UC Davis
UC Davis Previously Published Works

Title

Dietary Intake of Monosaccharides from Foods is Associated with Characteristics of the Gut Microbiota and Gastrointestinal Inflammation in Healthy US Adults

Permalink

https://escholarship.org/uc/item/2sx0q0zd

Journal

Journal of Nutrition, 153(1)

ISSN

0022-3166

Authors

Larke, Jules A
Bacalzo, Nikita
Castillo, Juan J
et al.

Publication Date

2023

DOI

10.1016/j.tjnut.2022.12.008

Peer reviewed

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

The Journal of Nutrition 153 (2023) 106–119

American Society for Nutrition
Excellence in Nutrition Research and Practice

JN THE JOURNAL OF NUTRITION

journal homepage: www.journals.elsevier.com/the-journal-of-nutrition

Nutrient Physiology, Metabolism, and Nutrient-Nutrient Interactions

Dietary Intake of Monosaccharides from Foods is Associated with Characteristics of the Gut Microbiota and Gastrointestinal Inflammation in Healthy US Adults

Jules A. Larke [1], Nikita Bacalzo [2], Juan J. Castillo [2], Garret Couture [2], Ye Chen [2], Zhengyao Xue [1,4], Zeynep Alkan [1], Mary E. Kable [1,3], Carlito B. Lebrilla [2,4], Charles B. Stephensen [1,3,†], Danielle G. Lemay [1,3,*]

[1] *USDA-ARS Western Human Nutrition Research Center, Davis, CA, USA;* [2] *Department of Chemistry, University of California, Davis, Davis, CA, USA;* [3] *Department of Nutrition, University of California, Davis, Davis, CA, USA;* [4] *Department of Food Science and Technology, University of California, Davis, Davis, CA, USA*

A B S T R A C T

Background: Current assessment of dietary carbohydrates does not adequately reflect the nutritional properties and effects on gut microbial structure and function. Deeper characterization of food carbohydrate composition can serve to strengthen the link between diet and gastrointestinal health outcomes.
Objectives: The present study aims to characterize the monosaccharide composition of diets in a healthy US adult cohort and use these features to assess the relationship between monosaccharide intake, diet quality, characteristics of the gut microbiota, and gastrointestinal inflammation.
Methods: This observational, cross-sectional study enrolled males and females across age (18-33 y, 34–49 y, and 50–65 y) and body mass index (normal, 18.5–24.99 kg/m^2; overweight, 25–29.99 kg/m^2; and obese, 30–44 kg/m^2) categories. Recent dietary intake was assessed by the automated self-administered 24-h dietary recall system, and gut microbiota were assessed with shotgun metagenome sequencing. Dietary recalls were mapped to the Davis Food Glycopedia to estimate monosaccharide intake. Participants with >75% of carbohydrate intake mappable to the glycopedia were included ($N = 180$).
Results: Diversity of monosaccharide intake was positively associated with the total Healthy Eating Index score (Pearson's $r = 0.520$, $P = 1.2 \times 10^{-13}$) and negatively associated with fecal neopterin (Pearson's $r = -0.247$, $P = 3.0 \times 10^{-3}$). Comparing high with low intake of specific monosaccharides revealed differentially abundant taxa (Wald test, $P < 0.05$), which was associated with the functional capacity to break down these monomers (Wilcoxon rank-sum test, $P < 0.05$).
Conclusions: Monosaccharide intake was associated with diet quality, gut microbial diversity, microbial metabolism, and gastrointestinal inflammation in healthy adults. As specific food sources were rich in particular monosaccharides, it may be possible in the future to tailor diets to fine-tune the gut microbiota and gastrointestinal function. This trial is registered at www.clinicaltrials.gov as NCT02367287.

Keywords: gut microbiota, gut inflammation, diet quality, carbohydrates, monosaccharide, glycan, dietary fiber, healthy adults

Introduction

Dietary carbohydrates have a considerable range in their degree of polymerization, conformational structure, and sugar composition, and each affects physiological handling by the host and their gut microbiota. However, assessment of carbohydrate intake is limited to a few broad categories to describe these parameters: dietary fiber, starches, and simple sugars. Recent advances in analytical chemistry have made it possible to analyze the monosaccharide composition of foods in high-throughput and

Abbreviations used: ASA24, automated self-administered 24-hr dietary recall system; FNDDS, Food and Nutrition Database for Dietary Studies; GalA, galacturonic acid; HEI, Healthy Eating Index; KEGG, Kyoto Encyclopedia of Genes and Genomes; KO, KEGG ortholog; LBP, lipopolysaccharide-binding protein; MPO, myeloperoxidase; RPKG, reads per kilobase per genome equivalent.

* Corresponding author. *E-mail address:* Danielle.Lemay@usda.gov (D.G. Lemay).

† CBS is a member of the Journal's Editorial Board.

https://doi.org/10.1016/j.tjnut.2022.12.008
Received 9 September 2022; Received in revised form 15 November 2022; Accepted 14 December 2022; Available online 26 December 2022
0022-3166/Published by Elsevier Inc. on behalf of American Society for Nutrition.

expand the assessment of carbohydrate intake [1]. Typical carbohydrate intake provides a significant source of calories from starches and sugars and the main carbon source for our gut microbes from fermentable dietary fiber. As such, dietary carbohydrates can be operationally defined as either available or resistant carbohydrates based on their glycosidic linkages, which determines the capacity for digestion and absorption in the host [2]. Additionally, the molecular properties of the sugar monomers that comprise these structures impact physiological handling based on the presence of pathways for carbohydrate metabolism [3, 4].

The inter- and intravariability in structures of dietary fiber across carbohydrate-rich foods lends itself to the complexity in understanding the relationship between diet and gut microbial ecology [5]. Fiber supplementation studies have aimed at isolating the impact of specific fiber types, e.g., inulin [6], pectin [7], and resistant starch [8], on health outcomes mediated through the gut microbiota. Given the interest in this area of research, a comprehensive database of intervention studies on fiber and health outcomes has been developed to summarize the literature on this topic [9]. Additionally, the study of whole plant foods provides another perspective on diet-based modulation of gut microbiota [10]. However, diets are not single foods nor specific fibers; they consist of a range of nondigestible carbohydrates that act in concert to affect the gut microbiota structure and function. Thus, identifying the glycomic signature of food is imperative in predicting diet-induced shifts in the gut microbiota and related health outcomes. We recently published a novel food glycan database (Davis Food Glycopedia) which describes the abundances of 10 monosaccharides in over 800 foods [1]. Leveraging fine-scale food carbohydrate data is needed in the effort to accurately capture the structural diversity and molecular detail of our diet and describe the relationship between diet and health outcomes for the advancement of precision nutrition.

Here, we characterize the comprehensive monosaccharide composition of recent dietary intake from healthy US adults. These data represent the monomeric composition (10 monosaccharides) of carbohydrate intakes from foods in the Davis Food Glycopedia. Our goal is to determine the levels of monosaccharides consumed by adults and the relationship between their consumption, diet quality, i.e., the 2015 Healthy Eating Index (HEI) [11], the gut microbiota, and gastrointestinal health. This is the first study to investigate population-based dietary intake at this resolution of food glycan composition and connect these dietary features to gut microbiota composition and markers of gastrointestinal health.

Subjects and Methods

Study population

During the period between May 2015 and July 2019, a total of 393 subjects were enrolled in the cross-sectional, observational USDA Nutritional Phenotyping Study (clinicaltrials.gov identifier: NCT02367287) conducted in Davis, California [12]. Briefly, subjects were recruited in an 18-bin sampling scheme balanced for age, sex, and BMI. Male and female subjects were individually recruited into 3 bins for age (18–33, 34–49, and 50–65 y) and 3 BMI categories (normal, 18.5–24.99 kg/m^2; overweight, 25–29.99 kg/m^2; and obese, 30–44 kg/m^2). From the initial enrollment, 358 subjects successfully completed the study and 343 had 2 or more complete Automated Self-Administered 24-h dietary recall (ASA24) questionnaires. Participants with >75% of total carbohydrate intake mappable to the glycopedia were included in the analysis, resulting in a final sample size of 180. A flow chart of participant inclusion is shown in Figure 1A.

Primary and secondary trial outcomes

The primary outcome for this study was gut microbial α diversity assessed by observed taxa (number of detectable gut taxa) and the Shannon diversity index. Secondary outcomes included HEI score and markers of gastrointestinal health: fecal calprotectin, myeloperoxidase (MPO), neopterin, and plasma lipopolysaccharide-binding protein (LBP). Exploratory analyses included differential abundance of gut microbial taxa across individual monosaccharide intake and microbial genes related to metabolism of these monosaccharides.

Food glycopedia

Currently, the Davis Food Glycopedia consists of 913 foods across 9 food groups with absolute quantities of 10 monosaccharides: D-glucose, D-galactose, D-fructose, L-arabinose, D-xylose, D-mannose, L-rhamnose, L-fucose, D-ribose, and D-galacturonic acid (GalA). Selection of foods for the Davis Food Glycopedia was based in part on foods most commonly consumed by the USDA Nutritional Phenotyping Study cohort used in the present study. Further details on food selection as well as methods of sample preparation and analysis are described in Castillo et al. [1].

Dietary assessment

Recent dietary intake was assessed using the ASA24, versions 2014 and 2016 [13]. Participants received e-mail prompts to complete dietary recalls on 2 weekdays and 1 weekend during the interim of the 2 study visits (between 10–14 d). Subjects completed 1 training recall with a study staff member, and subsequently, 3 were completed at home in response to unscheduled prompts. The 24-h recalls used in the current study were at-home recalls that passed quality control [14]. Nutrient densities were calculated per 1000 kcal. Diet quality was estimated using the HEI. HEI scores were based on individuals with 2 or 3 at-home dietary recalls.

