Title

Circulating Metabolite Profiles of Diet Patterns in South Asians in the United States

Journal

Current Developments in Nutrition, 4(Suppl 2)

ISSN

2475-2991

Authors

Gadgil, Meghana
Sands, Caroline
Lewis, Matthew
et al.

Publication Date

2020-06-01

DOI

10.1093/cdn/nzaa061_030

Diet patterns are associated with circulating metabolites and lipid profiles of South Asians in the United States

Running head: Diet patterns and metabolomics in US South Asians

Meghana D. Gadgil[a], Alka M. Kanaya[a], Caroline Sands[b], Elena Chekmeneva[b]
Matthew R. Lewis[b], Namratha R. Kandula[c], David M. Herrington[d]

[a] Division of General Internal Medicine, Department of Medicine, University of California, San Francisco, CA; 1545 Divisadero Street, Suite 320, San Francisco, CA 94143-0320

[b] National Phenome Centre, Imperial College London, IRDB Building 5th Floor, Hammersmith Hospital Campus, London, W12 0NN

[c] Division of General Internal Medicine, Department of Medicine, Northwestern University Feinberg School of Medicine, 750 N. Lakeshore Dr. 10th Floor
Chicago, IL 60611

[d] Section on Cardiovascular Medicine, Department of Internal Medicine, Wake Forest School of Medicine; Medical Center Boulevard, Winston-Salem, NC 27157

Corresponding Author:
Meghana D. Gadgil
c/o University of California, San Francisco, Box 0320
1545 Divisadero Street
San Francisco, CA 94143-0320
Phone : (415) 353-7922
E-mail : meghana.gadgil@ucsf.edu

Word Count:
Abstract: 288
Text: 3007
Tables: 4
Supplemental Tables: 4

Abbreviations:
Mediators of Atherosclerosis in South Asians Living in America (MASALA); ultra-performance liquid chromatography mass spectrometry (UPLC-MS); non-alcoholic fatty liver disease (NAFLD); Metabolic-equivalent (MET) minutes/week

CONTRIBUTION STATEMENT

The roles for each author are as follows. MDG conceived of the project idea, analytic design and performed the analysis, CS, ML, AMK, NRK and DH contributed to the interpretation of the results and reviewed and edited the manuscript; MDG wrote the manuscript. MDG had primary responsibility for final content. All authors have read and approved the manuscript.

ABSTRACT

BACKGROUND: South Asians are at higher risk for cardiometabolic disease than many other racial/ethnic minority groups. Diet patterns in U.S. South Asians have unique components associated with cardiometabolic disease.

OBJECTIVES: We aimed to characterize the metabolites associated with three representative diet patterns.

METHODS: We included 722 participants in the Mediators of Atherosclerosis in South Asians Living in America (MASALA) cohort study aged 40-84 years without known cardiovascular disease. Fasting serum specimens and diet and demographic questionnaires were collected at baseline and diet patterns previously generated through principal components analysis. LC-MS-based untargeted metabolomic and lipidomic analysis was conducted with targeted integration of known metabolite and lipid signals. Linear regression models of diet pattern factor score and log-transformed metabolites adjusted for age, sex, caloric intake and body mass index and adjusted for multiple comparisons was performed, followed by elastic net linear regression of significant metabolites.

RESULTS: There were 443 metabolites of known identity extracted from the profiling data. The 'Animal protein' diet pattern was associated with 61 metabolites and lipids; including glycerophospholipids PE(O-16:1/20:4) and/or PE(P-16:0/20:4) (β 0.13; 95% Confidence Interval (CI) [0.11, 0.14]) and N-acyl phosphatidylethanolamines (NAPE) NAPE(O-18:1/20:4/18:0) and/or NAPE(P-18:0/20:4/18:0) (β: 0.13; 95%CI [0.11, 0.14]), LPI (22:6/0:0) (β 0.14; 95% CI [0.12, 0.17]) and fatty acids FA (22:6) (β: 0.15; 95% CI [0.13, 0.17]. The 'Fried snacks, sweets, high-fat dairy' pattern was associated with 12 lipids; including PC(16:0/22:6) (β: -0.08; 95%CI[-0.09, -0.06]) and FA(22:6) (β 0.14; 95%CI [-0.17, -0.10]). The 'Fruits,

Vegetables, Nuts and Legumes' pattern was associated with five metabolites including proline betaine (β: 0.17 [0.09, 0.25]) [$p < 0.0002$].

