Serum metabolomic signatures of plant-based diets and incident chronic kidney disease

Hyunju Kim,[1,2] Bing Yu,[3] Xin Li,[2] Kari E Wong,[4] Eric Boerwinkle,[3] Sara B Seidelmann,[5] Andrew S Levey,[6] Eugene P Rhee,[7] Josef Coresh,[1,2,8] and Casey M Rebholz[1,2,8]

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; [2]Welch Center for Prevention, Epidemiology, and Clinical Research, Johns Hopkins University, Baltimore, MD, USA; [3]Department of Epidemiology, Human Genetics & Environmental Sciences, University of Texas Health Sciences Center at Houston School of Public Health, Houston, TX, USA; [4]Metabolon, Research Triangle Park, NC, USA; [5]College of Physicians & Surgeons, Columbia University, New York, NY, USA; [6]Division of Nephrology, Tufts Medical Center, Boston, MA, USA; [7]Nephrology Division and Endocrine Unit, Massachusetts General Hospital, Boston, MA, USA; and [8]Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD, USA

ABSTRACT

Background: Greater adherence to plant-based diets is associated with a lower risk of incident chronic kidney disease (CKD). Metabolomics can help identify blood biomarkers of plant-based diets and enhance understanding of underlying mechanisms.

Objectives: Using untargeted metabolomics, we aimed to identify metabolites associated with 4 plant-based diet indices (PDIs) (overall PDI, provegetarian diet, healthful PDI, and unhealthful PDI) and incident CKD in 2 subgroups within the Atherosclerosis Risk in Communities study.

Methods: We calculated 4 PDIs based on participants' responses on an FFQ. We used multivariable linear regression to examine the association between 4 PDIs and 374 individual metabolites, adjusting for confounders. We used Cox proportional hazards regression to evaluate associations between PDI-related metabolites and incident CKD. Estimates were meta-analyzed across 2 subgroups ($n_1 = 1762$; $n_2 = 1960$). We calculated C-statistics to assess whether metabolites improved the prediction of those in the highest quintile compared to the lower 4 quintiles of PDIs, and whether PDI- and CKD-related metabolites predicted incident CKD beyond the CKD prediction model.

Results: We identified 82 significant PDI–metabolite associations (overall PDI = 27; provegetarian = 17; healthful PDI = 20; unhealthful PDI = 18); 11 metabolites overlapped across the overall PDI, provegetarian diet, and healthful PDI. The addition of metabolites improved prediction of those in the highest quintile as opposed to the lower 4 quintiles of PDIs compared with participant characteristics alone (range of differences in C-statistics = 0.026–0.104; P value $\leq$ 0.001 for all tests). Six PDI-related metabolites (glycerate, 1,5-anhydroglucitol, γ-glutamylalanine, γ-glutamylglutamate, γ-glutamylleucine, γ-glutamylvaline), involved in glycolysis, gluconeogenesis, pyruvate metabolism, and γ-glutamyl peptide metabolism, were significantly associated with incident CKD and improved prediction of incident CKD beyond the CKD prediction model (difference in C-statistics for 6 metabolites = 0.005; P value = 0.006).

Conclusions: In a community-based study of US adults, we identified metabolites that were related to plant-based diets and predicted incident CKD. These metabolites highlight pathways through which plant-based diets are associated with incident CKD. *Am J Clin Nutr* 2022;116:151–164.

Keywords: plant-based diets, metabolomics, chronic kidney disease, biomarkers, US adults

Introduction

Chronic kidney disease (CKD) is an important health problem. In the United States, nearly 15% of adults have CKD, including >780,000 adults with kidney failure receiving replacement therapy (KFRT) (1, 2). Dietary modification is an effective approach for preventing the onset of CKD and delaying progression to KFRT (3–5).

Evidence has suggested that plant-based diets, which are comprised predominantly of plant foods and are low in animal products, may be beneficial for kidney health. In the Atherosclerosis Risk in Communities (ARIC) study, those in the highest compared with the lowest quintile of vegetarian diets and healthful plant-based diets had 10%–14% lower risk of incident CKD, whereas those in the highest compared with the lowest quintiles of unhealthful plant-based diets had 11% higher risk of incident CKD (6). In addition, overall plant-based diets and healthful plant-based diets were associated with slower decline in kidney function (6). Despite the evidence, there is a lack of objective biomarkers of plant-based diets and our understanding of the mechanisms through which plant-based diets are associated with a lower risk of CKD is limited.

Metabolomic techniques measure hundreds of low-molecular-weight metabolites (usually <1500 daltons) in biofluids simultaneously (7–9). Dietary intake alters the metabolome (9).

Am J Clin Nutr 2022;116:151–164. Printed in USA.

Metabolites involved in the pathophysiology of CKD and the physiological changes that occur in response to food intake can be evaluated through metabolomics. This approach can help identify objective biomarkers of dietary patterns and, as such, improve upon existing methods for assessing dietary intake. In addition, metabolomic profiling is useful for elucidating metabolic pathways that are relevant for diet–kidney disease associations.

Using untargeted metabolomics, we aimed to extend the previous findings on plant-based diets and incident CKD by identifying metabolites associated with 4 plant-based diet indices (PDIs) (overall PDI, provegetarian diet, healthful PDI, and unhealthful PDI) and evaluating whether candidate biomarkers of plant-based diets are associated with incident CKD in 2 subgroups within a large prospective study.

Methods

Study design

The ARIC study, a community-based cohort of predominantly black and white middle-aged men and women, was designed to investigate the etiology of atherosclerosis and cardiovascular disease (10). At baseline (visit 1, 1987–1989), participants aged 45–64 y in 4 US communities (Washington County, MD; Forsyth County, NC; Minneapolis, MN; and Jackson, MS) were recruited into the study. Follow-up study visits occurred in 1990–1992 (visit 2), 1993–1995 (visit 3), 1996–1998 (visit 4), 2011–2013 (visit 5), and 2016–2017 (visit 6). Participants provided informed consent, and procedures were followed in accordance with the ethical standards of Institutional Review Boards at all study sites.

In 2010 and 2014, metabolomic profiling was conducted in 2 subgroups using stored, fasting (≥8 h, stored at −80°C) serum specimens collected at baseline (visit 1, 1987–1989). The first subgroup was a random sample of black participants from the Jackson, MS study site. The second subgroup was a random sample of black and white participants from all 4 study sites with genetic sequencing data. No participant overlapped between the 2 subgroups.

From 1880 and 2152 participants in subgroups 1 and 2, respectively, we excluded participants with implausible dietary intake (defined as men with total energy intake <700 or >4500 kcal/d and women with total energy intake <500 or >3500 kcal/d) ($n_1 = 16$; $n_2 = 61$), missing covariates ($n_1 = 75$; $n_2 = 84$), having prevalent CKD at baseline ($n_1 = 24$; $n_2 = 43$), missing CKD outcome at follow-up (none missing), or missing plant-based diet scores ($n_1 = 3$; $n_2 = 4$) (**Supplemental Figure 1**). After these exclusions, our analytic sample was 1762 and 1960 participants in subgroups 1 and 2, respectively.

Plant-based diet indices

At baseline (visit 1, 1987–1989), trained interviewers collected information on participants' usual consumption of foods and beverages using a modified 66-item semiquantitative Willett FFQ. Participants reported the frequency with which they had consumed foods and beverages of a specific serving size over the past year. Participants had 9 frequency options to choose from, ranging from "almost never" to ">6 per day." Visual guides were provided to help participants estimate serving sizes. Nutrient intakes were derived by multiplying frequency of consumption by the nutrient composition of each food item. Reproducibility of the FFQ was high in a random subsample of participants ($n = 419$) who were selected from all 4 study sites (11).

We used responses on the FFQ to calculate adherence to 4 different types of previously published PDIs (12, 13). These indices were developed to characterize diets that are higher in plant foods and lower in animal products and to distinguish healthful and unhealthful plant-based diets (12, 13). Detailed description of the calculation of the scores in the ARIC study is provided in a previous publication (6). Briefly, we classified foods and beverages in the FFQ into 17 food groups for the overall PDI, healthful PDI, and unhealthful PDI, and 11 food groups for the provegetarian diet index. For the overall PDI, healthful PDI, and unhealthful PDI, the 17 food groups were broadly classified into healthful plant foods (whole grains, fruits, vegetables, nuts, legumes, tea and coffee), unhealthful plant foods (refined grains, potatoes, fruit juices, sugar-sweetened and artificially sweetened beverages, sweets and desserts), and animal products (animal fat, dairy, eggs, fish or seafood, meat, miscellaneous animal foods). For the provegetarian diet index, the 11 food groups were classified into plant foods (grains, fruits, vegetables, nuts, legumes, potatoes) and animal products (animal fat, dairy, eggs, fish or seafood, meat).

Then, we used the residual method to calculate energy-adjusted consumption of each of the food groups. Using the residuals, we ranked participants by quintiles. For the overall PDI, consumption of all plant foods (regardless of the healthfulness of the plant foods) received higher scores (positively scored). For instance, those in the highest quintile of vegetable consumption received a score of 5, whereas those in

Supported by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grant R03 DK128386 (to CMR). The Atherosclerosis Risk in Communities (ARIC) study was supported by National Heart, Lung, and Blood Institute (NHLBI), NIH, Department of Health and Human Services contracts HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I, and HHSN268201700005I. Funding support for "Building on genome-wide association study (GWAS) for NHLBI-diseases: the U.S. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium" was provided by the NIH through the American Recovery and Reinvestment Act of 2009, grant 5RC2HL102419. Metabolomics measurements were supported by National Human Genome Research Institute grant 3U01HG004402-02S1. Laboratory testing at ARIC visit 6 was supported by NIDDK grant R01DK089174. CMR was in addition supported by NIDDK grant K01 DK107782 and NHLBI grants R56 HL153178 and R01 HL153178.

Supplemental Figures 1–3 and Supplemental Tables 1–4 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/ajcn/.

Address correspondence to CMR (e-mail: crebhol1@jhu.edu).

Abbreviations used: ARIC, Atherosclerosis Risk in Communities; CKD, chronic kidney disease; DASH, Dietary Approaches to Stop Hypertension; eGFR, estimated glomerular filtration rate; $eGFR_{Cr}$, estimated glomerular filtration rate based on creatinine; $eGFR_{Cys}$, estimated glomerular filtration rate based on cystatin C; GFR, glomerular filtration rate; HEI, Healthy Eating Index; KFRT, kidney failure receiving replacement therapy; PDI, plant-based diet index; UPLC, ultra-high-performance liquid chromatography; 1,5-AG, 1,5-anhydroglucitol.

Received November 11, 2021. Accepted for publication February 24, 2022.