“Ingredientization” of dietary intake data

Food items in the Davis Food Glycopedia consist mainly of single-ingredient foods rather than multi-ingredient foods (which could be of infinite varieties). Ingredientization of diet recall data and integration with the Davis Food Glycopedia was performed using custom python scripts. Diet recalls passing quality control were converted from Food and Nutrition Database for Dietary Studies (FNDDS) versions 4.1 and 2011–2012 [15] (corresponding to ASA24 versions 2014 and 2016, respectively) to version 2017–2018 to retrieve ingredient level information. FoodCodes that were missing or discontinued across versions of FNDDS were manually curated with the closest suitable description with 2 foods having no match. Ingredient codes that were represented as an 8-digit FoodCode were iteratively remapped to obtain the corresponding ingredient code(s). This process reduced the number of unique food descriptions from 2284 to 1199 single-ingredient foods (Figure 1B). Preprocessing of food recall data including conversion of discontinued FoodCodes, ingredientization of FoodCodes, and final mappings of ingredients to the Davis Food Glycopedia are included in the online Supplemental Materials (Supplemental File 1).

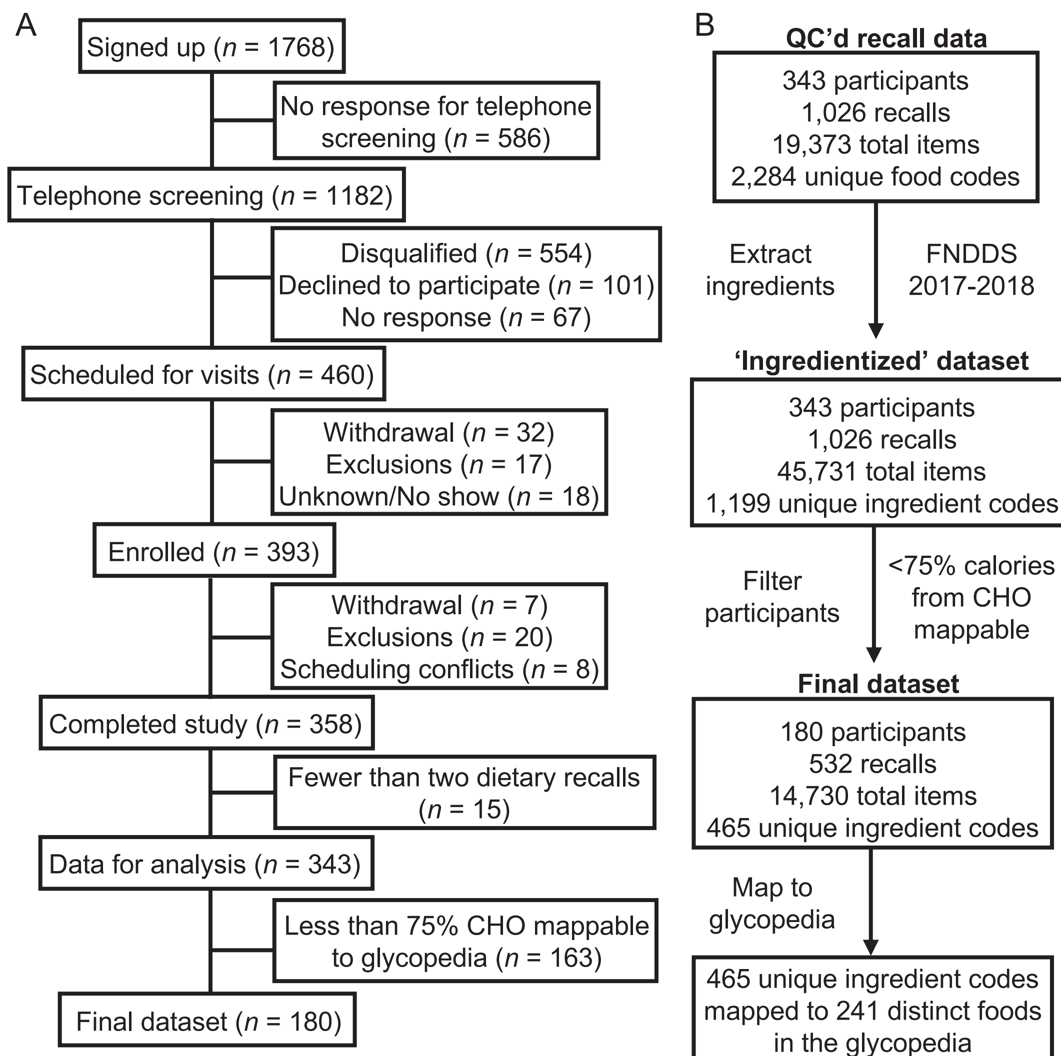

FIGURE 1. Flowcharts for (A) CONSORT diagram and (B) the process of converting dietary recalls to foods and ingredients for mapping to the glycopedia. Participants with less than 75% of calories from carbohydrates mappable to the glycopedia were excluded in the final dataset. Ingredient level information was extracted to increase coverage of diet recalls mappable to the glycopedia. Details for processing and mapping diet recall data can be found in Supplemental File 1. CHO, carbohydrate; CONSORT, Consolidated Standards of Reporting Trials; FNDDS, Food and Nutrient Database for Dietary Studies.

Assessment of monosaccharide intake

Diet recall data from the ASA24 and monosaccharide values from the Davis Food Glycopedia were merged to link monosaccharide content with foods and ingredients consumed by individuals in the cohort. The 'ingredientized' list of food recalls were merged with the Davis Food Glycopedia by manual matching to ensure accuracy for each individual food/ingredient. Following removal of participants with <75% of total carbohydrate intake mappable to the glycopedia, 465 unique foods/ingredients were mapped to 241 foods in the glycopedia. Figure 1B shows the process of converting diet recalls to foods and ingredients for mapping to the glycopedia. Supplemental File 1 lists food and ingredient descriptions for mapping across FNDDS versions and the Davis Food Glycopedia.

Monosaccharide intakes per ingredient were calculated by first taking the proportion of the ingredient in a multi-ingredient food (Equation 1). These proportions were multiplied by FoodAmt, the total amount consumed of the multi-ingredient food, to yield the gram quantities of consumed ingredients (Equation 2). Last, the quantity of ingredient consumed was multiplied by the concentration of monosaccharide for the matching food item in the Davis Food Glycopedia (Equation 3).

$$Proportion\ of\ ingredeient = \frac{Ingredient\ weight\ (g)}{sum(Ingredients\ in\ multi-ingredient\ food\ (g))} \quad (1)$$

$$Ingredient\ consumed\ (g) = FoodAmt\ (g) \times Proportion\ of\ ingredient \quad (2)$$

$$Glucose_{intake}\ (g) = Ingredient\ consumed\ (g) \times Glucose_{glycopedia} \left(\frac{g}{g}\right) \quad (3)$$

Stool collection and processing

One stool specimen was collected within 3 days of the second study visit at the end of the 10–14 day dietary recall period and

processed as described previously [12]. Briefly, stool samples were kept on blue ice and transported to the research center as soon as possible for same-day processing. A Stomacher paddle blender was used to homogenize samples prior to freezing at −80°C [12].

DNA extraction, library preparation, and sequencing

DNA was isolated using the ZymoBiomics DNA miniprep kit (Zymo Research) from 100 mg of homogenized stool as described in detail before [16]. Quality of the eluted DNA was assessed with Nanodrop (ThermoFisher), and the majority (>95%) of samples had $A_{260/280}$ and $A_{260/230}$ ratios above 1.80, with the lowest ratios of sequenced samples at 1.78 and 1.72, respectively. Representative DNA samples were confirmed to be intact and RNA-free prior to library construction by gel electrophoresis. DNA preps were quantified with the Qubit double-stranded DNA (dsDNA) broad-range assay (ThermoFisher), and all samples were diluted to 100 ng/μL.

Library preparation for shotgun genome sequencing, quality control, quantification, and pooling were performed by DNA Technologies and Expression Analysis Core Laboratory at the University of California at Davis, Genome and Biomedical Sciences Facility as previously described [16].

Metagenomic sequence analysis

BMTagger [17] was used to remove reads that aligned to the human genome version GRCh38.p13 [18]. Following this, reads were trimmed and adapter sequences removed with Trimmomatic version 0.33 [19] as described previously [20]. Reads were deduplicated using FastUniq version 1.1 with default settings [21]. Paired-end reads were then assembled using FLASH version 1.2.11 [22], setting overlapping length to 10 and 100 bp and mismatch ratio to 0.1. Taxonomy profiling was performed with Kraken2 [23] and aligned to a custom database [release 95 (13.07.2020)] using Struo [24]. Kraken2mpa.py from KrakenTools was used to format Kraken outputs for downstream analysis, and taxa with only 1 sequence read were dropped.

Paired-end deduplicated reads were mapped against the Kyoto Encyclopedia of Genes and Genomes (KEGG) database release 2019/10 to obtain KEGG ortholog (KO) counts [25]. Raw KO counts were normalized to reads per kb per genome equivalent [RPKG; RPKG = raw counts / (gene length × genome equivalent)], where genome equivalents based on the total paired-end, deduplicated reads for each sample were estimated using MicrobeCensus version 1.1.1 [26].

Markers of gastrointestinal inflammation and plasma LBP

Plasma LBP, fecal calprotectin, and fecal MPO were measured with enzyme-linked immunosorbent assays ELISAs as described previously [16]. The following kits were used: LBP (Human) ELISA (Abnova catalog number KA0448), Calprotectin ELISA (Immundiagnostik catalog number K6927), and MPO ELISA (Immundiagnostik catalog number KR6630). In order to measure fecal neopterin, approximately 100 mg of sample was taken from frozen homogenized stool specimens, weighed in tared tubes, and the exact sample weights were recorded. Fecal aliquots were extracted into 1 mL sterile 0.9% NaCl saline solution by vortexing at maximum speed for 30 min on a Vortex Genie 2 (Scientific Industries) equipped with a multiple tube holder adaptor. Extracts were then centrifuged at 1200 × *g* for 20 min to clarify, and the supernatants were kept frozen at -20°C for up to 4 days. Neopterin in undiluted fecal extracts was assayed with ELISAs (B·R·A·H·M·S/ThermoFisher catalog number 14-HD-99.1), and neopterin concentrations in the original stool samples were determined using the exact sample weights.

