e.g., guanosine and adenosine), and fatty acids, and increased pathway abundances for biosyntheses of tetrahydrofolate, acetyl-CoA, L-methionine, and dTDP-N-acetylthomosamine and for degradation of sucrose, 4-aminobutanoate (GABA), methylglyoxal, and sulfate. Meanwhile, the healthy diet group had a higher prevalence of pathways for degrading aromatic compounds (e.g., catechol and toluene). The adjusted mean relative abundances (95% CIs) of 20 common pathways and the prevalence of 4 rare pathways with $P < 0.01$ are shown in **Figures 2** and **3**, respectively. The MetaCyC classes of those pathways are summarized in **Supplemental Table 2**. We did not find significant effect modifications by participants' age, sex, lifestyle (smoking, drinking, and exercise), or obesity status on the diet–microbial pathway associations.

We then evaluated the associations of 8 major food groups with microbial metabolic pathways. At $P < 0.01$, many food groups were associated with the CMP-legionaminate biosynthesis pathway, including inverse associations for fruits, vegetables, fish/seafood, and nuts/legumes, and a positive association for refined grains (**Supplemental Table 3**). Of note, associations of refined grains, fruits, and fish/seafood with CMP-legionaminate biosynthesis reached an FDR <0.05. Meanwhile, dairy intake was associated with reduced tetrapyrrole biosynthesis at FDR = 0.03.

Discussion

Among generally healthy, older, urban Chinese adults, we found that a long-term healthy diet is associated with increased α-diversity of fecal microbial gene families and metabolic pathways and enhanced pathways related to cofactor, carrier, and vitamin biosynthesis (e.g., folate, acetyl-CoA, and L-methionine) and tricarboxylic acid (TCA) cycle, but increased degradation (or reduced biosynthesis) of certain sugars, nucleotides, amines, methylglyoxal, sulfate, and aromatic compounds. These findings support the influence of habitual diets on gut microbial functionality and reveal potential microbial pathways through which long-term healthy or unhealthy diets may subsequently affect human health.

Diet quality and gut microbial functional diversity

We observed an increased microbial functional diversity in the healthy diet group. Diversity is generally considered a feature of healthy gut microbiota, as higher diversity usually means a more resilient and efficient ecosystem (5, 15, 24). In line with our finding, the LifeLines DEEP cohort, a population-based study of 1135 Dutch adults, reported that more frequent intakes of fruit and tea and lower intakes of bread, snacks, and soda were associated with higher microbial gene richness (10), although they did not evaluate an overall diet quality. Meanwhile, like our previ-

TABLE 2 Fecal microbiome metabolic pathways associated with long-term diet quality[1]