First published online February 26, 2022; doi: https://doi.org/10.1093/ajcn/nqac054.

the lowest quintile received a score of 1. For the provegetarian diet index, similar to the overall PDI, consumption of plant foods received higher scores, but some plant foods were not scored (e.g., coffee and tea; sugar-sweetened and artificially sweetened beverages; sweets and desserts). For the healthful PDI, only healthful plant foods were positively scored. For the unhealthful PDI, only unhealthful plant foods were positively scored. For all diet indices, consumption of animal products received lower scores (negatively scored). For instance, those in the highest quintile of meat consumption received a score of 1, whereas those in the lowest quintile received a score of 5. Scores for all of the food groups in each index were summed.

These diet indices do not assess dietary patterns that exclude animal products, rather they rank individuals in a study population by relative intake of plant foods and animal products. As a result, those with higher scores for the overall PDI and provegetarian diet index had relatively higher intake of plant foods and relatively lower intake of animal products. Those with higher scores for the healthful PDI had relatively higher intake of healthful plant foods and relatively lower intake of unhealthful plant foods and animal products. Those with higher scores for the unhealthful PDI had relatively higher intake of unhealthful plant foods and relatively lower intake of healthful plant foods and animal products. The theoretical range for the overall PDI, healthful PDI, and unhealthful PDI was 17–85 and for the provegetarian diet index was 11–55.

Definition of CKD

CKD is defined as structural or functional abnormalities of the kidney for >3 mo, with implications for health (14, 15). The functional criterion for CKD is glomerular filtration rate (GFR) <60 mL $\cdot$ min^{-1} $\cdot$ 1.73 m^{-2}. In the ARIC study, a composite definition was used to define incident CKD: *1*) estimated GFR (eGFR) <60 mL $\cdot$ min^{-1} $\cdot$ 1.73 m^{-2} with eGFR decline $\geq$25% at any study visit compared with baseline (visit 1, 1987–1989), *2*) hospitalizations or death related to CKD using the relevant International Classification of Diseases (ICD-9/10) codes, or *3*) KFRT (also known as "end-stage renal disease"), defined as the initiation of renal replacement therapy (dialysis or transplant) and identified through linkage to the US Renal Data System (USRDS) (16).

In the primary analyses, serum or plasma creatinine was used to calculate estimated glomerular filtration rate ($eGFR_{Cr}$) using the 2009 CKD Epidemiology Collaboration (CKD-EPI) equation at every visit (17). Creatinine in serum was assessed at visits 1 and 2, and creatinine in plasma was assessed at visit 4 using the modified kinetic Jaffe method. Creatinine in serum was assessed at visits 3, 5, and 6 using the Roche enzymatic method. To account for variability between measurements, creatinine values were calibrated to the National Institute of Standards and Technology standard (18, 19).

As a sensitivity analysis, we used eGFR decline based on creatinine ($eGFR_{Cr}$) and cystatin C ($eGFR_{Cys}$) as the outcome. Serum cystatin C was used to calculate eGFR at every visit except for baseline (visit 1, 1987–1989) (20). $eGFR_{Cys}$ is less sensitive to confounding by diet and muscle mass than $eGFR_{Cr}$ (21, 22). At visits 2, 5, and 6, cystatin was assessed using the Gentian immunoassay. At visit 3, cystatin was assessed using the Roche Cobas 6000 chemistry analyzer. At visit 4, cystatin was assessed using a BNII nephelometer.

Metabolomic profiling

Untargeted metabolomic profiling was conducted by Metabolon using GC-MS and ultra-high-performance liquid chromatography–tandem MS (UPLC-MS/MS) on a Thermo Scientific Orbitrap MS analyzer. Details on metabolomic profiling have been reported previously (23, 24). Briefly, samples were extracted using an automated liquid handling robot (Hamilton Labstar, Hamilton Robotics). For GC-MS, the supernatants were injected onto a Thermo Scientific 5% diphenyl/95% dimethyl polysiloxane fused silica column (20 m $\times$ 0.18-mm ID; 0.18-μm film thickness) using helium as the carrier gas (temperature ramp from 60°C to 340°C) for a 17.5-min period. For UPLC, the supernatants were injected onto a 2.1 $\times$ 100-mm Waters BEH C18 1.7-μm column. Raw data were processed and peaks were identified using Metabolon's in-house software by matching to an extensive chemical library with >5000 commercially available standard compounds. To account for variation due to instrument interday tuning differences, each compound was adjusted in run-day blocks by normalizing data points proportionately to the median.

All known metabolites in the data set were confirmed using reference standards, and were classified as either tier 1 or tier 2 identification. In order to be tier 1 identification, metabolites had to meet $\geq$2 orthogonal measurements (e.g., accurate mass, retention index, fragmentation pattern) when compared to a reference standard (25, 26). Metabolites were considered a tier 2 identification if a reference standard was not available but there was evidence on physiochemical properties or spectral similarities (25, 26). Tier 2 metabolites are denoted in the tables with an asterisk.

Our analyses were restricted to named metabolites ($n = 385$) available in both subgroup 1 and subgroup 2. We excluded metabolites if >80% of values were missing across specimen ($n = 6$). For the rest of the metabolites, we imputed the lowest detectable value for each metabolite, scaled metabolites to a median of 1 by dividing by the subgroup-specific median, and used log transformation ($\log_e$) to improve normality. We excluded metabolites with very low variance (<0.01 on a log scale) or no variance ($n = 5$), then capped outliers at 5 SDs above or below the mean. These preprocessing steps yielded 374 metabolites in subgroups 1 and 2. In subgroup 2, a greater number of named metabolites were available than in subgroup 1 owing to improvement in the metabolomics platform over time (2010 compared with 2014). Twenty-seven metabolites that were unknown at the time of metabolomic profiling in subgroup 1 were retroactively named and were included in the present analyses.

Covariates

At baseline, participants' sociodemographic characteristics (age, sex, race/ethnicity, education), smoking, physical activity, alcohol intake, medication use, and diagnosed diseases were collected using a structured questionnaire. We adjusted for margarine consumption and total energy intake to be consistent

with prior studies of plant-based diets and chronic diseases (6, 27, 28), and because the composition of margarine may have been high in *trans* fats at the time diet was assessed in the ARIC study (12). Trained staff measured participants' height (cm) and weight (kg), which were used to calculate BMI (in kg/m^2). A certified technician measured systolic and diastolic blood pressure 3 times using a random-0 sphygmomanometer. The mean of the second and the third measurements was used. The modified hexokinase/glucose-6-phosphate dehydrogenase method was used to assess blood glucose concentration. Diabetes was defined as self-report of doctor's diagnosis of diabetes, diabetes medication use in the past 2 wk, fasting glucose ≥126 mg/dL, or nonfasting glucose ≥200 mg/dL. Hypertension was defined as self-report of antihypertensive medication use in the past 2 wk, systolic blood pressure ≥140 mm Hg, or diastolic blood pressure ≥90 mm Hg.

Statistical analyses

We examined baseline characteristics of the participants and nutritional characteristics (energy-adjusted macro- and micronutrients, fiber, cholesterol) of the diet according to quintiles of 4 different plant-based diet scores using means ± SDs for continuous variables and proportions for categorical variables.

For the cross-sectional analyses, we examined the associations between a 1-unit higher score in the 4 PDIs and 374 individual metabolites using multivariable linear regression models. In the cross-sectional analyses, we adjusted for age, sex, race-center (only in subgroup 2 because subgroup 1 consisted exclusively of African-American participants from the Jackson, MS site), education, physical activity, smoking, alcohol intake, margarine intake, BMI, eGFR, and total energy intake. A variable that combined race and study center was used in subgroup 2 to account for the nonuniform distribution of participants by race/ethnicity across the 4 ARIC study sites. Then, we meta-analyzed β coefficients of the 374 metabolites across the 2 subgroups using fixed-effects models (29), and adjusted the statistical significance level using the Bonferroni method to account for multiple comparisons ($\alpha = 3.34 \times 10^{-5}$; 0.05/4 dietary patterns/374 metabolites).

For the prospective analyses, we used multivariable Cox proportional hazards regression models to assess the associations between 1 SD higher in 51 diet-related metabolites and incident CKD from subgroups 1 and 2. We adjusted for the same covariates as the cross-sectional analyses, and in addition adjusted for diabetes, hypertension, and prevalent coronary artery disease at baseline. Similar to our approaches in the cross-sectional analyses, we meta-analyzed β coefficients of 51 diet-related metabolites across subgroup 1 and subgroup 2 using fixed-effects models (29) and used a Bonferroni-adjusted *P* value threshold of 9.80×10^{-4} (0.05/51 diet-related metabolites). For 6 diet-related metabolites that were significantly associated with incident CKD, we calculated Spearman rank correlation coefficients to describe the interrelation between these metabolites.

In subgroup 1 and subgroup 2 separately, and together, we used C-statistics to assess whether addition of plant-based diet–related metabolites improved the prediction of those in the highest quintile as opposed to the lower 4 quintiles of each of the PDIs beyond participant characteristics (the covariates used for the cross-sectional analyses). We used Harrell's C-statistics to examine whether plant-based diet– and CKD-related metabolites improved the prediction of incident CKD beyond the CKD risk prediction equation (30). The CKD risk prediction equation was developed using data from 34 multinational cohorts (30). The 5-y CKD risk prediction model for those without diabetes includes age, sex, race/ethnicity, baseline eGFR, smoking status, hypertension, BMI, history of cardiovascular disease, and albuminuria (30). We did not have albuminuria at baseline, thus we did not include this variable for our calculation of Harrell's C-statistics. Owing to lower sample sizes after excluding participants with diabetes ($n_1 = 280$, $n_2 = 221$), we pooled data from subgroup 1 and subgroup 2 ($n = 3221$) to increase statistical power.

We conducted 2 sensitivity analyses by *1*) stratifying the association between 1,5-anhydroglucitol (1,5-AG) and unhealthful PDI and incident CKD by diabetes status, given that this metabolite is more strongly associated with kidney disease among individuals with diabetes (31); and *2*) repeating the analyses on plant-based diet–related metabolites associated with CKD using eGFR decline as the outcome. We separately assessed the association between metabolites and $eGFR_{Cr}$ and $eGFR_{Cys}$. We used linear regression models with generalized estimating equations (exchangeable correlation structures and robust variance) and conducted meta-analysis across the 2 subgroups using fixed-effects models (29).

All analyses were conducted in Stata software version 15.0 (StataCorp) and R software version 4.1.0 (R Foundation for Statistical Computing).

Results

Baseline characteristics of the study population and nutritional characteristics of the dietary patterns

In subgroup 1, the overall PDI ranged from 34 to 72, provegetarian diet ranged from 19 to 49, the healthful PDI ranged from 34 to 74, and the unhealthful PDI ranged from 28 to 75 (**Table 1**). These ranges were similar in subgroup 2. Those in the highest quintiles of overall PDI, provegetarian diet, and healthful PDI were more likely to be slightly older, women, high school graduates, more physically active, have lower eGFR, and have coronary artery disease at baseline, and were less likely to be current smokers and to have diabetes than those in the lowest quintiles of these indices. Most of these trends (age, sex, education, physical activity, smoking status) were the opposite for the unhealthful PDI. Characteristics of the study population were similar for subgroup 2, except that those in the highest quintiles of all PDIs were less likely to be black.