Statistical analyses

All statistical analyses and graphics were generated using R version 4.1.0. Partial Pearson's correlations were used to assess the relationship between monosaccharide intake variables and measures of microbial diversity and gastrointestinal markers of inflammation, adjusted for age, sex, and BMI. For continuous variables, normality was assessed by the Shapiro-Wilk test and observing deviations in the residuals of quantile-quantile plots. For nonnormally distributed data, appropriate transformations were used to approximate normal distributions. Mean centered values for the covariates age and BMI were used to estimate β coefficients for linear models of individual monosaccharide intake regressed on demographic variables. Alpha-diversity for monosaccharide intake and gut microbiota were calculated using the vegan package version 2.5-7 [27]. The vegan package was also used to assess β-diversity with PERMANOVA (permutational multivariate analysis of variance) by the adonis function with default parameters after checking for differences in group dispersion with betadisper. Differential abundance of microbial taxa was analyzed using the Wald test and likelihood ratio test with DeSeq2 version 1.34.0 [28]. Raw count data of genus level taxa were used as inputs filtered to include observations in the top and bottom quartiles of intake for a given monosaccharide. Wilcoxon rank-sum test was used for comparing genes from KOs across high compared with low intakes of monosaccharides. Samples for gut markers of inflammation were excluded if processed more than 24 h from sample collection or collected after the second study visit. For any given analysis, the maximum number of samples available were included. *P* values were adjusted for multiple comparisons using the Benjamini-Hochberg method, and statistical significance was set at $P < 0.05$.

Results

Participant characteristics

The demographic characteristics of the cohort included in this substudy ($n = 180$, Figure 1) are summarized in Table 1 and Supplemental Figure 1. The target population of the USDA Nutritional Phenotyping Study, from which these participants are sourced, were generally healthy individuals. Thus, subjects were excluded if they had recently undergone surgery, received antibiotic therapy, or had a recent hospitalization. Individuals taking medication for a diagnosed chronic illness were also excluded. A complete list of exclusion criteria can be found in Baldiviez et al. [12].

Validation of the Davis Food Glycopedia to assess dietary carbohydrates

Food composition databases traditionally calculate total carbohydrate in a food using the "by difference" method that takes

TABLE 1
Demographic characteristics of the participants

	Male	Female	All
Participants, *n* (%)	83 (46.1)	97 (53.9)	180 (100)
Age, y (mean ± SD)	40.7 ± 13.9	41.6 ± 13.4	41.4 ± 13.6
BMI, kg/m^2 (mean ± SD)	26.0 ± 4.2	27.2 ± 5.5	26.6 ± 5.0
Ethnicity, n (%)[1]			
Asian	12 (15.0)	10 (10.3)	22 (12.4)
Black	2 (2.5)	4 (4.1)	6 (3.4)
Hispanic	12 (15.0)	12 (12.4)	24 (13.6)
White	54 (67.5)	71 (73.2)	125 (70.6)

[1] $n = 177$; participants with ethnicity data and 75% of carbohydrate intake mappable to the Davis Food Glycopedia.

the difference between 100 and the sum of the percentages of the other measurable components. In comparison, the total monosaccharide abundances are absolute, analytically derived values that approximate the total carbohydrate intake. Given these differences in methodology, we compared total carbohydrate intake by taking the total intake averaged across recalls using "by difference" values reported in USDA FoodData Central [29] and the monosaccharide data. Total carbohydrate values from the 2 methods were highly correlated (Figure 2, Pearson's $r = 0.954$, $P < 6.9 \times 10^{-95}$), providing assurance for the novel method.

Monosaccharide composition of diets in a healthy US cohort

Glucose comprised the majority of dietary monosaccharides consumed by the cohort [83.4% ± 5.3%, mean ± standard deviation (SD)] followed by fructose (5.9% ± 2.9%), galactose (4.7% ± 2.7%), arabinose (2.1% ± 0.9%), xylose (1.3% ± 0.4%), GalA (1.2% ± 0.8%) and mannose (0.8% ± 0.6%) as shown in Figure 3A. Ribose, fucose and rhamnose were consumed at < 0.5% of mean intake. Aggregation of monosaccharide consumption by food group revealed glucose as the major (>50%) constituent apart from 2 groups: Eggs, and Fats, Oils, and Salad Dressings (Figure 3B). Overall, grains represented the largest contributors of monosaccharide intake.

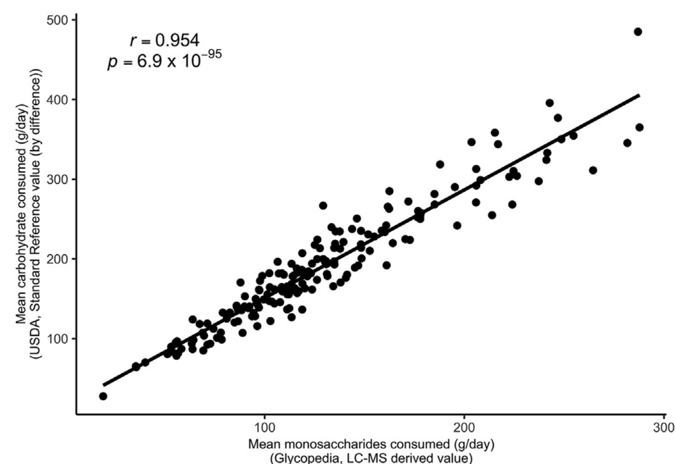

FIGURE 2. Pearson's correlation representing the relationship between carbohydrate and monosaccharide intake across the cohort ($n = 180$). Each point represents one participant's mean intake across diet recalls. Carbohydrate intakes are based on the Standard Reference value from FNDDS whereas monosaccharide intakes from the Davis Food Glycopedia are analytically derived by HPLC/MS. FNDDS, Food and Nutrient Database for Dietary Studies; HPLC/MS, high performance liquid chromatography/mass spectroscopy.

Excluding glucose from total monosaccharides (total nonglucose monosaccharides, Figure 3C) provided qualitative insights on the monosaccharide composition that reflect fiber content in the food categories. Both Vegetables and Fruits were abundant in GalA, reflecting the pectin content of these foods, whereas Grain Products contained the most arabinose and xylose, indicative of the arabinoxylan fibers in wheat. The category "Sugars, Sweets, and Beverages" contained the greatest amount of mannose from which coffee is a principal source in this cohort. Monosaccharide intake at the level of food category (mean ± SD) is provided in Supplemental Table 1.

The top food sources of individual monosaccharides were also examined (Figure 4). As expected, dairy products were the major source of galactose, whereas fruits and/or high-fructose beverages were the major source of fructose. Contributions to minor-abundance monosaccharides were more surprising. Coffee was the top contributor of mannose (199 ± 282 mg/d/1000 kcal); avocados were the largest source of GalA (201 ± 341 mg/d/1000 kcal); and the meat products, chicken, and beef, were the main contributors of ribose at 68 ± 81 mg/d/1000 kcal and 32 ± 41 mg/d/1000 kcal, respectively. Table sugar appeared as a top contributor to fucose intake; however, this result is likely to be an artifact of the data considering refined table sugar is virtually pure sucrose (glucose and fructose).

Dietary monosaccharide intake by participant characteristics

Monosaccharide intake was examined in relationship to age, sex, and BMI using multiple linear regression (Table 2). Overall, men consumed more glucose than women, and there was a slight increase in the consumption of several monosaccharides with age: fructose, arabinose, xylose, fucose, rhamnose, and GalA. With increasing BMI, there were small decreases in arabinose, xylose, and GalA intake, whereas intake of mannose and ribose increased with increasing BMI.

Dietary monosaccharide diversity and diet quality assessment

A diet high in simple sugars is likely to be a poor diet. Therefore, we hypothesized that a diverse monosaccharide diet would be positively correlated with diet quality. Given the distinct composition of monosaccharides across foods, we calculated the Shannon diversity index for total monosaccharide and nonglucose monosaccharide intake to assess the relationship between monosaccharide diversity and diet quality. Using the HEI score as a measure of diet quality, a statistically significant correlation was observed between total monosaccharide diversity and diet quality after adjusting for age, sex, and BMI (Figure 5A, Pearson's $r = 0.52$, $P < 1.2 \times 10^{-13}$). However, no relationship was observed between HEI score and diversity of nonglucose monosaccharide intake (Supplemental Figure 2).

Dietary monosaccharide diversity and microbial diversity

Given that dietary fiber provides a major source of fermentable substrates for the gut microbiota, identifying relationships between the monosaccharide constituents, dietary