CONCLUSIONS: Three predominant dietary patterns in U.S. South Asians are associated with circulating metabolites differentiated by lipids including glycerophospholipids and polyunsaturated fatty acids and the amino acid proline betaine.

Keywords: diet patterns; South Asian; cardiovascular risk; metabolomics; lipids

BACKGROUND

South Asians (individuals of Indian, Pakistani, Bangladeshi, Nepali, and Sri Lankan descent) are at higher risk for cardiovascular disease and diabetes than many other racial and ethnic groups.(1-3) About 23% of South Asians have diabetes, which often precedes coronary heart disease.(4) On average, South Asians develop coronary heart disease 10 years earlier than people identifying as a different race or ethnic group, and 50% of heart attacks in South Asians occur before the age of 50.(5)

In this population, diet quality and pattern of intake is a strong, modifiable risk factor for cardiometabolic disease.(6, 7) Prior investigations with South Asian populations in the diaspora have characterized unique diet patterns influenced by both heritage and emigration.(8, 9) We previously examined diet patterns in the Mediators of Atherosclerosis in South Asians Living in America (MASALA) cohort, in which habitual dietary intake was characterized with a culturally-concordant food frequency questionnaire.(10) We identified three major diet patterns: Animal Protein; Fried snacks, Sweets, High-fat Dairy, and Fruits, Vegetables, Nuts, Legumes, (11) which have unique associations with traditional risk factors for cardiometabolic disease. The increased risk associated with certain diet patterns may be tied to intermediate metabolic markers seen in South Asians, such as a pattern of atherogenic dyslipidemia, tendency towards larger ectopic adipose tissue deposits and lower muscle mass and poor beta cell function.(2, 12). Little is known about the mechanisms and process by which these particular diet profiles translate into metabolic phenotypes that cause higher cardiometabolic risk.

The identification of metabolites, easily measurable and present in serum, urine or tissue can help to shed light on the phenotypic links between diet and cardiometabolic disease in this

high-risk population.(13) A panel of metabolites may both be able to serve as a biomarker for diet intake and help clarify and measure the metabolic effects of that diet intake.

In this analysis, we aimed to establish representative metabolites for predominant diet patterns in South Asians who are part of the full MASALA study.

METHODS

Participants

Data were from South Asians who participated in the MASALA community-based cohort study and had complete diet and metabolomic data. The detailed methods have been described elsewhere.(10) Briefly, MASALA is a prospective cohort study which enrolled community-dwelling individuals living in the San Francisco Bay Area and the greater Chicago areas from 2010-2013. Participants self-identified as having South Asian ancestry and were aged 40-84 years and without known cardiovascular disease. Those on nitroglycerin, with active cancer, with impaired cognitive ability, a life expectancy less than five years, who lived in a nursing home, or who had plans to relocate were excluded. The University of California, San Francisco and Northwestern University Institutional Review Board approved the study protocol and all study participants provided written informed consent.

Demographic and diet data

Each participant underwent in-person interviews to determine age, sex, medical history, physical activity, smoking status and alcohol intake. Food group intake was collected with the Study of Health Assessment and Risk in Ethnic (SHARE) groups South Asian Food Frequency Questionnaire, which was developed and validated in South Asians in Canada.(14) The food frequency questionnaire included 163 items, with 61 items unique to the South Asian diet, and assessed usual eating habits, frequency and serving sizes over the prior 12 months.

Dietary pattern creation

Individual food items from the SHARE food frequency questionnaire were divided into 29 predefined subgroups reflecting likeness, underlying nutrient composition and South Asian culinary usage. Several foods (e.g. coffee) were kept as individual categories given their high reported intake. We excluded 1 individual with incomplete food frequency questionnaire data and another 6 who did not meet *a priori* criteria of daily caloric ranges (600-6000 kcal/24 h).

Principal components analysis with varimax rotation was previously used to identify the most prevalent groupings of major food group categories in our population.(15) After identifying three patterns that explained the majority of variance, the patterns were named according to their major components: “Animal protein” (9.3% variance), “Fried snacks, sweets, high-fat dairy” (7.4% variance) “Fruits, vegetables, nuts legumes” (6.5% variance). Each participant was assigned a factor score for each dietary pattern based on the correlation of his or her food frequency questionnaire data with the food groupings in the three prevalent patterns. The diet patterns: Animal Protein, Fried Snacks, Sweets, High-fat Dairy each had continuous factor scores which were divided into tertiles for ease of interpretation.