Those in the highest quintile of overall PDI, provegetarian diet, and healthful PDI had higher intake of carbohydrate, plant protein, fiber, and micronutrients (potassium; magnesium; iron; β-carotene; vitamins C, B-6, and E; and folate) and lower intake of total fat than those in the lowest quintile (**Supplemental Table 1**). Those in the highest compared with the lowest quintile of unhealthful PDI had lower intake of fiber and micronutrients (sodium; phosphorus; calcium; potassium; magnesium; iron; β-carotene; vitamins B-1, B-2, B-6, B-12, C, and E; folate; zinc) and higher carbohydrate and added sugar intake. Those in the highest quintiles of all PDIs had lower intake of animal products than those in the lowest quintiles.

TABLE 1 Characteristics of study participants in subgroup 1 and subgroup 2 by quintiles of PDIs[1]

	Overall PDI		Provegetarian		Healthful PDI		Unhealthful PDI	
	Quintile 1	Quintile 5	Quintile 1	Quintile 5	Quintile 1	Quintile 5	Quintile 1	Quintile 5
Subgroup 1 ($n = 1762$)								
Sample size, n	378	291	381	269	413	303	402	314
Median score [range]	44 [34–46]	59 [57–72]	28 [19–29]	40 [38–49]	44 [34–46]	59 [57–74]	42 [28–45]	61 [58–75]
Age, y	52.4 ± 5.7	52.8 ± 5.7	52.2 ± 5.9	52.8 ± 5.8	52.2 ± 5.7	53.5 ± 5.7	53.3 ± 5.6	52.1 ± 5.7
Women	204 (54.0)	212 (72.9)	225 (59.1)	180 (66.9)	241 (58.4)	215 (71.0)	293 (72.9)	166 (52.9)
Black	378 (100)	291 (100)	381 (100)	269 (100)	413 (100)	303 (100)	402 (100)	314 (100)
High school graduate	201 (53.2)	190 (65.3)	223 (58.5)	172 (63.9)	235 (56.9)	188 (62.0)	248 (61.7)	167 (53.2)
Smoking status								
Never smokers	165 (43.7)	156 (53.6)	171 (44.9)	138 (51.3)	203 (49.2)	150 (49.5)	207 (51.5)	151 (48.1)
Former smokers	76 (20.1)	62 (21.3)	89 (23.4)	59 (21.9)	94 (22.8)	71 (23.4)	98 (24.4)	63 (20.1)
Current smokers	137 (36.2)	73 (25.1)	121 (31.8)	72 (26.8)	116 (28.1)	82 (27.1)	97 (24.1)	100 (31.8)
Physical activity index	2.1 ± 0.6	2.2 ± 0.7	2.1 ± 0.6	2.2 ± 0.7	2.1 ± 0.7	2.3 ± 0.7	2.2 ± 0.7	2.0 ± 0.6
Total energy intake, kcal	1730 ± 651	1652 ± 582	1662 ± 660	1746 ± 583	1528 ± 640	1810 ± 585	1505 ± 604	1899 ± 598
$eGFR_{Cr}$, $mL \cdot min^{-1} \cdot 1.73\ m^{-2}$	114.5 ± 16.9	113.2 ± 16.4	114.7 ± 17.2	112.3 ± 17.0	114.0 ± 17.2	112.0 ± 16.9	114.5 ± 16.9	114.0 ± 15.8
BMI, kg/m^2	29.5 ± 6.0	29.8 ± 6.3	30.0 ± 6.3	29.5 ± 6.6	29.8 ± 6.1	29.6 ± 5.9	30.4 ± 6.0	28.5 ± 5.5
Diabetes	64 (16.9)	43 (14.8)	67 (17.6)	37 (13.8)	74 (17.9)	46 (15.2)	84 (20.9)	33 (10.5)
Hypertension	199 (52.6)	145 (49.8)	199 (52.2)	138 (51.3)	206 (49.9)	150 (49.5)	199 (49.5)	157 (50.0)
Prevalent CAD	9 (2.4)	16 (5.5)	10 (2.6)	10 (3.7)	12 (2.9)	20 (6.6)	15 (3.7)	9 (2.9)
Subgroup 2 ($n = 1960$)								
Sample size, n	455	372	485	361	400	391	475	387
Median score [range]	44 [32–46]	59 [57–70]	27 [18–29]	40 [38–51]	43 [33–45]	59 [57–74]	43 [30–46]	59 [57–76]
Age, y	53.7 ± 5.7	54.5 ± 5.9	53.9 ± 5.7	54.8 ± 5.9	53.1 ± 5.6	54.8 ± 5.8	54.4 ± 5.6	53.7 ± 5.7
Women	209 (45.9)	225 (60.5)	249 (51.3)	201 (55.7)	197 (49.2)	226 (57.8)	308 (64.8)	186 (48.1)
Black	187 (41.1)	49 (13.2)	171 (35.3)	67 (18.6)	139 (34.8)	85 (21.7)	137 (28.8)	92 (23.8)
High school graduate	316 (69.5)	312 (83.9)	346 (71.3)	285 (78.9)	283 (70.8)	317 (81.1)	360 (75.8)	292 (75.5)
Smoking status								
Never smokers	162 (35.6)	161 (43.3)	178 (36.7)	152 (42.1)	165 (41.2)	161 (41.2)	210 (44.2)	153 (39.5)
Former smokers	140 (30.8)	130 (34.9)	140 (28.9)	136 (37.7)	118 (29.5)	137 (35.0)	151 (31.8)	123 (31.8)
Current smokers	153 (33.6)	81 (21.8)	167 (34.4)	73 (20.2)	117 (29.2)	93 (23.8)	114 (24.0)	111 (28.7)
Physical activity index	2.3 ± 0.8	2.5 ± 0.8	2.3 ± 0.7	2.6 ± 0.8	2.3 ± 0.8	2.6 ± 0.8	2.5 ± 0.8	2.3 ± 0.8
Total energy intake, kcal	1734 ± 642	1733 ± 547	1614 ± 637	1785 ± 536	1602 ± 641	1867 ± 560	1587 ± 572	1868 ± 585
$eGFR_{Cr}$, $mL \cdot min^{-1} \cdot 1.73\ m^{-2}$	104.8 ± 15.3	100.3 ± 14.8	103.7 ± 15.0	100.9 ± 13.6	103.5 ± 16.0	102.1 ± 13.2	103.6 ± 14.5	100.4 ± 15.9
BMI, kg/m^2	28.3 ± 5.5	27.1 ± 5.6	28.3 ± 5.4	27.3 ± 5.8	28.7 ± 5.5	27.2 ± 5.4	28.5 ± 5.8	27.8 ± 5.3
Diabetes	54 (11.9)	33 (8.9)	52 (10.7)	36 (10.0)	46 (11.5)	39 (10.0)	69 (14.6)	32 (8.3)
Hypertension	169 (37.1)	135 (36.3)	175 (36.1)	146 (40.4)	141 (35.2)	135 (34.5)	172 (36.2)	147 (38.0)
Prevalent CAD	23 (5.1)	33 (8.9)	22 (4.5)	38 (10.5)	17 (4.2)	34 (8.7)	27 (5.7)	30 (7.8)

[1]Values are means ± SDs for continuous variables and n (%) for categorical variables. CAD, coronary artery disease; $eGFR_{Cr}$, estimated glomerular filtration rate based on creatinine; PDI, plant-based diet index.

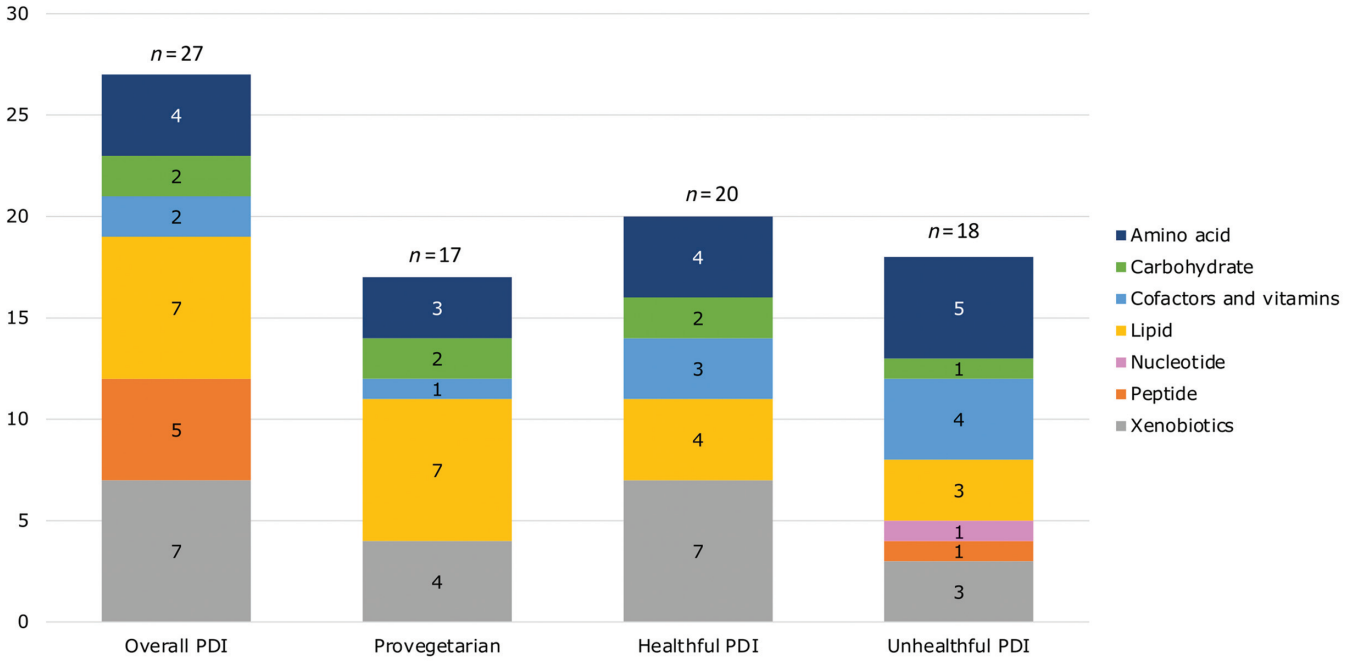

FIGURE 1 Metabolites significantly associated with PDIs in the Atherosclerosis Risk in Communities study. We used multivariable linear regression models to study the associations between 1 unit higher in PDIs and individual metabolites, adjusting for important confounders. Metabolites were meta-analyzed across 2 subgroups ($n_1 = 1762$; $n_2 = 1960$) using fixed-effects models. At the Bonferroni threshold of 3.34×10^{-5} (0.05/4 dietary patterns/374 metabolites), there were 82 significant associations (overall PDI = 27; provegetarian = 17; healthful PDI = 20; unhealthful PDI = 18). For reference, of the 374 individual metabolites in the data set, 23% ($n = 87$) were amino acids, 3% ($n = 13$) were carbohydrates, 3% ($n = 11$) were cofactors and vitamins, 1% ($n = 5$) were energy metabolites, 37% ($n = 138$) were lipids, 4% ($n = 15$) were nucleotides, 13% ($n = 49$) were peptides, and 15% ($n = 56$) were xenobiotics. PDI, plant-based diet index.