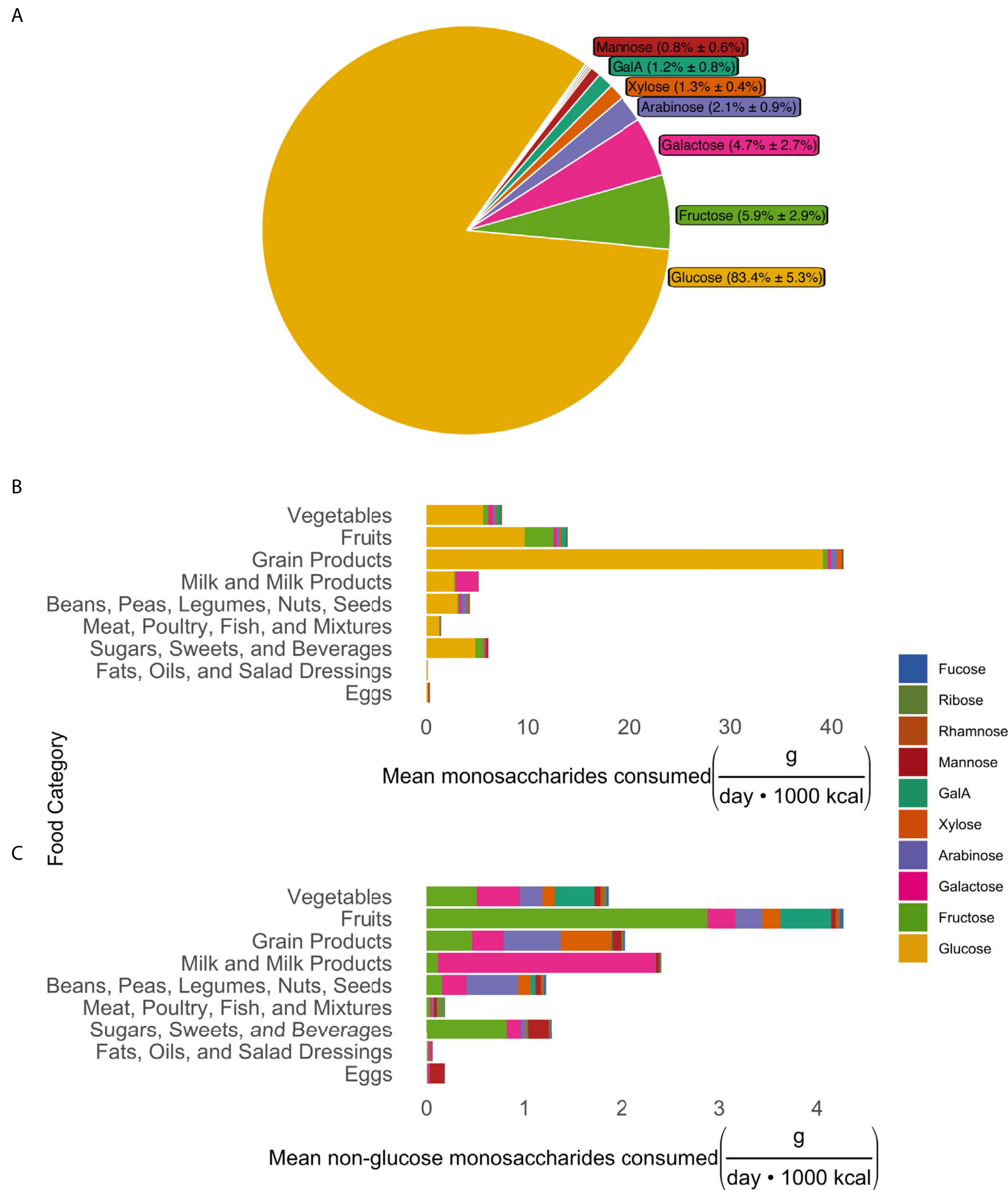

FIGURE 3. Mean monosaccharide intakes per participant ($n = 180$). (A) Pie chart representing the mean intake (mean ± SD) across diet recalls of the 7 most abundant monosaccharides consumed. Three additional monosaccharides with mean intake of less than 0.5% are not shown. (B) Stacked bar plots show the energy adjusted mean monosaccharide intake mappable to the glycopedia across diet recalls by food categories for all monosaccharides and (C) nonglucose monosaccharides.

fiber, and gut microbial diversity could provide additional insight into dietary influence on the gut microbiota in a free-living community. Total monosaccharide intake had a weaker association with dietary fiber intake derived from the ASA24 diet recalls (Pearson's $r = 0.314$, $P = 1.8 \times 10^{-5}$) compared with nonglucose monosaccharide intake (Pearson's $r = 0.702$, $P = 2.2 \times 10^{-16}$) (Supplemental Figure 3A, B). Moreover, a strong relationship between glucose monosaccharide intake and nonfiber carbohydrate intake (total carbohydrate – total dietary fiber) was observed (Pearson's $r = 0.875$, $P = 6.7 \times 10^{-58}$) (Supplemental Figure 3C).Given this observation, and the understanding that our monosaccharide data from certain food sources are not a component of fiber (e.g., glucose in the form of starch or table sugar), we estimated the α diversity from nonglucose monosaccharide intake and found that the Shannon diversity of nonglucose monosaccharides was predictive of observed gut microbial taxa (Figure 5B, Pearson's $r = 0.205$, $P = 0.012$). However, there was no statistically significant relationship between observed taxa and the diversity of total monosaccharide intake or between the Shannon diversity of the gut microbiome and diversity of total or nonglucose monosaccharide intake (Supplemental Figure 4A–C).

We next examined β diversity of gut microbiota across quartiles of monosaccharide intake. No differences were observed for either total monosaccharide intake (Pairwise betadisper, $P >$

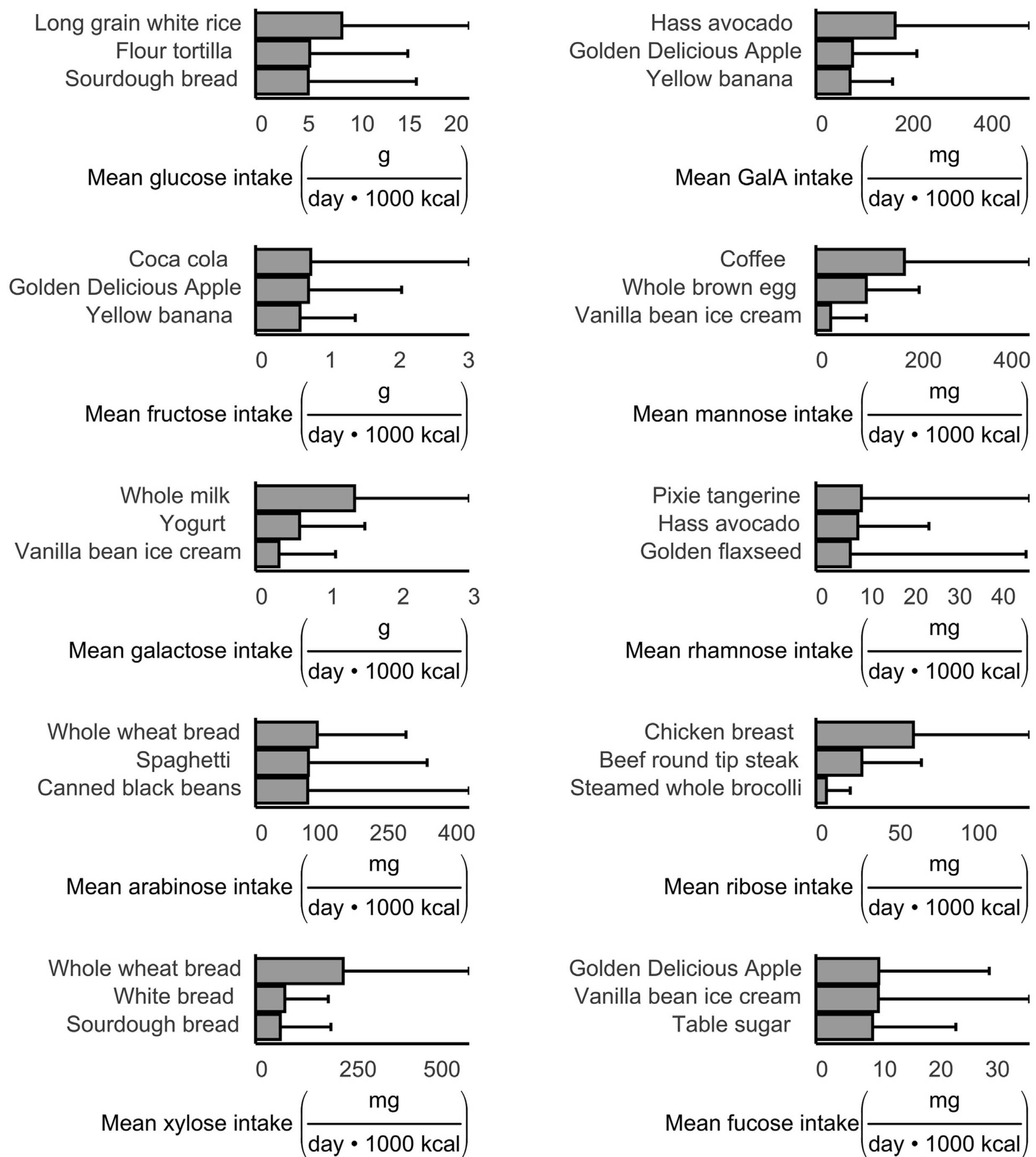

FIGURE 4. Top contributors to monosaccharide intakes by food and ingredient ($n = 180$). Bar plots represent the mean intake (mean ± SD) for the top 3 foods or ingredients contributing to a given monosaccharide intake. For each item, the mean is calculated by first summing unique foods or ingredients for each participant recall and dividing by the mean of that item for all participant recalls ($n = 532$).

0.83; PERMANOVA, $P = 0.766$) or total nonglucose monosaccharide intake (Pairwise betadisper, $P > 0.29$; PERMANOVA, $P = 0.304$) (Supplemental Figure 4D, E).

Specific monosaccharide intake is associated with gut microbial structure and function

As the data in the current version of the Davis Food Glycopedia do not yet provide information on glycosidic linkages, we selected the 3 most abundant monosaccharides: arabinose, xylose, and GalA, that are only present in the form of dietary fibers (i.e, not found in foods as mono-, di-, or trisaccharides and resistant to digestion) to investigate their effect on the gut microbiota.

Using DeSeq2 to identify taxa associated with consumption of specific monosaccharides, participants were split into high and low consumers based on the top and bottom quartiles of intake. Figure 6 shows the top 10 most abundant genera that are significantly differentially abundant (Wald test, $P < 0.05$) in high

TABLE 2
Coefficients for monosaccharide intake regressed on age, BMI, and sex[1]

Monosaccharide	Age[2]		BMI[2]		Sex (Male)		Transformation
	β (SE)	P	β (SE)	P	β (SE)	P	
Glucose	−0.038 (0.097)	0.701	−0.243 (0.269)	0.367	5.825 (2.659)	0.03	None
Fructose	0.004 (0.001)	0.004	−0.006 (0.004)	0.141	0.024 (0.037)	0.511	Log_{10}
Galactose	0.002 (0.001)	0.126	−0.003 (0.003)	0.414	−0.064 (0.034)	0.06	Log_{10}
Arabinose	0.002 (0.001)	0.045	−0.007 (0.002)	0.005	−0.018 (0.024)	0.46	Log_{10}
Xylose	0.002 (0.001)	0.044	−0.007 (0.002)	0.002	0.011 (0.023)	0.635	Log_{10}
Fucose	0.003 (0.001)	0.007	−0.003 (0.003)	0.312	−0.024 (0.025)	0.343	Log_{10}
Rhamnose	0.004 (0.001)	0.01	−0.007 (0.004)	0.064	−0.061 (0.037)	0.099	Log_{10}
GalA	0.003 (0.001)	0.032	−0.012 (0.004)	0.004	−0.057 (0.041)	0.161	Square root
Mannose	0.002 (0.001)	0.096	0.006 (0.003)	0.049	−0.015 (0.029)	0.597	Log_{10}
Ribose	−0.0006 (0.0004)	0.094	0.002 (0.001)	0.028	0.01 (0.01)	0.338	Square root

[1] $n = 180$.
[2] Mean centered continuous variable.