Metabolic Profiling by UPLC-MS

A total of 754 serum samples obtained at the baseline exam (2010-2013) were analyzed by ultra-performance liquid chromatography mass spectrometry (UPLC-MS) using previously described analytical and quality control procedures.(16, 17) Sample analysis was performed in an order designed to be orthogonal to clinical and demographic data metadata. For quality control assessment and data pre-processing, a study reference (SR) sample was prepared by pooling equal parts of each study sample.

Serum samples were prepared and analyzed using UPLC-MS as previously published. (16, 17) In brief, 50 µL aliquots were taken from each sample, diluted 1:1 with ultrapure water for lipid profiling and 1:1.4 for small molecule profiling. Protein was removed by addition of organic solvent to the diluted sample (four volumes isopropanol per volume of diluted sample for lipidomic profiling and three volumes of acetonitrile per volume of diluted sample for small molecule profiling) followed by mixing and centrifugation to yield a homogenous supernatant. Aliquot sets of prepared samples were subjected to chromatographic separation using an ACQUITY UPLC (Waters Corp., Milford, MA, USA) system. Lipidomic profiling was performed using reversed-phase chromatography (RPC) with a 2.1 × 100 mm Acquity BEH C8 column maintained at 55°C. The chromatographic separation was performed using a binary mobile phase system consisting of (A) a 50:25:25 mixture of H_2O:ACN:IPA with 5mm ammonium acetate, 0.05% acetic acid, and 20µM phosphoric acid and (B) 50:50 ACN:IPA with 5mm ammonium acetate, 0.05% acetic acid. Polar metabolite profiling was performed using hydrophilic interaction liquid chromatography (HILIC) with a 2.1 × 150 mm Acquity BEH HILIC column maintained at 40°C. The chromatographic separation was performed using a binary mobile phase system consisting of (A) acetonitrile with 0.1% formic acid and (B) 20 mM ammonium formate in water with 0.1% formic acid. Both separation types were coupled to high resolution mass spectrometry (Xevo G2-S TOF mass spectrometers, Waters Corp., Manchester, UK) via a Z-spray electrospray ionization source. The lipidomic profiling assay was conducted in both positive and negative ion modes (generating Lipid RPC+ and Lipid RPC- datasets), while the HILIC assay was performed in the positive ion mode only (generating the HILIC+ dataset). A SR sample was acquired every 10 study samples throughout the analysis. In addition, a

dilution series was created from the SR and analyzed immediately prior to and after the study sample analysis for use in signal filtering as described previously (16).

Raw data was converted to the mzML open source format and signals below an absolute intensity threshold of 100 counts were removed using the MSConvert tool in ProteoWizard.(18) Metabolite signal extraction was performed using PeakPantheR, an open-source package to detect, integrate and report pre-defined and annotated lipids and metabolites from an in-house database.(19) Elimination of potential run-order effects and filtering of the extracted metabolites was performed using the nPYc-Toolbox, an open-source package for data pre-processing.(20) Only those measured with high accuracy (relative coefficient of variance in SR samples less than 20%) and high precision (correlation to dilution in SR dilution series greater than 0.8) were retained and put forward for biological analysis. Of the 754 total study samples, 32 were not included in our analysis due to insufficient sample volume and five were excluded due to missed injection in the HILIC assay.

Cardiometabolic factors measured at baseline:

Weight was determined using a digital scale, height with a stadiometer, and waist circumference using a measuring tape halfway between the lower ribs and the anterior superior iliac spine, at the site of greatest circumference. Blood samples were obtained after a requested 12-hour fast. Fasting plasma glucose was measured using the hexokinase method (Quest diagnostics, San Jose, CA). An oral glucose tolerance test was performed, in which participants consumed a 75g oral glucose solution, and blood samples for plasma glucose and insulin were taken after 120 minutes. Type 2 diabetes was defined as a fasting glucose $\geq$ 126 mg/dl, 2-hour post-challenge glucose $\geq$ 200 mg/dl or use of a glucose-lowering medication. 717 participants were included in our analysis.