Metabolites associated with plant-based dietary patterns

We identified 82 significant associations between PDIs and metabolites (overall PDI = 27; provegetarian = 17; healthful PDI = 20; unhealthful PDI = 18) at the Bonferroni threshold of 3.34×10^{-5}. Of these, 51 metabolites were unique metabolites, indicating that some metabolites were associated with multiple dietary patterns (e.g., tryptophan betaine was associated with overall PDI, provegetarian diet, and healthful PDI). Across the 4 dietary patterns, xenobiotics (range of number of metabolites = 3–7), lipids (range = 3–7), and amino acids (range = 3–5) were the most common categories of metabolites that were significantly associated with PDIs (**Figure 1**). For overall PDI, peptides were another metabolite category with a large number of significant associations ($n = 5$). Eleven metabolites {4 xenobiotics (4-vinylphenol sulfate, O-methylcatechol sulfate, catechol sulfate, hippurate), 3 amino acid derivatives (N-acetylornithine, tryptophan betaine, indolepropionate), 2 lipids [10-nonadecenoate (19:1n–9) and 10-heptadecenoate (17:1n–7)], a carbohydrate (glycerate), and a cofactor and vitamin (threonate)} overlapped across overall PDI, provegetarian diet, and healthful PDI (**Supplemental Figure 2**). The overall PDI ($n = 12$) and unhealthful PDI ($n = 13$) had a similar number of metabolites that did not overlap with the other PDIs.

The majority of lipids were negatively associated with all 4 PDIs (**Table 2**; **Figure 2**). N-methylproline and stachydrine were positively associated with the overall PDI (**Figure 2A**). Xenobiotics (e.g., 4-vinylphenol sulfate, hippurate, catechol sulfate), amino acid derivatives (e.g., tryptophan betaine, N-acetylornithine), a carbohydrate (glycerate), and cofactors and vitamins were positively associated with the overall PDI, provegetarian diet, and healthful PDI (Figure 2A–C). Four γ-glutamyl peptides (γ-glutamylvaline, γ-glutamylleucine, γ-glutamylglutamate, γ-glutamylalanine) were positively associated with the overall PDI (Figure 2A). All 3 significant xenobiotics (catechol sulfate, homostachydrine, 3-hydroxypyridine sulfate) were negatively associated with the unhealthful PDI (Figure 2D), whereas 1,5-AG was positively associated with the unhealthful PDI (Table 2).

PDI-related metabolites associated with incident CKD

Over a median follow-up of 23 y, there were 606 and 807 incident CKD cases in subgroups 1 and 2, respectively. Six PDI-related metabolites (glycerate, 1,5-AG, γ-glutamylalanine, γ-glutamylglutamate, γ-glutamylleucine, γ-glutamylvaline) were significantly associated with incident CKD (**Table 3**; **Supplemental Table 2**). Four out of these 6 metabolites were involved in γ-glutamyl peptide metabolism. Per SD higher levels of all 6 metabolites were associated with 10%–13% lower risk of incident CKD. Metabolites in the γ-glutamyl peptide pathway were strongly associated with each other (range of $\rho = 0.72$–0.96) (**Supplemental Figure 3**).

TABLE 2 Metabolites significantly associated with PDIs[1]

Metabolite	Superpathway	Subpathway	HMDB ID	β	SE	P value
Overall PDI ($n = 27$)						
N-acetylornithine	Amino acid	Urea cycle; arginine and proline metabolism	—	0.0069	0.0014	3.01×10^{-7}
N-methylproline	Amino acid	Urea cycle; arginine and proline metabolism	—	0.0181	0.0031	3.60×10^{-9}
Tryptophan betaine	Amino acid	Tryptophan metabolism	HMDB61115	0.0240	0.0029	2.91×10^{-16}
Indolepropionate	Amino acid	Tryptophan metabolism	HMDB02302	0.0119	0.0023	2.69×10^{-7}
Erythronate*	Carbohydrate	Aminosugar metabolism	HMDB00613	0.0051	0.0008	7.11×10^{-11}
Glycerate	Carbohydrate	Glycolysis, gluconeogenesis, and pyruvate metabolism	HMDB00139	0.0127	0.0013	3.91×10^{-21}
γ-CEHC	Cofactors and vitamins	Tocopherol metabolism	HMDB01931	0.0105	0.0025	2.34×10^{-5}
Threonate	Cofactors and vitamins	Ascorbate and aldarate metabolism	HMDB00943	0.0153	0.0016	1.28×10^{-20}
1-Palmitoylplasmenylethanolamine*	Lipid	Lysolipid	—	−0.0065	0.0012	2.01×10^{-7}
Scyllo-inositol	Lipid	Inositol metabolism	HMDB06088	0.0114	0.0018	2.10×10^{-10}
Myo-inositol	Lipid	Inositol metabolism	HMDB00211	0.0027	0.0006	1.96×10^{-5}
CMPF	Lipid	Fatty acid, dicarboxylate	HMDB61112	−0.0181	0.0026	2.04×10^{-12}
10-Nonadecenoate (19:1n–9)	Lipid	Long-chain fatty acid	HMDB13622	−0.0063	0.0011	3.10×10^{-8}
10-Heptadecenoate (17:1n–7)	Lipid	Long-chain fatty acid	HMDB60038	−0.0062	0.0010	1.19×10^{-10}
Margaric acid (17:0)	Lipid	Long-chain fatty acid	HMDB02259	−0.0042	0.0009	2.46×10^{-6}
Pyroglutamylglycine	Peptide	Dipeptide	—	−0.0100	0.0023	1.08×10^{-5}
γ-Glutamylvaline	Peptide	γ-Glutamyl amino acid	HMDB11172	0.0063	0.0015	1.62×10^{-5}
γ-Glutamylleucine	Peptide	γ-Glutamyl amino acid	HMDB11171	0.0067	0.0015	1.02×10^{-5}
γ-Glutamylglutamate	Peptide	γ-Glutamyl amino acid	HMDB11737	0.0115	0.0023	8.51×10^{-7}
γ-Glutamylalanine	Peptide	γ-Glutamyl amino acid	HMDB29142	0.0097	0.0021	2.37×10^{-6}
Stachydrine	Xenobiotics	Food component/plant	HMDB04827	0.0312	0.0039	8.40×10^{-16}
Cinnamoylglycine	Xenobiotics	Food component/plant	HMDB11621	0.0117	0.0027	1.39×10^{-5}
Paraxanthine	Xenobiotics	Xanthine metabolism	HMDB01860	0.0150	0.0035	2.00×10^{-5}
4-Vinylphenol sulfate	Xenobiotics	Benzoate metabolism	HMDB04072	0.0176	0.0029	1.80×10^{-9}
O-methylcatechol sulfate	Xenobiotics	Benzoate metabolism	—	0.0099	0.0020	6.88×10^{-7}
Catechol sulfate	Xenobiotics	Benzoate metabolism	HMDB59724	0.0142	0.0018	1.49×10^{-15}
Hippurate	Xenobiotics	Benzoate metabolism	HMDB00714	0.0147	0.0024	9.45×10^{-10}
Provegetarian ($n = 17$)						
N-acetylornithine	Amino acid	Urea cycle; arginine and proline metabolism	—	0.0105	0.0016	1.04×10^{-10}
Tryptophan betaine	Amino acid	Tryptophan metabolism	HMDB61115	0.0313	0.0036	1.19×10^{-18}
Indolepropionate	Amino acid	Tryptophan metabolism	HMDB02302	0.0127	0.0028	6.61×10^{-6}
Erythronate*	Carbohydrate	Aminosugar metabolism	HMDB00613	0.0041	0.0009	1.49×10^{-5}
Glycerate	Carbohydrate	Glycolysis, gluconeogenesis, and pyruvate metabolism	HMDB00139	0.0093	0.0016	1.25×10^{-8}
Threonate	Cofactors and vitamins	Ascorbate and aldarate metabolism	HMDB00943	0.0111	0.0020	2.25×10^{-8}
Stearoyl sphingomyelin	Lipid	Sphingolipid metabolism	HMDB01348	−0.0038	0.0009	2.65×10^{-5}
2-Aminooctanoate	Lipid	Fatty acid, amino	HMDB00991	0.0088	0.0020	1.65×10^{-5}
CMPF	Lipid	Fatty acid, dicarboxylate	HMDB61112	−0.0198	0.0031	2.27×10^{-10}
Docosapentaenoic acid (DPA; 22:5n–6)	Lipid	PUFA (n–3 and n–6)	HMDB01976	−0.0079	0.0018	1.21×10^{-5}
Linoleic acid (18:2n–6)	Lipid	PUFA (n–3 and n–6)	HMDB00673	0.0034	0.0008	1.04×10^{-5}
10-Nonadecenoate (19:1n–9)	Lipid	Long-chain fatty acid	HMDB13622	−0.0061	0.0014	1.08×10^{-5}
10-Heptadecenoate (17:1n–7)	Lipid	Long-chain fatty acid	HMDB60038	−0.0058	0.0012	7.28×10^{-7}
4-Vinylphenol sulfate	Xenobiotics	Benzoate metabolism	HMDB04072	0.0183	0.0036	2.69×10^{-7}
O-Methylcatechol sulfate	Xenobiotics	Benzoate metabolism	—	0.0105	0.0024	1.45×10^{-5}
Catechol sulfate	Xenobiotics	Benzoate metabolism	HMDB59724	0.0138	0.0022	2.56×10^{-10}

(*Continued*)

TABLE 2 (*Continued*)