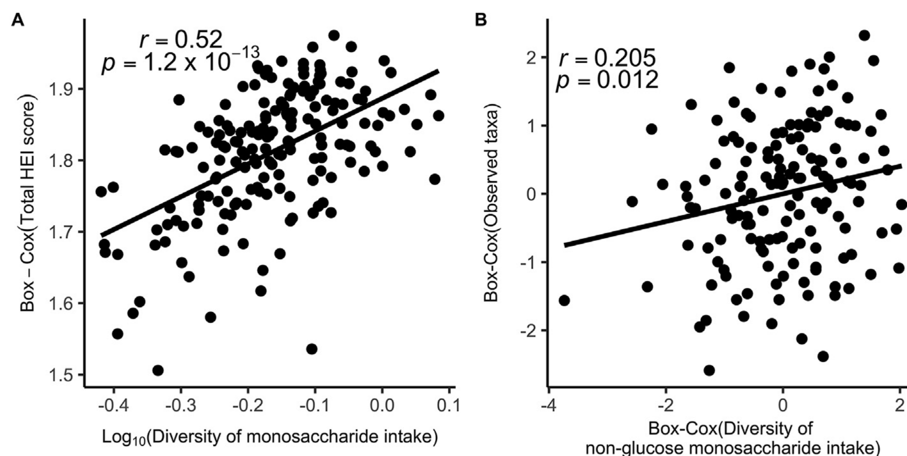

FIGURE 5. Associations between diversity of monosaccharide intake, diet quality, and gut microbial diversity. (A) Partial Pearson's correlation between diversity of monosaccharide intake calculated by the Shannon index and the total HEI score while controlling for the effects of age, sex, and BMI ($n = 180$). (B) Partial Pearson's correlation between diversity of nonglucose monosaccharide intake and the number of observed gut microbial taxa while controlling for the effects of age, sex, and BMI (n = 152). HEI, Healthy Eating Index.

compared to low consumers of arabinose, xylose, and GalA. The most abundant genera in the high monosaccharide consumers were Ruminiclostridium_E, CAG-180 [Acutalibacteraceae], and *Lachnospira* for arabinose, xylose, and GalA respectively. In contrast, the most abundant genera in the low monosaccharide intake groups were *Blautia* for arabinose and xylose, and *Faecalitalea* in participants consuming low amounts of GalA.

We then investigated the microbiota for the functional capacity to break down and metabolize monosaccharides. KOs were selected that corresponded to microbial genes involved in the cleavage and metabolism of arabinose, xylose, and GalA. The sum of all microbial genes across high compared with low consumers for each of the arabinose, xylose, and GalA degradation pathways exhibited no statistically significant differences. However, genes in the GalA degradation pathway mapping to *Lachnospira*, the most abundant taxon across high and low GalA consumers, were significantly increased in the participants with the highest quartile of GalA intake relative to the lowest quartile (Wilcoxon rank-sum test, $P < 0.05$) (Figure 7). Mapping genes to the most abundant taxon discriminating high to low consumption of arabinose, *Ruminiclostridium*_E, to the KEGG taxonomy revealed no differences in the arabinose degradation pathway across high compared with low intakes of arabinose (Supplemental Figure 5). However, participants within the highest quartile of xylose intake had a significantly greater abundance of *Ruminococcus*-derived xylose isomerase, a rate limiting enzyme in bacterial xylose catabolism (Wilcoxon rank-sum test, $P < 0.05$) (Supplemental Figure 5). Overall, while global changes in microbial genes in metabolic pathways for monosaccharide metabolism were not significantly affected by dietary monosaccharide intake, detection of species-specific gene enrichment, i.e., genes from *Lachnospira* involved in the metabolism of GalA, was observed.

Monosaccharide intake and markers of gut inflammation

Intake of nondigestible carbohydrates promotes gut health. We next examined the relationship between monosaccharide intake and diversity with markers of gut inflammation. Although no relationship was found with fecal calprotectin or MPO, fecal neopterin showed a positive trend with total monosaccharide intake (Figure 8A, $r = 0.147$, $P = 0.076$), which appears to be largely driven by glucose intake (Figure 8B). Conversely, fecal neopterin was negatively correlated with monosaccharide intake diversity (Figure 8C, $r = -0.247$, $P = 0.003$) but uncorrelated with nonglucose monosaccharide intake diversity (Supplemental Figure 6).

A similar finding was observed for plasma LBP, a plasma biomarker used as a proxy for gut barrier function as it is an

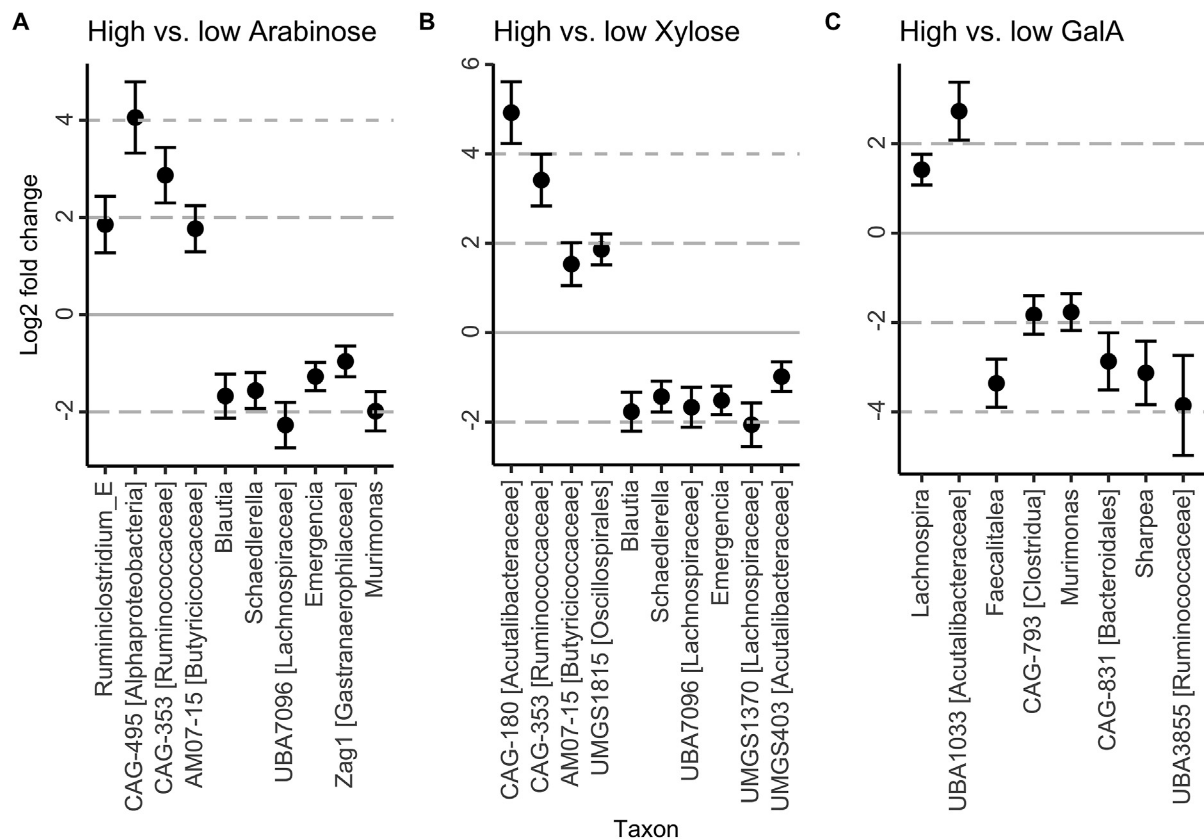

FIGURE 6. Differential abundance of microbial taxa across high and low consumers (n = 152) of specific monosaccharides arabinose, xylose, and GalA. The Wald test was used to detect pairwise differences in taxa between high and low consumers (top and bottom quartiles) for (A) arabinose, (B) xylose, and (C) GalA. Taxa are arranged from high to low abundance for each panel with positive log2 fold change being enriched in the top quartile and negative log2 fold change enriched in the bottom quartile of intake. Only the top 10 statistically significant, differentially abundant taxa for arabinose and xylose intake are shown. Unclassified taxa are annotated with the lowest annotated taxonomic rank shown in brackets. All taxa shown are statistically significant at $P < 0.05$ after adjusting for multiple comparisons. GalA, galacturonic acid.

indicator of the amount of lipopolysaccharide absorbed from the intestinal lumen. Total monosaccharide intake had a positive correlation with plasma LBP (Figure 8D, $r = 0.184$, $P = 0.014$), whereas the association with other monosaccharide intake metrics were not statistically significant (Supplemental Figure 6). These observations suggest that specific constituents of carbohydrate intake have modest yet variable associations with markers of gut health.

Discussion

This study provides a quantitative characterization of monosaccharide intakes from healthy US adults by "ingredientizing" diets and leveraging a novel food glycan database. Herein, we explored associations between dietary intake of monosaccharides and diet quality, characteristics of the gut microbiota, and markers of gut inflammation. Our method of ingredientizing the ASA24 recall data to map to a special interest database, the Davis Food Glycopedia, is novel and our data offer new perspectives on the dietary assessment of usual intake in free-living participants.