	B	SE	P	FDR
Common pathways				
PWY-6749: CMP-legionaminate biosynthesis I	– 0.604	0.175	0.0007	0.10
DTDPRHAMSYN-PWY: dTDP-l-rhamnose biosynthesis I	– 0.171	0.051	0.001	0.13
PWY-6125: super-pathway of guanosine nucleotides de novo biosynthesis II	– 0.153	0.048	0.002	0.13
PWY-7228: super-pathway of guanosine nucleotides de novo biosynthesis I	– 0.159	0.053	0.003	0.13
PWY-7208: super-pathway of pyrimidine nucleobases salvage	– 0.176	0.059	0.003	0.13
PWY-6612: super-pathway of tetrahydrofolate biosynthesis	0.508	0.173	0.004	0.13
FOLSYN-PWY: super-pathway of tetrahydrofolate biosynthesis and salvage	0.482	0.165	0.004	0.13
PWY-5384: sucrose degradation IV (sucrose phosphorylase)	0.487	0.167	0.004	0.13
PWY-5173: super-pathway of acetyl-CoA biosynthesis	0.643	0.222	0.004	0.13
PWY-5345: super-pathway of l-methionine biosynthesis (by sulfhydrylation)	0.532	0.187	0.005	0.13
PWY-5022: 4-aminobutanoate degradation V	0.553	0.197	0.006	0.13
FASYN-INITIAL-PWY: super-pathway of fatty acid biosynthesis initiation (*Escherichia coli*)	– 0.170	0.061	0.006	0.13
PWY-6126: super-pathway of adenosine nucleotides de novo biosynthesis II	– 0.102	0.037	0.007	0.13
PWY-7229: super-pathway of adenosine nucleotides de novo biosynthesis I	– 0.094	0.034	0.007	0.13
PWY-7315: dTDP-N-acetylthomosamine biosynthesis	0.479	0.177	0.008	0.14
METHGLYUT-PWY: super-pathway of methylglyoxal degradation	0.427	0.159	0.008	0.14
SO4ASSIM-PWY: sulfate reduction I (assimilatory)	0.690	0.259	0.009	0.14
GLYCOLYSIS: glycolysis I (from glucose 6-phosphate)	– 0.116	0.043	0.009	0.14
SULFATE-CYS-PWY: super-pathway of sulfate assimilation and cysteine biosynthesis	0.540	0.204	0.009	0.14
PWY-6969: TCA cycle V (2-oxoglutarate:ferredoxin oxidoreductase)	0.209	0.081	0.01	0.15
Rare pathways				
PWY-5417: catechol degradation III (ortho-cleavage pathway)	1.356	0.463	0.003	0.13
PWY-5431: aromatic compounds degradation via B-ketoadipate	1.356	0.463	0.003	0.13
PWY-5181: toluene degradation III (aerobic) (via p-cresol)	1.162	0.420	0.006	0.13
PWY-6344: l-ornithine degradation II (Stickland reaction)	0.826	0.301	0.006	0.13
PWY-6185: 4-methylcatechol degradation (ortho cleavage)	1.204	0.458	0.009	0.15

[1]General linear models were used to evaluate the healthy diet group in association with the relative abundance of common pathways after centered log-ratio transformation and zero imputation. Logistic regression models were used to evaluate the healthy diet group in association with the presence of rare pathways. Covariates included age at sample collection, sex, income, cigarette smoking, alcohol drinking, leisure-time exercise, BMI, and total calorie intake. Results with $P \leq 0.01$ are shown. CMP, cytidine monophosphate; CoA, coenzyme A; dTDP, deoxythymidine diphosphate; FDR, false discovery rate; PWY, pathway; TCA, tricarboxylic acid.

ous finding on microbial taxonomic diversity, the effect of diet quality on microbial functional diversity was modest, although statistically significant (12). It is important to note that all of our participants, regardless of the healthiness of their diets, were free of major diseases, had normal bowel movements, and were without recent use of antibiotics or other medications, nor did they eat specific restricted diets [e.g., vegan, keto, or low-FODMAP (-fermentable oligosaccharides, disaccharides, monosaccharides, and polyols) diet]; therefore, their gut microbiota should be in stable, balanced states. Our results support previous findings that community-level measures of gut microbiota (i.e., α- or β-diversity) are only modestly related to usual diets among generally healthy populations (25, 26).

Diet quality and gut microbial metabolic pathways

Among specific pathways, the most significant findings were related to decreased biosyntheses of CMP-legionaminate and dTDP-L-rhamnose, 2 nucleotide sugars, in the healthy diet group. A reduced CMP-legionaminate biosynthesis pathway was also related to multiple food groups, including higher intakes of fruits, vegetables, fish/seafood, and nuts/legumes and a lower intake of refined grains. Legionaminate and L-rhamnose are building blocks of the glycan component (e.g., O-antigens) of lipopolysaccharide (LPS) of gram-negative bacteria. Many pathogenic bacteria (e.g., *Legionella pneumophila*, *Campylobacter coli*, and *Salmonella enterica*) display O-antigens on their surface to help them invade host cells (27, 28). Bacteria can also release LPS into the environment, which then triggers host inflammatory cascades (29). Our findings may suggest a mechanism underlying the anti-inflammatory effect of healthy diets via reducing gut bacteria–derived, proinflammatory nucleotide sugars. However, we did not observe a significant change in the overall LPS biosynthesis super-pathway ($P = 0.30$), which may be because we also observed counteractively increased biosyntheses of other LPS-contributing molecules such as dTDP-N-acetylthomosamine and lipid IV_A (tetra-acylated lipid A) ($P = 0.008$ and 0.03, respectively). The evidence regarding the gut microbiome LPS biosynthesis pathway and human health remains inconsistent. Studies have reported higher abundances of LPS pathway in the fecal microbiome of morbidly obese individuals than in lean individuals (BMI: 37 vs. 20) and among colorectal adenoma/carcinoma patients than healthy controls (30, 31), but also higher abundances among healthy controls than rheumatoid arthritis patients and after metformin treatment among diabetic patients with no concurrent change in circulating C-reactive protein (32, 33). It is unclear to what extent the metagenomic abundance of this pathway affects secreted LPS concentration and, more importantly, the host's immune response to LPS exposure. Unfortunately, we did not collect blood samples at the time of stool collection for participants involved in this study. Future studies incorporating fecal metagenomics with blood concentrations of LPS or LPS-binding protein and inflammatory markers are needed.