Metabolite	Superpathway	Subpathway	HMDB ID	β	SE	*P* value
Hippurate	Xenobiotics	Benzoate metabolism	HMDB00714	0.0131	0.0029	7.03×10^{-6}
Healthful PDI ($n = 20$)						
N-acetylornithine	Amino acid	Urea cycle; arginine and proline metabolism	—	0.0066	0.0013	3.54×10^{-7}
Tryptophan betaine	Amino acid	Tryptophan metabolism	HMDB61115	0.0232	0.0028	2.19×10^{-16}
Indolepropionate	Amino acid	Tryptophan metabolism	HMDB02302	0.0117	0.0022	1.33×10^{-7}
3-Phenylpropionate (hydrocinnamate)	Amino acid	Phenylalanine and tyrosine metabolism	HMDB00764	0.0112	0.0023	1.15×10^{-6}
Threitol	Carbohydrate	Pentose metabolism	HMDB04136	0.0059	0.0013	3.68×10^{-6}
Glycerate	Carbohydrate	Glycolysis, gluconeogenesis, and pyruvate metabolism	HMDB00139	0.0068	0.0013	2.06×10^{-7}
Pyridoxate	Cofactors and vitamins	Vitamin B-6 metabolism	HMDB00017	0.0067	0.0016	3.28×10^{-5}
Threonate	Cofactors and vitamins	Ascorbate and aldarate metabolism	HMDB00943	0.0076	0.0016	2.12×10^{-6}
N1-methyl-2-pyridone-5-carboxamide	Cofactors and vitamins	Nicotinate and nicotinamide metabolism	HMDB04193	0.0053	0.0013	3.33×10^{-5}
Stearoyl sphingomyelin	Lipid	Sphingolipid metabolism	HMDB01348	−0.0039	0.0007	1.57×10^{-7}
Myo-inositol	Lipid	Inositol metabolism	HMDB00211	0.0029	0.0006	3.66×10^{-6}
10-Nonadecenoate (19:1n–9)	Lipid	Long-chain fatty acid	HMDB13622	−0.0047	0.0011	1.84×10^{-5}
10-Heptadecenoate (17:1n–7)	Lipid	Long-chain fatty acid	HMDB60038	−0.0044	0.0009	1.62×10^{-6}
3-Hydroxypyridine sulfate	Xenobiotics	Chemical	—	0.0179	0.0029	8.07×10^{-10}
Quinate	Xenobiotics	Food component/plant	HMDB03072	0.0243	0.0035	4.82×10^{-12}
1-Methylurate	Xenobiotics	Xanthine metabolism	HMDB03099	0.0129	0.0029	9.33×10^{-6}
4-Vinylphenol sulfate	Xenobiotics	Benzoate metabolism	HMDB04072	0.0161	0.0028	1.13×10^{-8}
O-Methylcatechol sulfate	Xenobiotics	Benzoate metabolism	—	0.0149	0.0019	4.14×10^{-15}
Catechol sulfate	Xenobiotics	Benzoate metabolism	HMDB59724	0.0182	0.0017	1.48×10^{-26}
Hippurate	Xenobiotics	Benzoate metabolism	HMDB00714	0.0176	0.0023	2.11×10^{-14}
Unhealthful PDI ($n = 18$)						
Proline	Amino acid	Urea cycle; arginine and proline metabolism	HMDB00162	0.0029	0.0005	8.69×10^{-8}
S-methylcysteine	Amino acid	Methionine, cysteine, SAM, and taurine metabolism	HMDB02108	0.0156	0.0035	6.93×10^{-6}
3-Methyl-2-oxovalerate	Amino acid	Leucine, isoleucine, and valine metabolism	HMDB03736	0.0258	0.0044	3.95×10^{-9}
3-(4-Hydroxyphenyl)lactate	Amino acid	Phenylalanine and tyrosine metabolism	HMDB00755	0.0032	0.0007	1.23×10^{-5}
Pyroglutamine*	Amino acid	Glutamate metabolism	—	0.0062	0.0013	2.31×10^{-6}
1,5-Anhydroglucitol (1,5-AG)	Carbohydrate	Glycolysis, gluconeogenesis, and pyruvate metabolism	HMDB02712	0.0098	0.0019	2.46×10^{-7}
Pyridoxate	Cofactors and vitamins	Vitamin B-6 metabolism	HMDB00017	−0.0064	0.0015	9.03×10^{-6}
Bilirubin (Z,Z)	Cofactors and vitamins	Hemoglobin and porphyrin metabolism	HMDB00054	0.0161	0.0038	2.21×10^{-5}
γ-CEHC	Cofactors and vitamins	Tocopherol metabolism	HMDB01931	0.0097	0.0021	4.35×10^{-6}
Pantothenate	Cofactors and vitamins	Pantothenate and CoA metabolism	HMDB00210	−0.0057	0.0011	2.32×10^{-7}
CMPF	Lipid	Fatty acid, dicarboxylate	HMDB61112	−0.0131	0.0022	4.16×10^{-9}
DHA (22:6n–3)	Lipid	PUFA (n–3 and n–6)	HMDB02183	−0.0037	0.0007	5.33×10^{-7}
Nonadecanoate (19:0)	Lipid	Long-chain fatty acid	HMDB00772	−0.0027	0.0006	1.60×10^{-5}
N2,N2-dimethylguanosine	Nucleotide	Purine metabolism, guanine containing	HMDB04824	0.0041	0.0009	1.10×10^{-5}
Bradykinin	Peptide	Polypeptide	HMDB04246	0.0294	0.0042	3.08×10^{-12}
3-Hydroxypyridine sulfate	Xenobiotics	Chemical	—	−0.0116	0.0026	8.82×10^{-6}
Homostachydrine*	Xenobiotics	Food component/plant	HMDB33433	−0.0060	0.0013	7.02×10^{-6}
Catechol sulfate	Xenobiotics	Benzoate metabolism	HMDB59724	−0.0081	0.0015	1.58×10^{-7}

[1] β coefficients and *P* values were calculated from multivariable linear regression models which assessed the associations between 1 unit higher in PDIs and individual metabolites, adjusting for age, sex, race-center (only in subgroup 2), education, physical activity, smoking, alcohol consumption, margarine consumption, BMI, baseline estimated glomerular filtration rate, and total energy intake. Metabolites were meta-analyzed across 2 subgroups ($n_1 = 1762$; $n_2 = 1960$) using fixed-effects models. Meta-analyzed metabolites were considered statistically significant at the Bonferroni threshold of 3.34×10^{-5} (0.05/4 dietary patterns/374 metabolites). Positive coefficients indicate that serum level of the metabolite is higher with higher plant-based diet scores. Negative coefficients indicate that serum level of the metabolite is lower with higher plant-based diet scores. CEHC, carboxyethyl hydroxychroman; CMPF, 3-carboxy-4-methyl-5-propyl-2-furanpropanoate; HMDB ID, human metabolome database identification number; PDI, plant-based diet index; SAM, S-Adenosyl methionine.

*Metabolites have not been confirmed based on a reference standard.

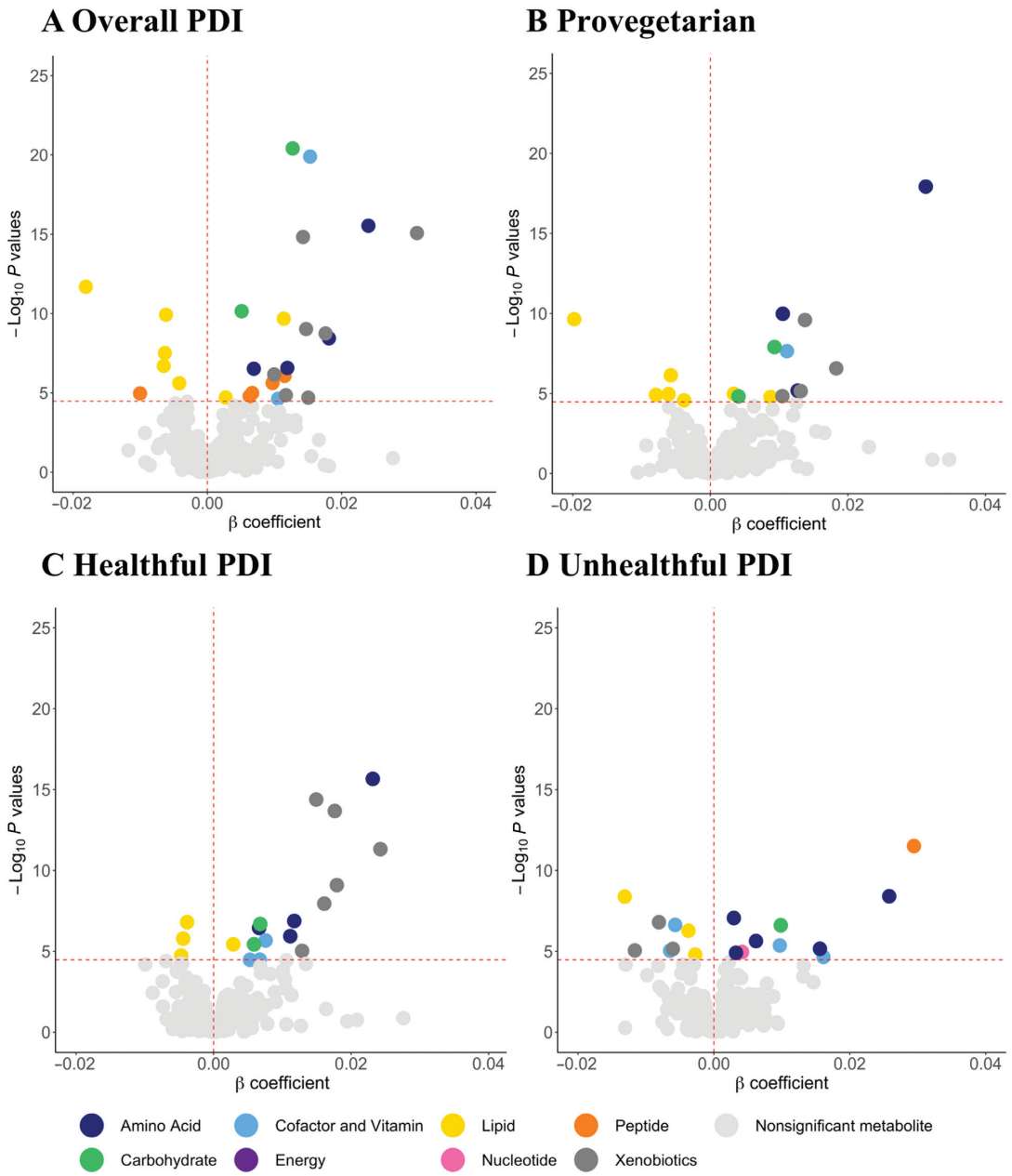

FIGURE 2 β coefficients and P values for the associations between 4 PDIs and individual metabolites. (A) Overall PDI, (B) provegetarian diet, (C) healthful PDI, (D) unhealthful PDI. β coefficients and P values were calculated from multivariable linear regression models which assessed the associations between PDIs and individual metabolites, adjusting for age, sex, race-center (only in subgroup 2), education, physical activity, smoking, alcohol consumption, margarine consumption, BMI, baseline estimated glomerular filtration rate, and total energy intake. Metabolites were meta-analyzed across 2 subgroups ($n_1 = 1762$; $n_2 = 1960$) using fixed-effects models. The horizontal line is set at the Bonferroni threshold of 3.34×10^{-5} (0.05/4 dietary patterns/374 metabolites), and the vertical line is set at a β coefficient of 0. PDI, plant-based diet index.