From a food composition perspective, the total monosaccharide and carbohydrate content are equivalent. However, the methodology used to assess these components differ. Given that this is the first instance of using the Davis Food Glycopedia database for nutrient assessment, we evaluated the relationship between total carbohydrate intake and total monosaccharide intake to validate its applicability in estimating carbohydrate consumption. Quantities of monosaccharides in foods were assessed by a recently developed liquid chromatography-mass spectroscopy platform [1] and were compared with the USDA Standard Reference values for carbohydrate that are based on a "by difference" calculation. The strong correlation suggests a high correspondence across various levels of intake; however, the estimate of the monosaccharide content of foods is consistently lower than that of carbohydrates. The carbohydrate estimate is based on the "by difference" method and is prone to overestimate this quantity as it will count the mass of components other than water, fat, protein, ash, and alcohol toward the total estimate. Additionally, the inclusion criteria; a minimum of 75% of calories from carbohydrates mappable to the glycopedia, would tend to underestimate the overall monosaccharide intake. Given these considerations, the data suggest that the Davis Food Glycopedia can be used for the assessment of carbohydrate intake in this cohort with the caveat that monosaccharide content will be systematically underestimated until all food ingredients are included in the database.

In the diets of these free-living adults, we observed patterns of food monosaccharide composition consistent with known food chemistry, which improves our confidence in the results: high GalA content from pectin fibers in fruits and vegetables [30], greater amounts of arabinose and xylose from arabinoxylans in grain products [31], and high galactose in dairy products. Several

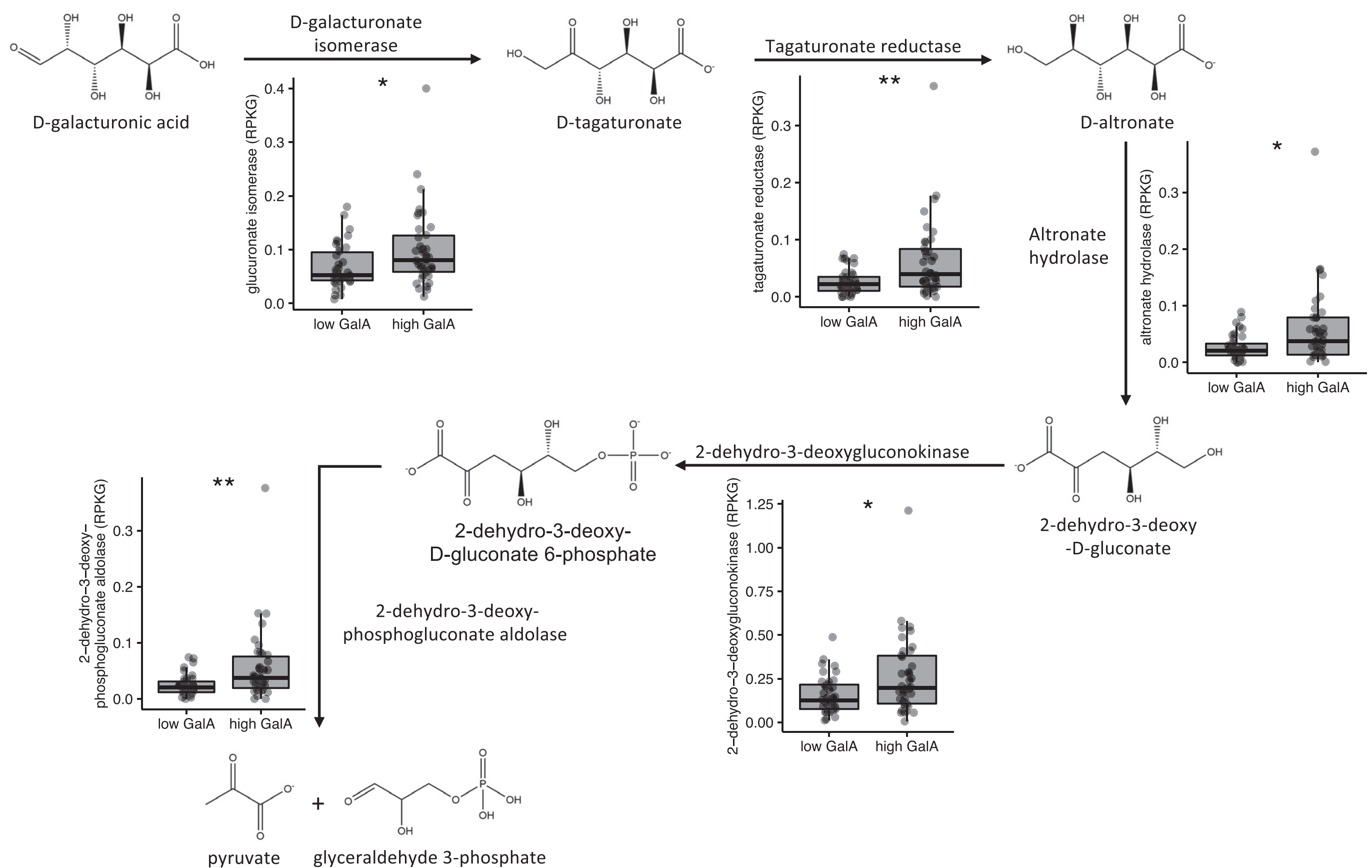

FIGURE 7. Metabolism of galacturonic acid (GalA) by *Lachnospira* across high and low consumers of GalA ($n = 152$). Metagenome data were mapped to Kyoto Encyclopedia of Genes and Genomes orthologs to obtain genes in the metabolic pathway, and counts were normalized to reads per kb per genome equivalent (RPKG). Wilcoxon rank-sum tests were performed to assess differences in genes from *Lachnospira* involved in the degradation of GalA. * $P < 0.05$, ** $P < 0.01$.

insightful patterns were observed from the contribution of individual foods to monosaccharide intakes. Coffee was the top contributor of mannose intake whereas avocado was the greatest source GalA. Thus, a few key food sources like coffee and avocado may have an outsized role in providing diverse carbon sources for the gut microbiota. Given this data, diets could be tailored to include foods that are rich sources of specific monosaccharides known to support the growth of particular gut microbes.

The Dietary Guidelines for Americans recommends incorporating a variety of foods within and across food groups to increase diet quality and nutritional adequacy [32]. Similarly, we hypothesized that a diet with diverse sources of monosaccharides, compared to a diet high in simple sugars (i.e., glucose and fructose) would track with diet quality. To address this, the Shannon diversity of monosaccharide intake was calculated, thereby providing a metric based on the relative proportion of monosaccharides consumed. Using the HEI score, a composite score that measures how well a diet aligns with the Dietary Guidelines for Americans, we found a positive correlation between monosaccharide intake diversity and HEI score, suggesting that a diet more varied in monosaccharide content corresponds to a healthier eating pattern. As dietary fiber is a significant determinant of gut microbiota composition, we predicted that a more diverse intake of monosaccharides would correspond to increased gut microbial diversity. However, we found that the Shannon diversity of the nonglucose monosaccharide intake, rather than total monosaccharide intake, was associated with α diversity. Glucose in the form of nonresistant starch, sucrose, and other simple sugars is largely absorbed in the small intestine, having minimal impact on the distal gut microbiota. As such, we observed a greater effect on gut community composition from nonglucose monosaccharides. However, the distinction between resistant and nonresistant starch deserves consideration. Our database does not distinguish the types of carbohydrate polymers in foods, including various types of starch. Collecting data on the food composition of resistant starch would benefit future analyses that examine the impact of glucose intake from these forms on gut microbial composition.

Previous observational studies have demonstrated the influence of diet on the gut microbiota; however, the small effect sizes may, in part, reflect the low resolution of traditional dietary predictors such as food groups and dietary fiber [33–35]. In our analyses, specific monosaccharides were used as predictors to discriminate microbial taxa across high and low intakes. Participants in the top quartiles of arabinose and xylose consumption relative to the bottom quartile had microbiota enriched in CAG-353, an unclassified member of the *Ruminococcaceae* family. This increase may reflect the higher consumption of whole grains, providing arabinoxylans as substrates to support the growth of this taxon. A previous study in which subjects

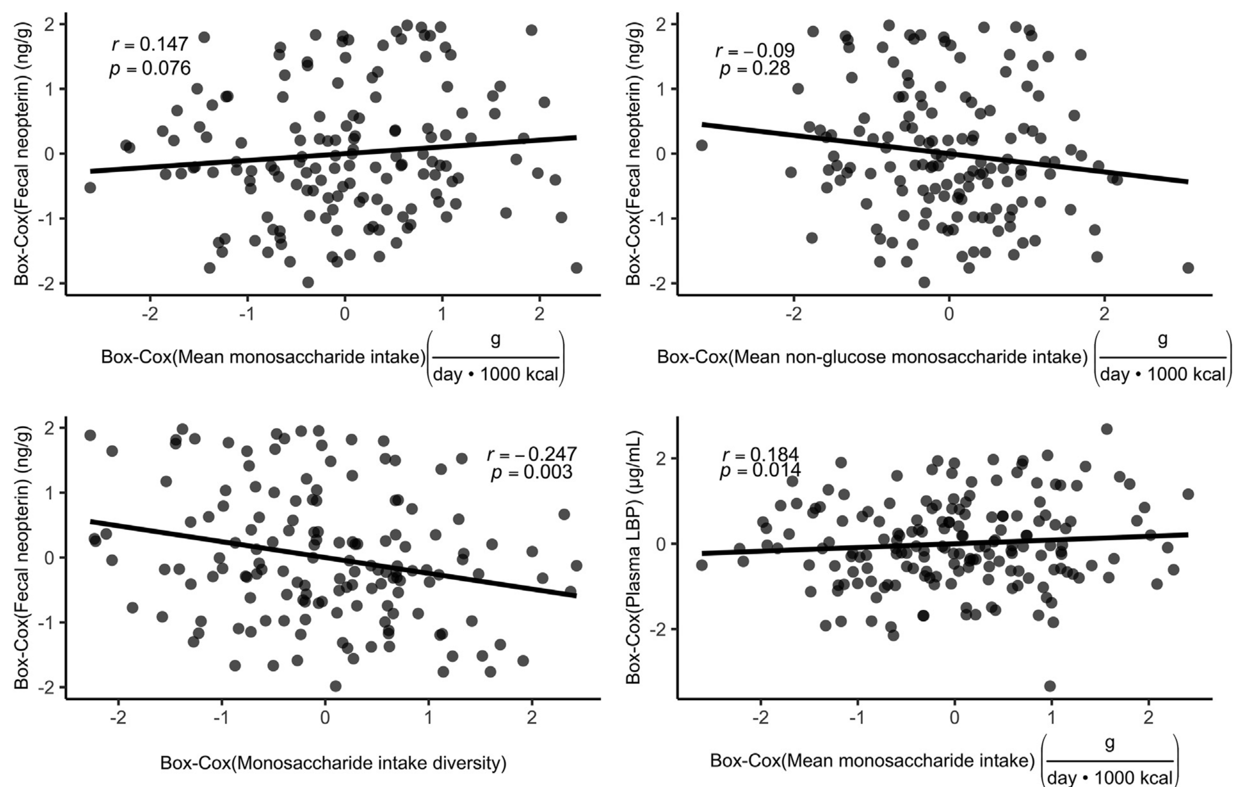

FIGURE 8. Associations between monosaccharide intake and diversity with markers of gastrointestinal health. (A–C) Partial Pearson's correlations between metrics of monosaccharide intake and fecal neopterin adjusted for the effects of age, sex, and BMI ($n = 151$). (D) Partial Pearson's correlation between mean monosaccharide intake and plasma lipopolysaccharide-binding protein (LBP) adjusted for the effects of age, sex, and BMI ($n = 180$).