Next, we observed pathways related to cofactor, carrier, and vitamin biosynthesis enriched in the healthy diet group, involving tetrahy-

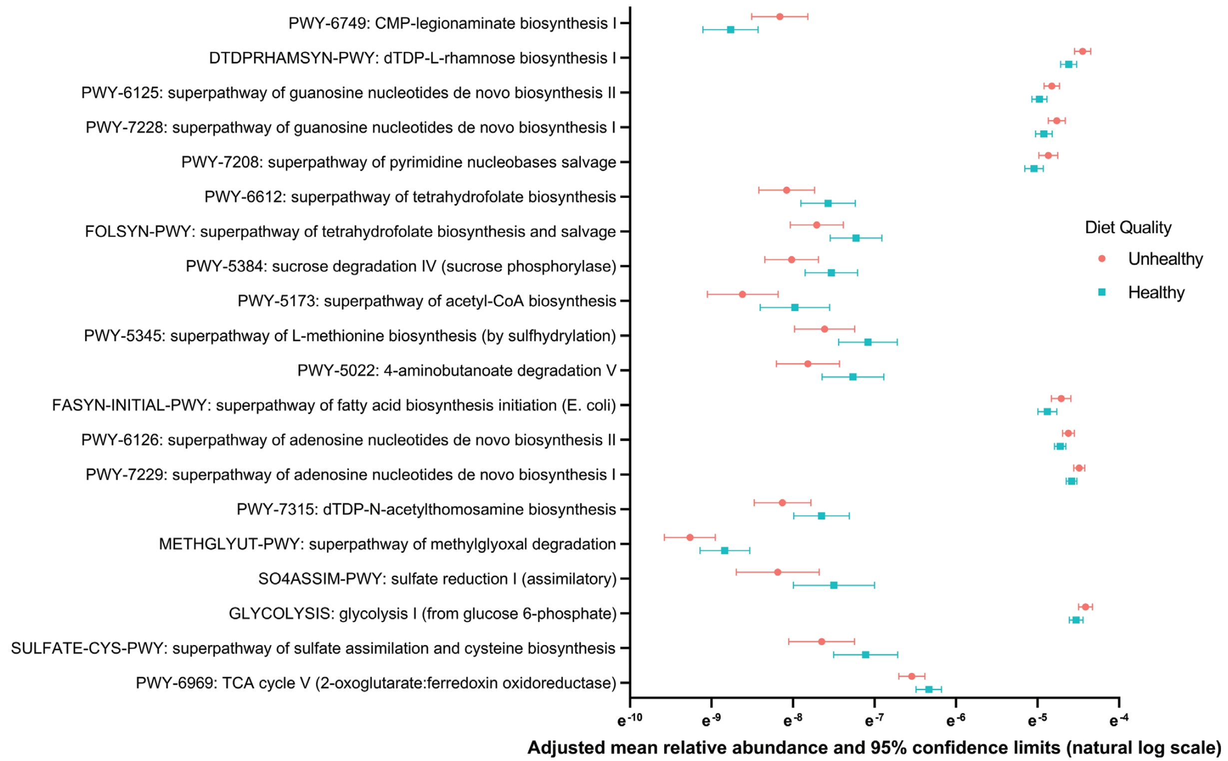

FIGURE 2 Common fecal microbial metabolic pathway relative abundance by long-term diet quality[1]. [1]Common pathways were defined as those present in >50% of samples. Mean relative abundances (95% confidence limits) were obtained from general linear models, adjusting for age at sample collection, sex, income, cigarette smoking, alcohol drinking, leisure-time exercise, BMI, and total calorie intake. Results with $P \leq 0.01$ are shown. A natural log scale is used for better visualization. CMP, cytidine monophosphate; CoA, coenzyme A; dTDP, deoxythymidine diphosphate; PWY, pathway; TCA, tricarboxylic acid.