Prediction of plant-based dietary patterns with metabolites

The ability of plant-based diet–related metabolites to predict those in the highest quintile as opposed to the lower 4 quintiles of each of the PDIs improved significantly, compared with the model with only participant characteristics (range of differences in C-statistics = 0.026–0.104; P value for difference in C-statistics $\leq$ 0.001) (**Table 4**). In subgroup 1, the magnitude of increase was the highest for overall PDI (difference in C-statistics = 0.104), followed by healthful PDI (difference in C-statistics = 0.089). In subgroup 2, the magnitude of increase was the highest for provegetarian diet (difference in C-statistics = 0.076), followed by overall PDI (difference in C-statistics = 0.061). When we pooled subgroups 1 and 2, the magnitude of increase was the highest for overall PDI (difference in C-statistics = 0.074).

TABLE 3 Association of plant-based diet–related metabolites with incident CKD in Atherosclerosis Risk in Communities study participants[1]

Metabolite	Superpathway	Subpathway	HMDB ID	HR	SE	*P* value
1,5-anhydroglucitol (1,5-AG)	Carbohydrate	Glycolysis, gluconeogenesis, and pyruvate metabolism	HMDB02712	0.9002	0.0257	4.36×10^{-5}
Glycerate	Carbohydrate	Glycolysis, gluconeogenesis, and pyruvate metabolism	HMDB00139	0.8843	0.0258	1.91×10^{-6}
γ-Glutamylvaline	Peptide	γ-Glutamyl amino acid	HMDB11172	0.8872	0.0249	1.45×10^{-6}
γ-Glutamylleucine	Peptide	γ-Glutamyl amino acid	HMDB11171	0.8770	0.0248	1.23×10^{-7}
γ-Glutamylglutamate	Peptide	γ-Glutamyl amino acid	HMDB11737	0.8793	0.0247	2.04×10^{-7}
γ-Glutamylalanine	Peptide	γ-Glutamyl amino acid	HMDB29142	0.8743	0.0248	6.42×10^{-8}

[1]HR, SE, and *P* values were calculated from multivariable Cox proportional hazards models which assessed the associations between per 1 SD higher in plant-based diet–related metabolites and incident CKD, adjusting for age, sex, race-center (only in subgroup 2), education, physical activity, smoking, alcohol consumption, margarine consumption, BMI, baseline estimated glomerular filtration rate, total energy intake, diabetes, hypertension, and coronary artery disease. Metabolites were meta-analyzed across 2 subgroups ($n_1 = 1762$; $n_2 = 1960$) using fixed-effects models. Metabolites were considered statistically significant at the Bonferroni threshold of 9.80×10^{-4} (0.05/51 plant-based diet–related metabolites) for subgroups 1 and 2. Of 82 significant PDI–metabolite associations, 51 metabolites were unique (51 metabolites were associated with ≥1 dietary patterns). Over a median follow-up of 23 y, there were 606 and 807 incident CKD cases in subgroups 1 and 2, respectively. The 2 subgroups were mutually exclusive. CKD, chronic kidney disease; HMDB ID, human metabolome database identification number.

Prediction of incident CKD with metabolites

The addition of 6 plant-based diet–related metabolites modestly improved the prediction of incident CKD among those without diabetes relative to the CKD prediction equation [C-statistic for CKD prediction model = 0.702; C-statistic for CKD prediction model and 6 metabolites = 0.707; difference in C-statistics = 0.005 (95% CI: 0.002, 0.009); *P* value for difference in C-statistics = 0.006].

Sensitivity analyses

When we stratified 1,5-AG by diabetes status, 1,5-AG was positively associated with unhealthful PDI among those with and without diabetes (**Supplemental Table 3**). 1,5-AG was not significantly associated with incident CKD among those without diabetes, and was inversely associated with incident CKD among those with diabetes. The direction of the association was consistent between $eGFR_{Cr}$ and $eGFR_{Cys}$ for the 6 metabolites, with a slightly stronger magnitude for decline in $eGFR_{Cys}$ than for decline in $eGFR_{Cr}$ (**Supplemental Table 4**).

Discussion

In a large prospective study of individuals without CKD, we found 82 significant associations between PDIs and metabolites (overall PDI = 27; provegetarian = 17; healthful PDI = 20; unhealthful PDI = 18). Addition of these metabolites improved the prediction of individuals with high adherence to PDIs beyond sociodemographic characteristics, health behaviors, clinical factors, and total energy intake. We also identified 6 plant-based diet–related metabolites (glycerate, 1,5-AG, γ-glutamylalanine, γ-glutamylglutamate, γ-glutamylleucine, γ-glutamylvaline) which predicted the risk of incident CKD beyond traditional risk factors, highlighting potentially modifiable mechanisms leading to incident CKD.

To our knowledge, this is the first study to report metabolites associated with predefined PDIs. Of the metabolites associated with the overall PDI, provegetarian diet, and healthful PDI, several were previously reported as candidate biomarkers of other healthy dietary patterns, foods, and beverages. N-methylproline and stachydrine (also known as proline betaine), which were positively associated with the overall PDI in the present study, were also positively associated with the Healthy Eating Index (HEI) and Dietary Approaches to Stop Hypertension (DASH) diet in the ARIC study (32, 33), among postmenopausal women (34), and in a CKD population (35). N-methylproline and stachydrine are proline derivatives which are considered biomarkers of citrus fruit intake (36, 37). Importantly, they are osmoprotectants which accumulate in osmotically stressed conditions in human cells and provide protection against oxidative stress (38, 39). Tryptophan betaine (also known as hypaphorine), which was positively associated with overall PDI, provegetarian diet, and healthful PDI in our analysis, was reported as a biomarker of plant protein intake (chickpeas, lentils, total nuts) (36, 37, 40–43). Hypaphorine is found in extracts of legumes (44). This metabolite has been shown to reduce expression of inflammatory cytokines in human endothelial cells, and lower glucose in diabetic rats compared with rats that were administered metformin (45, 46). Hippurate, catechol sulfate, and quinate, which were also

TABLE 4 C-statistics and difference in C-statistics for prediction of individuals in the highest quintile as opposed to the lower 4 quintiles of PDIs in the Atherosclerosis Risk in Communities study[1]

	C-statistics for participant characteristics	Difference in C-statistics (95% CI)	*P* value[2]
Subgroup 1 (N = 1762)			
Overall PDI ($n = 27$)	0.623	0.104 (0.055, 0.152)	<0.001
Provegetarian ($n = 17$)	0.626	0.065 (0.028, 0.102)	0.001
Healthful PDI ($n = 20$)	0.679	0.089 (0.042, 0.136)	<0.001
Unhealthful PDI ($n = 18$)	0.725	0.051 (0.024, 0.078)	<0.001
Subgroup 2 (N = 1960)			
Overall PDI ($n = 27$)	0.639	0.061 (0.037, 0.085)	<0.001
Provegetarian ($n = 17$)	0.633	0.076 (0.049, 0.103)	<0.001
Healthful PDI ($n = 20$)	0.677	0.033 (0.013, 0.053)	0.001
Unhealthful PDI ($n = 18$)	0.665	0.037 (0.016, 0.058)	<0.001
Subgroups 1 and 2 (N = 3722)			
Overall PDI ($n = 27$)	0.626	0.074 (0.050, 0.097)	<0.001
Provegetarian ($n = 17$)	0.630	0.064 (0.043, 0.084)	<0.001
Healthful PDI ($n = 20$)	0.674	0.041 (0.023, 0.059)	<0.001
Unhealthful PDI ($n = 18$)	0.688	0.026 (0.011, 0.041)	<0.001

[1] We built logistic regression models with PDI (highest quintile compared with lower 4 quintiles) as the outcome and participant characteristics as exposures. Participant characteristics included age, sex, race-center (not included in subgroup 1), education, physical activity, smoking, alcohol consumption, margarine consumption, BMI, baseline estimated glomerular filtration rate, and total energy intake. Table 2 presents the list of metabolites significantly associated with plant-based diets. *n*, number of significant metabolites associated with the specific plant-based diet index; PDI, plant-based diet index.

[2] *P* value comparing C-statistics with participant characteristics and metabolites to C-statistics with only participant characteristics.

positively associated with the overall PDI, provegetarian diet, and healthful PDI, are metabolites that are derived from chlorogenic acid, which is found in fruits and coffee (44, 47). Such consistent associations with healthy dietary patterns could justify their prioritization as biomarkers of plant-based diets in the future, and suggest key pathways that are relevant for diet and CKD.

We found that the metabolites significantly associated with the unhealthful PDI were different from those associated with the overall PDI, provegetarian diet, and healthful PDI. Out of 18 metabolites associated with the unhealthful PDI, 13 were unique to this dietary pattern, and the direction of association was the opposite (e.g., catechol sulfate) for the metabolites also associated with the other PDIs. Pyridoxate and pantothenate, metabolites involved in vitamin B metabolism, were negatively associated with the unhealthful PDI in our study, which reflects nutritional characteristics of this dietary pattern (lower in B vitamins in general). In addition, we found that N2,N2-dimethylguanosine was positively associated with the unhealthful PDI. This metabolite is a purine nucleoside which results from degradation of transfer RNA, and a uremic solute which has been known to increase in plasma with lower urinary excretion. Greater plasma concentrations of uremic solutes have been associated with polycystic kidney disease (48) and progression to KFRT (49), because they can damage the proximal tubule through tissue hypoxia or endothelial cell injury. To date, no other studies have reported an association between N2,N2-dimethylguanosine and dietary patterns or foods. Because we did not find a statistically significant association between N2,N2-dimethylguanosine and incident CKD, future studies are warranted to assess whether N2,N2-dimethylguanosine represents a pathway through which unhealthful dietary patterns are associated with other adverse health outcomes.

Our observation that several γ-glutamyl peptides were positively associated with overall PDI, and inversely associated with incident CKD, is novel. These metabolites are involved in glutathione homeostasis, which is important for reducing oxidative stress (50). γ-glutamyl peptides are produced when γ-glutamyl transferase transfers the γ-glutamyl moiety of glutathione to amino acids (50). These metabolites are found in foods such as legumes (dry beans, soybeans), garlic, onion, and fermented foods, and are considered anti-inflammatory peptides (51). In vitro, γ-glutamyl valine has shown anti-inflammatory effects on human aortic endothelial cells by reducing production of IL-6 and IL-8 (52). In animal models, it has been reported that administration of γ-glutamyl valine inhibited TNF-α signaling in intestinal epithelial cells, and reduced expression of proinflammatory cytokines (TNF-α, IL-6, IL-1β) in the colon and adipocytes (53, 54). Another study found that when mice were fed high-fat diets for 6 wk, mice with the antiobesogenic phenotype gained less weight and had elevated levels of γ-glutamyl leucine compared with wild-type mice (55). These studies suggest that γ-glutamyl peptides may represent pathways that suppress inflammation and regulate adiposity. Inflammation and obesity are involved in the pathogenesis of several vascular conditions and CKD (56, 57). Taken together, the γ-glutamyl peptides highlight that plant-based diets may have anti-inflammatory and antiobesogenic properties.