consumed 100 g/d of wheat and barley for 2 months resulted in a statistically significant increase in gut ruminococci [36]. Similarly, providing 30 subjects classified as overweight or obese (BMI 25–40 kg/m^2) with 10.4 g/d of arabinoxylan oligosaccharides for 4 weeks resulted an increased abundance of *Ruminococcus sp.* [37]. Our study found that participants grouped as high arabinose consumers had an increased representation of *Ruminiclostridium* in their gut microbiota relative to the low arabinose consumers. In support of the *Ruminiclostridium*-arabinose interaction, a recent study of 18 subjects fed a fiber snack containing pea protein with high arabinose content (22.4% arabinose w/w) resulted in a greater amount of *Ruminiclostridium* after 6 weeks of intervention [38].

In our study, participants consuming the largest amount of GalA had significantly higher representation of *Lachnospira* in their gut microbiota compared to those consuming the lowest amount of GalA. A common source of GalA is pectin fibers, which predominantly consist of a GalA backbone with a degree of esterified methyl or acetyl groups [39]. The complexity of pectin fibers varies based on the plant source and can include branching chains of xylose, apiose, arabinose, and other sugars. Several *in vitro* human fecal fermentation studies have consistently observed a bloom in *Lachnospira* when pectin is provided as a substrate [5, 40–42]. Moreover, our previous study of the USDA Nutritional Phenotyping Study cohort found a positive association between *Lachnospira* abundance and fruit consumption [43]. Dietary interventions in human trials have observed similar effects. A trial of 24 subjects consuming a snack containing orange pulp fiber high in GalA (42.9% GalA w/w) resulted in statistically significant enrichment of several *Lachnospira* species [38]. Additionally, a randomized control trial that followed 163 men and women consuming isocaloric meals with and without avocado for 12 weeks detected a statistically significant increase in *Lachnospira* abundance [44]. Notably, avocado was the greatest contributor to dietary GalA in our study, further substantiating the link between the monosaccharide composition of dietary fiber and the gut microbiota. Although we found coffee to be a major source of mannose in this cohort, we did not find a significant association between coffee, specifically mannose intake, and the gut microbiota. Other studies have found associations between coffee consumption and the gut microbiome [33], which, in part, may be related to the polyphenol and alkaloid content in coffee [45].

Next, we examined the associations between gut microbial composition and microbial genes involved in the breakdown and metabolism of monosaccharides. Notably, participants in the highest quartile of GalA intake had a greater abundance of genes specific to *Lachnospira* in the GalA degradation pathway. These results suggest that the gut microbial function has the potential to be tuned by dietary intake. However, microbial genes in the saccharolytic pathways for arabinose and xylose (except xylose isomerase) were no different across high and low consumers. This may be, in part, due to the lack of correspondence between taxonomic groups for our metagenomic data and the KEGG organism database. Together, our data provides a high-resolution glycomic signature of the diet that can help link usual, free-living diets to gut microbiota composition and function.

Lastly, we examined the association between monosaccharide intake and markers of gut inflammation. Although we found no statistically significant relationships between fecal calprotectin or MPO with monosaccharide intake, the association between fecal neopterin and mean monosaccharide intake trended toward statistical significance. Mean nonglucose monosaccharide intake was not associated with fecal neopterin, indicating glucose intake as a significant component in this association. By far, grain products were the major source of glucose intake across the cohort. Previous reports have shown that consumption of whole grains, compared to refined grains, resulted in modest improvements to immune function that may be linked to changes in gut microbiota composition [46–48]. Fecal neopterin was negatively associated with the diversity of monosaccharides in the diet. This finding supports the idea that consumption of diverse sources of nondigestible carbohydrates are less inflammatory compared to a diet higher in simple sugars such as glucose and fructose.

In summary, our study delves into the nutritional "dark matter" of carbohydrates by examining their fine-scale molecular composition and characterizing the monosaccharide intakes of healthy US adults. We found associations between diversity of monosaccharide intake with diet quality, gut microbial diversity, and gut inflammation. Moreover, taxonomic differences in the gut microbiota across high and low GalA intake were linked to the functional capacity to metabolize these monosaccharides.

However, some limitations should be acknowledged. The data presented here represent the monosaccharides present in foods and not the glycan structures bound by glycosidic linkages. Hence, we are not able to draw inferences on the polysaccharide structures naturally found in foods. Additionally, although the Davis Food Glycopedia is the largest monosaccharide food composition database to date, the mapping of diets to the glycopedia represent, at minimum, 75% of the total carbohydrate intake, and this lack of complete dietary information is not fully representative in some individuals. Moreover, foods selected for inclusion in the Davis Food Glycopedia prioritized more commonly consumed items based on diet recall data from both the USDA Nutritional Phenotyping Study and What We Eat in America Database, which could explain the predominance of individuals that identified as White (~70%). Because the diets of individuals consuming uncommon ingredients could not be mapped to the Glycopedia foods, the impact of the most diverse monosaccharide profiles may not be fully investigated by this analysis. Additional foods that reflect cultural and ethnic choices need to be added to provide an equal representation of ethnicities and help generalize underrepresented populations.

In contrast, the quantitative monosaccharide intake data paired with markers of gut inflammation and assessment of the microbiota is a major strength of this study, providing a link between diet and gut microbiota composition and function. Dietary assessment in observational studies is constrained by the nutrients available in food composition databases. As such, public databases are limited to estimating total carbohydrate, fiber, simple sugars, and starch, which do not account for the specificity by which these sugar constituents are metabolized by the gut microbiota. Examining carbohydrate intake through the lens of monosaccharide composition can aid in the discovery of specific diet–health relationships that are less sensitive to traditional markers of carbohydrate intake. Future work will be needed to resolve the glycosidic linkages in foods to further uncover the complex diet–gut microbiome relationship.

Author disclosures

The authors report no conflicts of interest.

Funding

This research was supported by United States Department of Agriculture (USDA) Agricultural Research Service grant 2032-51530-026-00D and National Institutes of Health (NIH) award R01DK124193 and the USDA/NSF AI Institute for Next Generation Food Systems (AIFS), USDA, award number 2020-67021-32855. The USDA is an equal opportunity provider and employer.

Data Availability

Data described in the manuscript and analytic code will be made available upon request pending application and approval by the USDA ARS WHNRC Nutritional Phenotyping Study Investigators. Requests should be sent to the corresponding author. Scripts for processing raw sequence data can be found on GitHub at https://github.com/dglemay/ARG_metagenome. Metagenomic data described in the manuscript and code has been made publicly and freely available without restriction at NCBI Sequence Read Archive (SRA) under the study accession SRP354271 at https://www.ncbi.nlm.nih.gov/sra.

Acknowledgments

We acknowledge Ellen Bonnel for human studies management; Yasmine Bouzid and Eduardo Cervantes for human studies in the USDA Nutritional Phenotyping Study; Joanne Arsenault for quality control of dietary data; the Western Human Nutrition Research Center Bioanalytical Support Lab for sample management and processing of stool, and Sarah Spearman for assisting with fecal sample analysis. We also thank Andrew Oliver and Elizabeth Chin for helpful discussions related to data analysis. The author's responsibilities were as follows—D.G.L. conceived of the study; J.A.L., M.E.K., C.B.L., C.B.S., and D.G.L. designed research; J.A.L., Z.X., and Z.A. conducted research; N.B., J.J.C., G.C., Y.C., and C.B.L. provided essential monosaccharide data; J.A.L. analyzed data and performed statistical analysis; J.A.L. wrote the manuscript; N.B., Z.X., Z.A., M.E.K., C.B.L., C.B.S., and D.G.L. edited the manuscript. J.A.L. and D.G.L. had primary responsibility for final content. All authors have read and approved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tjnut.2022.12.008.

References

[1] J.J. Castillo, G. Couture, N.P. Bacalzo, Y. Chen, E.L. Chin, S.E. Blecksmith, et al., The development of the Davis Food Glycopedia-a glycan encyclopedia of food, Nutrients 14 (8) (2022) 1639.