drofolate, acetyl-CoA, L-methionine, and also NAD, *S*-adenosyl-L-methionine, heme, and pantothenate (vitamin B-5) (see Supplemental Table 1). These molecules are involved in biochemical reactions as acceptors or donors of electrons, hydrons, and methyl or acyl groups, and hence may affect energy metabolism, glycolysis, glucogenesis, oxidative stress, and epigenetic regulation. Previous case-control studies of type 2 diabetes, liver cirrhosis, and rheumatoid arthritis among Chinese and European adults have reported higher abundances of these pathways among healthy controls than in patients (32, 34, 35), suggesting a potential role of gut microbial cofactor/carrier/vitamin biosynthesis against host metabolic and inflammatory diseases. However, it is unclear to what extent gut microbial biosynthesis contributes to human needs, given that most of those molecules can be synthesized by the human body or ingested from foods. Nevertheless, vitamin deficiencies,

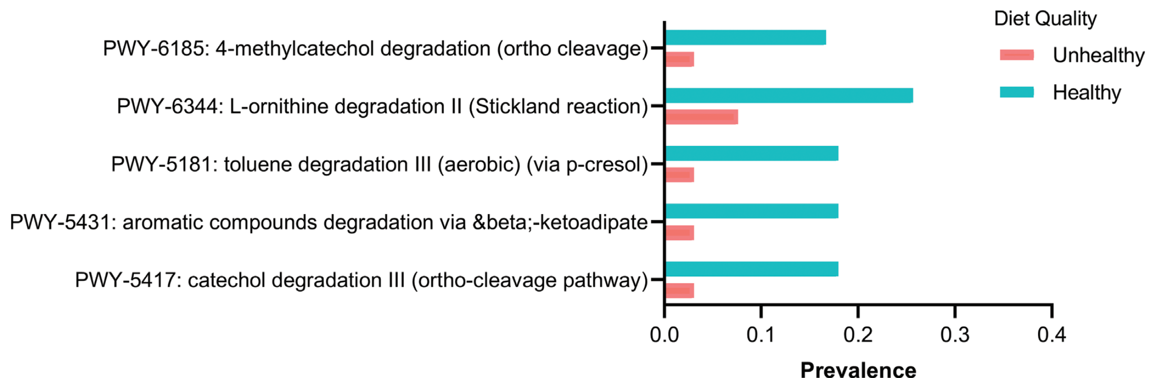

FIGURE 3 Rare fecal microbial metabolic pathway prevalence by long-term diet quality. Rare pathways were defined as those present in 10–50% of samples. Results with $P \leq 0.01$ are shown. PWY, pathway.

including folate deficiency, have been observed after antibiotic use, implying that commensal microbial vitamin production can be significant (36, 37).