Glycerate is another metabolite that was positively associated with the overall PDI, provegetarian diet, and healthful PDI, and negatively associated with incident CKD. Glycerate is a monosaccharide found in plant foods (tomatoes, plantains, grapes, peanuts), and a glycolytic intermediate which increases in the portal vein after consumption of dietary fructose (44, 58, 59). However, studies on the functional role of glycerate in

humans are limited. In our previous ARIC study and in a separate study of healthy male smokers, glycerate was positively associated with healthy dietary patterns (e.g., HEI-2015, Alterative Healthy Eating Index-2010, DASH, and Mediterranean-style diet), specifically fruits (32, 60). In a weight loss intervention, serum glycerate level increased from baseline to 1 y of follow-up in obese adults with ≥10% weight loss (61). In a cross-sectional study, plasma glycerate was negatively associated with insulin resistance (62). High blood sugar is another prominent factor of CKD (1). These results suggest that an increase in glycerate due to higher intake of fruit may play a role in glycemic control and weight loss, which may be beneficial for kidney health.

We found an unexpected association between 1,5-AG, unhealthful PDI, and incident CKD. 1,5-AG was positively associated with unhealthful PDI, but was inversely associated with incident CKD. 1,5-AG, a monosaccharide found in many foods, such as soybeans, is known as a marker of short-term (1–2 wk) hyperglycemia (63). In normoglycemia, the renal tubule reabsorbs nearly 100% of 1,5-AG (64). In the setting of hyperglycemia, 1,5-AG cannot be reabsorbed by the renal tubules owing to glucose and is excreted in the urine, which results in lower circulating concentrations of 1,5-AG (63). Using data from the ARIC study, low concentrations of 1,5-AG were associated with incident CKD, particularly among those with diabetes (31). In line with prior findings in the ARIC study, 1,5-AG was not significantly associated with incident CKD among those without diabetes when we stratified by diabetes status, but the positive association between 1,5-AG and unhealthful PDI remained significant. A recent review concluded that blood concentrations of 1,5-AG are unresponsive to dietary intake in the general population, because, in healthy individuals, blood glucose concentrations are not high enough to inhibit reabsorption of 1,5-AG (65). However, the review added that there is limited evidence on the extent to which food intake influences blood concentrations of 1,5-AG in individuals with normoglycemia. These results highlight the need for more research and suggests that our findings on 1,5-AG, unhealthful PDI, and incident CKD should be interpreted with caution.

The following limitations deserve consideration in interpreting our results. Metabolites were measured using serum specimens which had been in storage for >20 y. Metabolites may have degraded, but we would expect that to be nondifferential with respect to adherence to plant-based dietary patterns. In subgroup 1, fewer metabolites were available relative to subgroup 2 owing to differences in the timing of metabolomic profiling. The FFQ was modified from a validated questionnaire to assess dietary habits in the ARIC study population, but it is possible that the FFQ may not have covered all foods. We did not have data on albuminuria as a marker of kidney damage to include in our outcome definition of CKD and to include in the CKD prediction equation. Lastly, we did not require participants to have reduced eGFR for >3 mo when ascertaining incident CKD. However, omission of this chronicity criterion does not differentially bias the results by adherence to plant-based dietary patterns.

Our study has several strengths, including a large sample size, racially diverse sample, and the use of 2 different subgroups within a prospective study. We used established PDIs to evaluate participants' dietary habits, which will facilitate replication of our study design in other study populations. Despite the absence of albuminuria, ascertainment of CKD in the ARIC study was rigorous, which incorporated eGFR measured at follow-up visits and surveillance of CKD hospitalization, death, and KFRT over an extensive follow-up period. Such careful data collection allowed us to capture CKD events comprehensively, considering that CKD may have occurred in individuals who did not attend a follow-up visit. Further, we conducted a sensitivity analysis using eGFR decline as the outcome, and found that the direction of the association was the same with stronger associations when using $eGFR_{Cys}$ than using $eGFR_{Cr}$.

In conclusion, we identified 82 metabolites significantly associated with plant-based dietary patterns (overall PDI = 27; provegetarian = 17; healthful PDI = 20; unhealthful PDI = 18). Of these, 6 metabolites (glycerate, 1,5-AG, γ-glutamylalanine, γ-glutamylglutamate, γ-glutamylleucine, γ-glutamylvaline) were associated with incident CKD, and predicted incident CKD beyond traditional risk factors. These metabolites suggest that plant-based diets may be associated with a lower risk of incident CKD through reduced inflammation, prevention of adiposity, and glycemic control.

Some of the data reported here were supplied by the US Renal Data System (USRDS). The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as official policy or interpretation of the US government.

The authors' responsibilities were as follows—HK: wrote the manuscript, analyzed the data, and had primary responsibility for the final content; BY, XL, KEW, EB, SBS, ASL, EPR, and JC: contributed to interpretation of data and revised the manuscript critically; CMR: was involved in all aspects of the study from study design to analysis to critical revision of the manuscript; and all authors: read and approved the final manuscript. The authors report no conflicts of interest.

Conflict of Interest

All authors have no conflicts of interests to disclose.

Data Availability

ARIC (Atherosclerosis Risk in Communities) study data are available through the National Heart, Lung, and Blood Institute Biologic Specimen and Data Repository Information Coordinating Center. Interested researchers may also contact the ARIC Study Coordinating Center to access data and study materials. Analytic code will be made available upon reasonable request.

References

1. CDC. Chronic kidney disease in the United States, 2021. Atlanta (GA): US Department of Health and Human Services; 2021.
2. United States Renal Data System. USRDS annual data report: epidemiology of kidney disease in the United States. Bethesda (MD): NIH, National Institute of Diabetes and Digestive and Kidney Diseases; 2020.
3. Kelly JT, Palmer SC, Wai SN, Ruospo M, Carrero J-J, Campbell KL, Strippoli GFM. Healthy dietary patterns and risk of mortality and ESRD in CKD: a meta-analysis of cohort studies. Clin J Am Soc Nephrol 2017;12(2):272–9.
4. Bach KE, Kelly JT, Palmer SC, Khalesi S, Strippoli GFM, Campbell KL. Healthy dietary patterns and incidence of CKD: a meta-analysis of cohort studies. Clin J Am Soc Nephrol 2019;14(10):1441–9.
5. Hu EA, Steffen LM, Grams ME, Crews DC, Coresh J, Appel LJ, Rebholz CM. Dietary patterns and risk of incident chronic kidney