[2] K.N. Englyst, S. Liu, H.N. Englyst, Nutritional characterization and measurement of dietary carbohydrates, Eur J Clin Nutr 61 (Suppl 1) (2007) S19–39.
[3] E.D. Sonnenburg, H. Zheng, P. Joglekar, S.K. Higginbottom, S.J. Firbank, D.N. Bolam, et al., Specificity of polysaccharide use in intestinal *Bacteroides* species determines diet-induced microbiota alterations, Cell 141 (7) (2010) 1241–1252.
[4] M.L. Patnode, Z.W. Beller, N.D. Han, J. Cheng, S.L. Peters, N. Terrapon, et al., Interspecies competition impacts targeted manipulation of human gut bacteria by fiber-derived glycans, Cell 179 (1) (2019) 59–73.e13.
[5] N. Larsen, C.B. De Souza, L. Krych, T.B. Cahú, M. Wiese, W. Kot, et al., Potential of pectins to beneficially modulate the gut microbiota depends on their structural properties, Front Microbiol 10 (2019) 1–13.
[6] D. Vandeputte, G. Falony, S. Vieira-Silva, J. Wang, M. Sailer, S. Theis, et al., Prebiotic inulin-type fructans induce specific changes in the human gut microbiota, Gut 66 (11) (2017) 1968–1974.
[7] C.L. Adam, S.W. Gratz, D.I. Peinado, L.M. Thomson, K.E. Garden, P.A. Williams, et al., Effects of dietary fibre (pectin) and/or increased protein (casein or pea) on satiety, body weight, adiposity and caecal fermentation in high fat diet-induced obese rats, PLOS One 11 (5) (2016), e0155871.
[8] M. Vital, A. Howe, N. Bergeron, R.M. Krauss, J.K. Jansson, J.M. Tiedje, Metagenomic insights into the degradation of resistant starch by human gut microbiota, Appl Environ Microbiol 84 (23) (2018) e01562–18.
[9] K.A. Livingston, M. Chung, C.M. Sawicki, B.J. Lyle, D.D. Wang, S.B. Roberts, et al., Development of a publicly available, comprehensive database of fiber and health outcomes: rationale and methods, PLOS One 11 (6) (2016), e0156961.
[10] H.J. Willis, J.L. Slavin, The influence of diet interventions using whole, plant food on the gut microbiome: a narrative review, J Acad Nutr Diet 120 (4) (2020) 608–623.
[11] US Department of Health and Human Services, 2015–2020 Dietary Guidelines for Americans, 8th edition, Department of Health and Human Services, Washington, DC, 2015.
[12] L.M. Baldiviez, N.L. Keim, K.D. Laugero, D.H. Hwang, L. Huang, L.R. Woodhouse, et al., Design and implementation of a cross-sectional nutritional phenotyping study in healthy US adults, BMC Nutr 3 (1) (2017) 79.
[13] Epidemiology and Genomics Research Program, Automated Self-Administered 24-Hour (ASA24) Dietary Assessment Tool, Available from: https://epi.grants.cancer.gov/asa24/, 2014.
[14] Y. Bouzid, J. Arsenault, E. Bonnel, E. Cervantes, A. Kan, D. Burnett, et al., Manual data cleaning can significantly change nutrient values recorded in Automated Self-Administered 24-Hour Recalls (ASA24), Curr Dev Nutr 4 (Suppl 2) (2020) 1162.
[15] U.S. Department of Agriculture - Agricultural Research Service. USDA Food and Nutrient Database for Dietary Studies 2011-2018. Food Surveys Research Group [Internet]. Available from: http://www.ars.usda.gov/nea/bhnrc/fsrg.
[16] A. Oliver, Z. Xue, Y.T. Villanueva, B. Durbin-Johnson, Z. Alkan, D.H. Taft, et al., Association of diet and antimicrobial resistance in healthy U.S. adults, mBio 13 (2022) e00101–e00122.
[17] K. Rotmistrovsky, Agarwala. BMTagger, Best Match Tagger for removing human reads from metagenomics datasets, Available from: https://ftp.ncbi.nlm.nih.gov/pub/agarwala/bmtagger/, 2011.
[18] V.A. Schneider, T. Graves-Lindsay, K. Howe, N. Bouk, H.C. Chen, P.A. Kitts, et al., Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly, Genome Res 27 (5) (2017) 849–864.
[19] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (15) (2014) 2114–2120.
[20] D.H. Taft, J. Liu, M.X. Maldonado-Gomez, S. Akre, M.N. Huda, S.M. Ahmad, et al., Bifidobacterial dominance of the gut in early life and acquisition of antimicrobial resistance, mSphere 3 (5) (2018) e00441-18.
[21] H. Xu, X. Luo, J. Qian, X. Pang, J. Song, G. Qian, et al., FastUniq: a fast de novo duplicates removal tool for paired short reads, PLOS One 7 (12) (2012), e52249.
[22] T. Magoč, S.L. Salzberg, FLASH: fast length adjustment of short reads to improve genome assemblies, Bioinformatics 27 (21) (2011) 2957–2963.
[23] D.E. Wood, J. Lu, B. Langmead, Improved metagenomic analysis with Kraken 2, Genome Biol 20 (1) (2019) 257.
[24 J. de La Cuesta-Zuluaga, R.E. Ley, N.D. Youngblut, Struo: a pipeline for building custom databases for common metagenome profilers, Bioinformatics 36 (7) (2020) 2314–2315.
[25] M. Kanehisa, S. Goto, KEGG: Kyoto Encyclopedia of Genes and Genomes, Nucleic Acids Res 28 (1) (2000) 27–30.
[26] S. Nayfach, K.S. Pollard, Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome, Genome Biol 16 (1) (2015) 51.
[27] J. Oksanen, G.F. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, et al., vegan: Community Ecology Package, Available from: https://cran.r-project.org/package=vegan, 2020.
[28] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol 15 (12) (2014) 550.
[29] U.S. Department of Agriculture, Agricultural Research Service, FoodData Central (2019). Available from: https://fdc.nal.usda.gov.
[30] L.A. Campbell, G.H. Palmer, Pectin, in: G.A. Spiller (Ed.), Topics in Dietary Fiber Research, Springer US, Boston, MA, 1978, Available from: https://doi.org/10.1007/978-1-4684-2481-2_4.
[31] L. Saulnier, N. Peneau, J.F. Thibault, Variability in grain extract viscosity and water-soluble arabinoxylan content in wheat, J Cereal Sci 22 (3) (1995) 259–264.
[32] J.A. Phillips, Dietary guidelines for Americans, Workplace Health Saf 69 (8) (2021) 395, 2020–2025.
[33] F. Asnicar, S.E. Berry, A.M. Valdes, L.H. Nguyen, G. Piccinno, D.A. Drew, et al., Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals, Nat Med 27 (2) (2021) 321–332.
[34] A. Zhernakova, A. Kurilshikov, M.J. Bonder, E.F. Tigchelaar, M. Schirmer, T. Vatanen, et al., Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity, Science 352 (6285) (2016) 565–569.
[35] G. Falony, M. Joossens, S. Vieira-Silva, J. Wang, Y. Darzi, K. Faust, et al., Population-level analysis of gut microbiome variation, Science 352 (6285) (2016) 560–564.
[36] M. De Angelis, E. Montemurno, L. Vannini, C. Cosola, N. Cavallo, G. Gozzi, et al., Effect of whole-grain barley on the human fecal microbiota and metabolome, Appl Environ Microbiol 81 (22) (2015) 7945–7956.
[37] A. Benítez-Páez, L. Kjølbæk, E. Gómez del Pulgar, L. Brahe, A. Astrup, S. Matysik, et al., A multi-omics approach to unraveling the microbiome-mediated effects of arabinoxylan oligosaccharides in overweight humans, mSystems 4 (4) (2019) 1–16.
[38] O. Delannoy-Bruno, C. Desai, J.J. Castillo, G. Couture, R.A. Barve, V. Lombard, et al., An approach for evaluating the effects of dietary fiber polysaccharides on the human gut microbiome and plasma proteome, Proc Natl Acad Sci U S A 119 (20) (2022), e2123411119.
[39] M. Beukema, M.M. Faas, P. de Vos, The effects of different dietary fiber pectin structures on the gastrointestinal immune barrier: impact via gut microbiota and direct effects on immune cells, Exp Mol Med 52 (9) (2020) 1364–1376.
[40] S.J. Bang, G. Kim, M.Y. Lim, E.J. Song, D.H. Jung, J.S. Kum, et al., The influence of in vitro pectin fermentation on the human fecal microbiome, AMB Express 8 (1) (2018) 98.
[41] Y.E. Tuncil, C.H. Nakatsu, A.E. Kazem, S. Arioglu-Tuncil, B. Reuhs, E.C. Martens, et al., Delayed utilization of some fast-fermenting soluble dietary fibers by human gut microbiota when presented in a mixture, J Funct Foods 32 (2017) 347–357.
[42] T.M. Cantu-Jungles, N. Bulut, E. Chambry, A. Ruthes, M. Iacomini, A. Keshavarzian, et al., Dietary fiber hierarchical specificity: the missing link for predictable and strong shifts in gut bacterial communities, mBio 12 (3) (2021) e01028-21.
[43] M.E. Kable, E.L. Chin, D. Storms, D.G. Lemay, C.B. Stephensen, Tree-based analysis of dietary diversity captures associations between fiber intake and gut microbiota composition in a healthy US adult cohort, J Nutr 152 (3) (2022) 779–788.
[44] S.V. Thompson, M.A. Bailey, A.M. Taylor, J.L. Kaczmarek, A.R. Mysonhimer, C.G. Edwards, et al., Avocado consumption alters gastrointestinal bacteria abundance and microbial metabolite concentrations among adults with overweight or obesity: a randomized controlled trial, J Nutr 151 (4) (2021) 753–762.
[45] S. González, N. Salazar, S. Ruiz-Saavedra, M. Gómez-Martín, C.G. de los Reyes-Gavilán, M. Gueimonde, Long-term coffee consumption is

associated with fecal microbial composition in humans, Nutrients 12 (5) (2020) 1–11.

[46] I. Martínez, J.M. Lattimer, K.L. Hubach, J.A. Case, J. Yang, C.G. Weber, et al., Gut microbiome composition is linked to whole grain-induced immunological improvements, ISME J 7 (2) (2013) 269–280.

[47] S.M. Vanegas, M. Meydani, J.B. Barnett, B. Goldin, A. Kane, H. Rasmussen, et al., Substituting whole grains for refined grains in a 6-wk randomized trial has a modest effect on gut microbiota and immune and inflammatory markers of healthy adults, Am J Clin Nutr 105 (3) (2017) 635–650.

[48] P. Vitaglione, I. Mennella, R. Ferracane, A.A. Rivellese, R. Giacco, D. Ercolini, et al., Whole-grain wheat consumption reduces inflammation in a randomized controlled trial on overweight and obese subjects with unhealthy dietary and lifestyle behaviors: role of polyphenols bound to cereal dietary fiber, Am J Clin Nutr 101 (2) (2015) 251–261.