Finally, we observed a series of pathway changes in the healthy diet group regarding microbial degradation and fermentation of various substrates as sources of energy and nutrients or means of detoxification, including increased degradation of sugars (e.g., sucrose, fucose, and D-glucarate), amines (e.g., GABA, ornithine, and allantoin), methylglyoxal, sulfate, and aromatic compounds (e.g., catechol and toluene), which can generate energy, pyruvate, lactate, SCFAs (acetate, propanoate, and butanoate), and other metabolites (e.g., succinate and ethanol). Supporting our observational findings, intervention studies have found that a 5-d plant-based diet or an 8-wk Mediterranean diet increased abundances of microbial genes for carbohydrate and amino acid degradation, particularly glutamate-family amino acids: glutamate, arginine, and GABA (38, 39). Metformin treatment may also increase microbial pathways for carbohydrate and amino acid degradation (33). Meanwhile, in a unique study comparing the gut microbiome of centenarians (99–109 y) with young and middle-aged Italians (40), centenarians had significantly enriched microbial genes for aromatic xenobiotics degradation (including toluene, an industrial chemical used in paint and glue). Furthermore, research has increasingly focused on gut microbiota–related metabolites, which can mediate gut microbiota–host cross-talks, enter the host circulatory system, and affect multiple aspects of human health (41). For example, SCFAs are known for their potential health benefits to reduce obesity, inflammation, and insulin resistance (42, 43), whereas methylglyoxal is a toxic byproduct of glycolysis, which can cause endothelial cell loss and inflammation (44). Our current findings regarding increased fermentation and generation of lactate and SCFAs are in line with our previous results from a 16S study that higher diet quality was associated with higher abundances of fiber-fermenting, SCFA-/lactate-generating bacteria (e.g., *Faecalibacterium* and *Bifidobacterium*) (12). Together, previous studies and our current study have revealed certain microbial functions that may mediate the effects of habitual diets on human health and aging. Prospective studies are needed to evaluate potential causal relations of those gut microbial functions with health outcomes.

Strengths and limitations

To our knowledge, this is one of the few longitudinal, population-based studies on the influence of habitual diets on the gut microbiome and one of the first reports on overall diet quality and microbial functionality among Asian populations. Our strengths include the use of validated, repeated dietary surveys to assess participants' long-term diets and comprehensive metadata (e.g., other lifestyle factors, disease history, and medications) to minimize and control for potential confounding. Our limitations include a small sample size, and thus studies with larger sample sizes are needed to examine further the associations of habitual diets with microbial functionality. Second, although we used repeated dietary surveys and adjusted for potential confounders, we cannot rule out measurement errors and residual confounding. We also acknowledge a lack of dietary assessment proximal to stool collection as a limitation. We aimed to evaluate the effect of long-term, overall diet and thus selected participants who had maintained a stable diet quality during follow-up and excluded those who reported substantial recent diet changes; therefore, potential misclassification in the diet group should be minimized. Third, we could not further evaluate the potential health implications of those observed diet–microbial pathway associations due to excluding participants with major diseases and lack of disease risk biomarkers in the current study. Also, given that our current study participants were highly selected, our current findings may not be directly applicable to other populations (e.g., younger or rural Chinese adults or those with metabolic diseases). Finally, although shotgun metagenomics can imply microbial composition and functional potential, it is ideal to incorporate other data (e.g., fecal and/or blood metabolomics) and evaluate microbial functional activities or circulating microbial metabolites that may directly affect host physiology. The relations between diet, gut microbiota, and disease risk are yet to be evaluated in our future studies after enough incident cases have accrued.

Conclusions

In summary, among 2 groups of generally healthy, older, urban Chinese adults who had a healthy or an unhealthy long-term diet, we found the healthy diet to be associated with increased α-diversity of fecal microbial gene families and metabolic pathways and enhanced pathways related to cofactor, carrier, and vitamin biosynthesis and TCA cycle, and increased degradation (or reduced biosynthesis) of certain sugars, sugar nucleotides, amines, and aromatic compounds. Future large-scale prospective studies are needed to validate our findings and evaluate those gut microbial metabolic pathways for their associations with human disease risk.

Acknowledgments

The authors' responsibilities were as follows—DY and X-OS: designed the research; DY, YY, JL, WX, QC, JW, HC, WZ, and X-OS: conducted the research; DY and YY: analyzed data; DY: wrote the manuscript and had primary responsibility for the final content; and all authors: provided critical input and read and approved the final manuscript.

Data Availability

Data described in the manuscript will be made available upon research study application and approval by the cohort committees.