disease: the Atherosclerosis Risk in Communities study. Am J Clin Nutr 2019;110(3):713–21.
6. Kim H, Caulfield LE, Garcia-Larsen V, Steffen LM, Grams ME, Coresh J, Rebholz CM. Plant-based diets and incident CKD and kidney function. Clin J Am Soc Nephrol 2019;14(5):682–91.
7. Tzoulaki I, Ebbels TMD, Valdes A, Elliott P, Ioannidis JPA. Design and analysis of metabolomics studies in epidemiologic research: a primer on -omic technologies. Am J Epidemiol 2014;180(2):129–39.
8. Guasch-Ferré M, Bhupathiraju SN, Hu FB. Use of metabolomics in improving assessment of dietary intake. Clin Chem 2018;64(1): 82–98.
9. Whitfield PD, German AJ, Noble P-JM. Metabolomics: an emerging post-genomic tool for nutrition. Br J Nutr 2004;92(4):549–55.
10. The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) study: design and objectives. Am J Epidemiol 1989;129(4):687–702.
11. Stevens J, Metcalf PA, Dennis BH, Tell GS, Shimakawa T, Folsom AR. Reliability of a food frequency questionnaire by ethnicity, gender, age and education. Nutr Res 1996;16(5):735–45.
12. Satija A, Bhupathiraju SN, Rimm EB, Spiegelman D, Chiuve SE, Borgi L, Willett WC, Manson JE, Sun Q, Hu FB. Plant-based dietary patterns and incidence of type 2 diabetes in US men and women: results from three prospective cohort studies. PLoS Med 2016;13(6): e1002039.
13. Martínez-González MA, Sánchez-Tainta A, Corella D, Salas-Salvadó J, Ros E, Arós F, Gómez-Gracia E, Fiol M, Lamuela-Raventós RM, Schröder H, et al. A provegetarian food pattern and reduction in total mortality in the Prevención con Dieta Mediterránea (PREDIMED) study. Am J Clin Nutr 2014;100(suppl_1):320S–8S.
14. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Working Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int Suppl 2013;3(1):1–150.
15. Levey AS, Eckardt K-U, Dorman NM, Christiansen SL, Hoorn EJ, Ingelfinger JR, Inker LA, Levin A, Mehrotra R, Palevsky PM, et al. Nomenclature for kidney function and disease: report of a Kidney Disease: Improving Global Outcomes (KDIGO) Consensus Conference. Kidney Int 2020;97(6):1117–29.
16. Grams ME, Rebholz CM, McMahon B, Whelton S, Ballew SH, Selvin E, Wruck L, Coresh J. Identification of incident CKD stage 3 in research studies. Am J Kidney Dis 2014;64(2):214–21.
17. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF III, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150(9):604–12.
18. Coresh J, Astor BC, McQuillan G, Kusek J, Greene T, Van Lente F, Levey AS. Calibration and random variation of the serum creatinine assay as critical elements of using equations to estimate glomerular filtration rate. Am J Kidney Dis 2002;39(5):920–9.
19. Eckfeldt JH, Chambless LE, Shen YL. Short-term, within-person variability in clinical chemistry test results. Experience from the Atherosclerosis Risk in Communities Study. Arch Pathol Lab Med 1994;118(5):496–500.
20. Inker LA, Eckfeldt J, Levey AS, Leiendecker-Foster C, Rynders G, Manzi J, Waheed S, Coresh J. Expressing the CKD-EPI cystatin C equations for estimating GFR with standardized serum cystatin C values. Am J Kidney Dis 2011;58(4):682–4.
21. Inker LA, Coresh J, Sang Y, Hsu C, Foster MC, Eckfeldt JH, Karger AB, Nelson RG, Liu X, Sarnak M, et al. Filtration markers as predictors of ESRD and mortality: individual participant data meta-analysis. Clin J Am Soc Nephrol 2017;12(1):69–78.
22. Inker LA, Schmid CH, Tighiouart H, Eckfeldt JH, Feldman HI, Greene T, Kusek JW, Manzi J, Van Lente F, Zhang YL, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med 2012;367(1):20–9.
23. Evans AM, DeHaven CD, Barrett T, Mitchell M, Milgram E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. Anal Chem 2009;81(16):6656–67.
24. Ford L, Kennedy AD, Goodman KD, Pappan KL, Evans AM, Miller LAD, Wulff JE, Wiggs BR, Lennon JJ, Elsea S, et al. Precision of a clinical metabolomics profiling platform for use in the identification of inborn errors of metabolism. J Appl Lab Med 2020;5(2):342–56.
25. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW-M, Fiehn O, Goodacre R, Griffin JL, et al. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 2007;3(3):211–21.
26. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. Untargeted metabolomics strategies—challenges and emerging directions. J Am Soc Mass Spectrom 2016;27(12):1897–905.
27. Kim H, Rebholz CM, Garcia-Larsen V, Steffen LM, Coresh J, Caulfield LE. Operational differences in plant-based diet indices affect the ability to detect associations with incident hypertension in middle-aged US adults. J Nutr 2020;150(4):842–50.
28. Satija A, Bhupathiraju SN, Spiegelman D, Chiuve SE, Manson JE, Willett W, Rexrode KM, Rimm EB, Hu FB. Healthful and unhealthful plant-based diets and the risk of coronary heart disease in U.S. adults. J Am Coll Cardiol 2017;70(4):411–22.
29. Harris RJ, Deeks JJ, Altman DG, Bradburn MJ, Harbord RM, Sterne JAC. Metan: fixed- and random-effects meta-analysis. Stata J 2008;8(1):3–28.
30. Nelson RG, Grams ME, Ballew SH, Sang Y, Azizi F, Chadban SJ, Chaker L, Dunning SC, Fox C, Hirakawa Y, et al. Development of risk prediction equations for incident chronic kidney disease. JAMA 2019;322(21):2104–14.
31. Selvin E, Rawlings AM, Grams M, Klein R, Steffes M, Coresh J. Association of 1,5-anhydroglucitol with diabetes and microvascular conditions. Clin Chem 2014;60(11):1409–18.
32. Kim H, Hu EA, Wong K, Yu B, Steffen LM, Seidelmann SB, Boerwinkle E, Coresh J, Rebholz CM. Serum metabolites associated with healthy diets in African Americans and European Americans. J Nutr 2021;151(1):40–9.
33. Tariq A, Chen J, Yu B, Boerwinkle E, Coresh J, Grams ME, Rebholz CM. Metabolomics of dietary acid load and incident chronic kidney disease. J Ren Nutr 2021 Jul 19 (Epub ahead of print; doi: 10.1053/j.jrn.2021.05.005).
34. Playdon MC, Ziegler RG, Sampson JN, Stolzenberg-Solomon R, Thompson HJ, Irwin ML, Mayne ST, Hoover RN, Moore SC. Nutritional metabolomics and breast cancer risk in a prospective study. Am J Clin Nutr 2017;106(2):637–49.
35. Kim H, Anderson CA, Hu EA, Zheng Z, Appel LJ, He J, Feldman HI, Anderson AH, Ricardo AC, Bhat Z, et al. Plasma metabolomic signatures of healthy dietary patterns in the Chronic Renal Insufficiency Cohort (CRIC) study. J Nutr 2021;151(10):2894–907.
36. Guertin KA, Moore SC, Sampson JN, Huang W-Y, Xiao Q, Stolzenberg-Solomon RZ, Sinha R, Cross AJ. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr 2014;100(1):208–17.
37. Wang Y, Gapstur SM, Carter BD, Hartman TJ, Stevens VL, Gaudet MM, McCullough ML. Untargeted metabolomics identifies novel potential biomarkers of habitual food intake in a cross-sectional study of postmenopausal women. J Nutr 2018;148(6): 932–43.
38. Patriarca EJ, Cermola F, D'Aniello C, Fico A, Guardiola O, De Cesare D, Minchiotti G. The multifaceted roles of proline in cell behavior. Front Cell Dev Biol 2021;9:2236.
39. Thiemicke A, Neuert G. Kinetics of osmotic stress regulate a cell fate switch of cell survival. Sci Adv 2021;7(8):eabe1122.
40. Garcia-Aloy M, Ulaszewska M, Franceschi P, Estruel-Amades S, Weinert CH, Tor-Roca A, Urpi-Sarda M, Mattivi F, Andres-Lacueva C. Discovery of intake biomarkers of lentils, chickpeas, and white beans by untargeted LC–MS metabolomics in serum and urine. Mol Nutr Food Res 2020;64(13):1901137.
41. Playdon MC, Sampson JN, Cross AJ, Sinha R, Guertin KA, Moy KA, Rothman N, Irwin ML, Mayne ST, Stolzenberg-Solomon R, et al. Comparing metabolite profiles of habitual diet in serum and urine. Am J Clin Nutr 2016;104(3):776–89.
42. Wang Y, Hodge RA, Stevens VL, Hartman TJ, McCullough ML. Identification and reproducibility of plasma metabolomic biomarkers of habitual food intake in a US diet validation study. Metabolites 2020;10(10):382.
43. Zheng Y, Yu B, Alexander D, Steffen LM, Boerwinkle E. Human metabolome associates with dietary intake habits among African Americans in the Atherosclerosis Risk in Communities Study. Am J Epidemiol 2014;179(12):1424–33.

44. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, et al. HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res 2018;46(D1):D608–17.
45. Sun H, Zhu X, Lin W, Zhou Y, Cai W, Qiu L. Interactions of TLR4 and PPARγ, dependent on AMPK signalling pathway contribute to anti-inflammatory effects of Vaccariae hypaphorine in endothelial cells. Cell Physiol Biochem 2017;42(3):1227–39.
46. Chand K, Akanksha RN, Mishra DP, Srivastava AK, Maurya R. Major alkaloidal constituent from *Impatiens niamniamensis* seeds as antihyperglycemic agent. Med Chem Res 2011;20(9):1505–8.
47. Pallister T, Jackson MA, Martin TC, Zierer J, Jennings A, Mohney RP, MacGregor A, Steves CJ, Cassidy A, Spector TD, et al. Hippurate as a metabolomic marker of gut microbiome diversity: modulation by diet and relationship to metabolic syndrome. Sci Rep 2017;7(1):13670.
48. Grams ME, Tin A, Rebholz CM, Shafi T, Köttgen A, Perrone RD, Sarnak MJ, Inker LA, Levey AS, Coresh J. Metabolomic alterations associated with cause of CKD. Clin J Am Soc Nephrol 2017;12(11):1787–94.
49. Niewczas MA, Sirich TL, Mathew AV, Skupien J, Mohney RP, Warram JH, Smiles A, Huang X, Walker W, Byun J, et al. Uremic solutes and risk of end-stage renal disease in type 2 diabetes: metabolomic study. Kidney Int 2014;85(5):1214–24.
50. Lu SC. Glutathione synthesis. Biochim Biophys Acta 2013;1830(5):3143–53.
51. Yang J, Bai W, Zeng X, Cui C. Gamma glutamyl peptides: the food source, enzymatic synthesis, kokumi-active and the potential functional properties—a review. Trends Food Sci Technol 2019;91:339–46.
52. Guha S, Paul C, Alvarez S, Mine Y, Majumder K. Dietary γ-glutamyl valine ameliorates TNF-α-induced vascular inflammation via endothelial calcium-sensing receptors. J Agric Food Chem 2020;68(34):9139–49.
53. Zhang H, Kovacs-Nolan J, Kodera T, Eto Y, Mine Y. γ-Glutamyl cysteine and γ-glutamyl valine inhibit TNF-α signaling in intestinal epithelial cells and reduce inflammation in a mouse model of colitis via allosteric activation of the calcium-sensing receptor. Biochim Biophys Acta 2015;1852(5):792–804.
54. Xing L, Zhang H, Majumder K, Zhang W, Mine Y. γ-Glutamylvaline prevents low-grade chronic inflammation via activation of a calcium-sensing receptor pathway in 3T3-L1Mouse adipocytes. J Agric Food Chem 2019;67(30):8361–9.
55. Sansbury BE, Bhatnagar A, Hill BG. Impact of nutrient excess and endothelial nitric oxide synthase on the plasma metabolite profile in mice. Front Physiol 2014;5:453.
56. Shankar A, Sun L, Klein BEK, Lee KE, Muntner P, Nieto FJ, Tsai MY, Cruickshanks KJ, Schubert CR, Brazy PC, et al. Markers of inflammation predict the long-term risk of developing chronic kidney disease: a population-based cohort study. Kidney Int 2011;80(11):1231–8.
57. Rhee CM, Ahmadi S-F, Kalantar-Zadeh K. The dual roles of obesity in chronic kidney disease: a review of the current literature. Curr Opin Nephrol Hypertens 2016;25(3):208–16.
58. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 2000;28(1):27–30.
59. Jang C, Hui S, Lu W, Cowan AJ, Morscher RJ, Lee G, Liu W, Tesz GJ, Birnbaum MJ, Rabinowitz JD. The small intestine converts dietary fructose into glucose and organic acids. Cell Metab 2018;27(2):351–61.e3.
60. Playdon MC, Moore SC, Derkach A, Reedy J, Subar AF, Sampson JN, Albanes D, Gu F, Kontto J, Lassale C, et al. Identifying biomarkers of dietary patterns by using metabolomics. Am J Clin Nutr 2017;105(2):450–65.
61. Geidenstam N, Al-Majdoub M, Ekman M, Spégel P, Ridderstråle M. Metabolite profiling of obese individuals before and after a one year weight loss program. Int J Obes 2017;41(9):1369–78.
62. Hernández-Alonso P, García-Gavilán J, Camacho-Barcia L, Sjödin A, Hansen TT, Harrold J, Salas-Salvadó J, Halford JCG, Canudas S, Bulló M. Plasma metabolites associated with homeostatic model assessment of insulin resistance: metabolite-model design and external validation. Sci Rep 2019;9(1):13895.
63. Parrinello CM, Selvin E. Beyond HbA1c and glucose: the role of nontraditional glycemic markers in diabetes diagnosis, prognosis, and management. Curr Diab Rep 2014;14(11):548.
64. Yamanouchi T, Tachibana Y, Akanuma H, Minoda S, Shinohara T, Moromizato H, Miyashita H, Akaoka I. Origin and disposal of 1,5-anhydroglucitol, a major polyol in the human body. Am J Physiol 1992;263(2 Pt 1):E268–73.
65. Feskens E, Brennan L, Dussort P, Flourakis M, Lindner LME, Mela D, Rabbani N, Rathmann W, Respondek F, Stehouwer C, et al. Potential markers of dietary glycemic exposures for sustained dietary interventions in populations without diabetes. Adv Nutr 2020;11(5):1221–36.