References

1. Aron-Wisnewsky J, Clement K. The gut microbiome, diet, and links to cardiometabolic and chronic disorders. Nat Rev Nephrol 2016;12(3):169–81.
2. Gentile CL, Weir TL. The gut microbiota at the intersection of diet and human health. Science 2018;362(6416):776–80.
3. Hooper LV, Gordon JI. Commensal host-bacterial relationships in the gut. Science 2001;292(5519):1115–18.
4. Koppel N, Maini Rekdal V, Balskus EP. Chemical transformation of xenobiotics by the human gut microbiota. Science 2017;356(6344):eaag2770.
5. Levy M, Kolodziejczyk AA, Thaiss CA, Elinav E. Dysbiosis and the immune system. Nat Rev Immunol 2017;17(4):219–32.
6. Bowyer RCE, Jackson MA, Pallister T, Skinner J, Spector TD, Welch AA, Steves CJ. Use of dietary indices to control for diet in human gut microbiota studies. Microbiome 2018;6(1):77.
7. Maskarinec G, Hullar MAJ, Monroe KR, Shepherd JA, Hunt J, Randolph TW, Wilkens LR, Boushey CJ, Le Marchand L, Lim U, et al. Fecal microbial

diversity and structure are associated with diet quality in the Multiethnic Cohort Adiposity Phenotype Study. J Nutr 2019;149(9):1575–84.
8. Kaplan RC, Wang Z, Usyk M, Sotres-Alvarez D, Daviglus ML, Schneiderman N, Talavera GA, Gellman MD, Thyagarajan B, Moon JY, et al. Gut microbiome composition in the Hispanic Community Health Study/Study of Latinos is shaped by geographic relocation, environmental factors, and obesity. Genome Biol 2019;20(1):219.
9. Ericson U, Brunkwall L, Hellstrand S, Nilsson PM, Orho-Melander M. A health-conscious food pattern is associated with prediabetes and gut microbiota in the Malmo Offspring Study. J Nutr 2020;150(4):861–72.
10. Zhernakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T, Mujagic Z, Vila AV, Falony G, Vieira-Silva S, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science 2016;352(6285):565–9.
11. Kurilshikov A, van den Munckhof ICL, Chen L, Bonder MJ, Schraa K, Rutten JHW, Riksen NP, de Graaf J, Oosting M, Sanna S, et al. Gut microbial associations to plasma metabolites linked to cardiovascular phenotypes and risk. Circ Res 2019;124(12):1808–20.
12. Yu D, Nguyen SM, Yang Y, Xu W, Cai H, Wu J, Cai Q, Long J, Zheng W, Shu XO. Long-term diet quality is associated with gut microbiome diversity and composition among urban Chinese adults. Am J Clin Nutr 2021;113(3): 684–94.
13. Pedersen HK, Gudmundsdottir V, Nielsen HB, Hyotylainen T, Nielsen T, Jensen BA, Forslund K, Hildebrand F, Prifti E, Falony G, et al. Human gut microbes impact host serum metabolome and insulin sensitivity. Nature 2016;535(7612):376–81.
14. Vangay P, Johnson AJ, Ward TL, Al-Ghalith GA, Shields-Cutler RR, Hillmann BM, Lucas SK, Beura LK, Thompson EA, Till LM, et al. US immigration westernizes the human gut microbiome. Cell 2018;175(4):962–72, e10.
15. Moya A, Ferrer M. Functional redundancy-induced stability of gut microbiota subjected to disturbance. Trends Microbiol 2016;24(5):402–13.
16. Yu D, Zheng W, Cai H, Xiang YB, Li H, Gao YT, Shu XO. Long-term diet quality and risk of type 2 diabetes among urban Chinese adults. Diabetes Care 2018;41(4):723–30.
17. Shu XO, Li H, Yang G, Gao J, Cai H, Takata Y, Zheng W, Xiang YB. Cohort profile: the Shanghai Men's Health Study. Int J Epidemiol 2015;44(3):810–18.
18. Zheng W, Chow WH, Yang G, Jin F, Rothman N, Blair A, Li HL, Wen W, Ji BT, Li Q, et al. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. Am J Epidemiol 2005;162(11):1123–31.
19. Shu XO, Yang G, Jin F, Liu D, Kushi L, Wen W, Gao YT, Zheng W. Validity and reproducibility of the food frequency questionnaire used in the Shanghai Women's Health Study. Eur J Clin Nutr 2004;58(1):17–23.
20. Villegas R, Yang G, Liu D, Xiang YB, Cai H, Zheng W, Shu XO. Validity and reproducibility of the food-frequency questionnaire used in the Shanghai Men's Health Study. Br J Nutr 2007;97(5):993–1000.
21. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30(15):2114–20.
22. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9(4):357.
23. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci 2003;14(6):927–30.
24. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. Nature 2012;489(7415):220–30.
25. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, Kurilshikov A, Bonder MJ, Valles-Colomer M, Vandeputte D, et al. Population-level analysis of gut microbiome variation. Science 2016;352(6285):560–4.
26. Willis HJ, Slavin JL. The influence of diet interventions using whole, plant food on the gut microbiome: a narrative review. J Acad Nutr Diet 2020;120(4):608–23.
27. Schoenhofen IC, Vinogradov E, Whitfield DM, Brisson JR, Logan SM. The CMP-legionaminic acid pathway in Campylobacter: biosynthesis involving novel GDP-linked precursors. Glycobiology 2009;19(7):715–25.
28. Samuel G, Reeves P. Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. Carbohydr Res 2003;338(23):2503–19.
29. Raetz CRH, Whitfield C. Lipopolysaccharide endotoxins. Annu Rev Biochem 2002;71(1):635–700.
30. Liu R, Hong J, Xu X, Feng Q, Zhang D, Gu Y, Shi J, Zhao S, Liu W, Wang X, et al. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. Nat Med 2017;23(7):859–68.
31. Feng Q, Liang S, Jia H, Stadlmayr A, Tang L, Lan Z, Zhang D, Xia H, Xu X, Jie Z, et al. Gut microbiome development along the colorectal adenoma-carcinoma sequence. Nat Commun 2015;6(1):6528.
32. Zhang X, Zhang D, Jia H, Feng Q, Wang D, Liang D, Wu X, Li J, Tang L, Li Y, et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. Nat Med 2015;21(8): 895–905.
33. Wu H, Esteve E, Tremaroli V, Khan MT, Caesar R, Manneras-Holm L, Stahlman M, Olsson LM, Serino M, Planas-Felix M, et al. Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. Nat Med 2017;23(7): 850–8.
34. Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, Nielsen J, Backhed F. Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature 2013;498(7452):99–103.
35. Qin N, Yang F, Li A, Prifti E, Chen Y, Shao L, Guo J, Le Chatelier E, Yao J, Wu L, et al. Alterations of the human gut microbiome in liver cirrhosis. Nature 2014;513(7516):59–64.
36. Hill MJ. Intestinal flora and endogenous vitamin synthesis. Eur J Cancer Prev 1997;6(Suppl 1):S43–5.
37. Rossi M, Amaretti A, Raimondi S. Folate production by probiotic bacteria. Nutrients 2011;3(1):118–34.
38. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature 2014;505(7484):559–63.
39. Meslier V, Laiola M, Roager HM, De Filippis F, Roume H, Quinquis B, Giacco R, Mennella I, Ferracane R, Pons N, et al. Mediterranean diet intervention in overweight and obese subjects lowers plasma cholesterol and causes changes in the gut microbiome and metabolome independently of energy intake. Gut 2020;69(7):1258–68.
40. Rampelli S, Soverini M, D'Amico F, Barone M, Tavella T, Monti D, Capri M, Astolfi A, Brigidi P, Biagi E, et al. Shotgun metagenomics of gut microbiota in humans with up to extreme longevity and the increasing role of xenobiotic degradation. mSystems 2020;5(2):e00124–20.
41. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, Pettersson S. Host-gut microbiota metabolic interactions. Science 2012;336(6086): 1262–7.
42. Schwiertz A, Taras D, Schafer K, Beijer S, Bos NA, Donus C, Hardt PD. Microbiota and SCFA in lean and overweight healthy subjects. Obesity 2010;18(1):190–5.
43. Chambers ES, Viardot A, Psichas A, Morrison DJ, Murphy KG, Zac-Varghese SE, MacDougall K, Preston T, Tedford C, Finlayson GS, et al. Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. Gut 2015;64(11):1744–54.
44. Vulesevic B, McNeill B, Giacco F, Maeda K, Blackburn NJ, Brownlee M, Milne RW, Suuronen EJ. Methylglyoxal-induced endothelial cell loss and inflammation contribute to the development of diabetic cardiomyopathy. Diabetes 2016;65(6):1699–713.