Statistical methods

Before modeling, relative abundance of metabolites were log-transformed to reduce the potential for outliers to influence the model. Multivariable linear regression analyses were used to determine associations of diet pattern factor score and relative abundance of each independent metabolite. The analyses were adjusted for age, sex, calories per day, body mass index in Model 1 and further adjusted for presence of diabetes, hypertension, use of statin medication, smoking and alcohol intake of $\geq$ 1 drink/wk as categorical variables and exercise (metabolic-equivalent-minutes/week) as a continuous variable. We applied the conservative Bonferroni method to adjust for multiple comparisons, with an alpha<0.00002 deemed significant. To adjust for unreliable parameter estimates that may occur when using multiple regression models in the setting of multicollinearity, we performed an elastic net regularized regression model to evaluate metabolites that were significant in independent analyses. The elastic-net model allowed for a penalized logistic regression on all biomarkers simultaneously to identify the metabolites most highly associated with diet pattern score. Optimal parameters for the penalty value (α) and the regularization penalty (λ) were determined by 10-fold cross-validation. Briefly, data in the full dataset were randomly assigned to one of two equal sized datasets. Model performance was judged based on root mean square error, with the model chosen minimizing mean cross-validated error. Optimazationw as complete using STATA's "elasticnet" and postestimation commands for model prediction. We then further adjusted these linear regression models for physical activity, diabetes, and family history of diabetes.

The analysis was completed using STATA (version 16.1, 2021, College Station, TX, USA).

RESULTS

In total, 443 metabolites and lipids were examined in this analysis (Supplemental Table 1). MASALA participants in the highest tertile of factor score of the Animal Protein pattern were less likely to be women, had lower total daily energy intake but were of similar BMI than those who most often consumed the Fried snacks, Sweets, High-fat dairy or Fruits, Vegetables, Nuts, Legumes patterns (Table 1). A similar proportion of women most often consumed the Fried snacks, Sweets, High-fat dairy pattern and Fruits, Vegetables, Nuts, Legumes diet patterns (47%).

After elastic net regularized regression, and further adjustment for relevant covariates, the Animal Protein diet pattern was associated with 61 metabolite and lipid species. It was positively associated with phospholipids, sphingomyelins, ceramides and other lipid species including omega-3 fatty acids, and negatively associated with long-chain acylcarnitines and trigonelline. The metabolites most highly associated with the Animal Protein diet pattern were: PE(O-16:1/20:4) and/or PE(P-16:0/20:4) (0.13; 95% Confidence Interval [0.11, 0.14]) and NAPE (O-18:1/20:4/18:0) and/or NAPE(P-18:0/20:4/18) (0.13; 95%CI [0.11, 0.14]), LPI (22:6/0:0) (0.14; 95% CI [0.12, 0.17]) and fatty acids FA (22:6) (0.15; 95% CI [0.13, 0.17]. (Table 2) The Fried snacks, Sweets, High-fat dairy pattern was associated with 12 lipids, the top two hits of which were PC(16:0/22:6) (-0.08; 95%CI[-0.09, -0.06]) and FA(22:6) (0.14; 95%CI [-0.17, -0.10]) (Table 3). The Fruits, Vegetables, Nuts, Legumes diet was associated with five metabolites, including a positive association with proline betaine (0.17 [0.09, 0.25]) (Table 4).

DISCUSSION

Participants in the MASALA study consumed three predominant dietary patterns: Animal Protein; Fried Snacks, Sweets, High-Fat Dairy; Fruits, Vegetables, Nuts, Legumes, which were each associated with particular metabolite and lipid patterns. The metabolic profile associated

with Animal Protein pattern was represented by glycerophospholipids, acylcarnitines and ceramides, which carry high metabolic risk. The Fried snacks, Sweets, High-fat dairy pattern was inversely associated a number of lipids, including an omega-3 fatty acid derived from seafood and linked to lower cardiovascular risk.(21) Higher consumption of the Fruits, Vegetables, Nuts, Legumes pattern was associated with higher abundance of proline betaine, a marker of citrus consumption and lower risk for type 2 diabetes in prior studies(22), and lower relative abundance of several lipid subspecies.

The metabolite and lipid patterns associated with high consumption of each diet pattern have implications for metabolic health. In particular, proline betaine was positively associated with the most "prudent" diet pattern, Fruits, Vegetables, Nuts, Legumes, and negatively associated with the Animal Protein diet pattern. There is a correlation between proline betaine and fruit intake in this sample (Supplemental Table 4). This amino acid, and its analogue, glycine betaine, have been associated with lower risk for diabetes in the Diabetes Prevention Program and other trials and cohort studies.(22, 23) Betaine is derived from the amino acid glycine, and acts as a methyl donor to allow the conversion of homocysteine to methionine. (24) Proline betaine is also a biomarker of citrus consumption.(25) Deficiency of betaine is additionally linked with increased severity of non-alcoholic fatty liver disease (NAFLD).(26) In our prior work, the Fruits, Vegetables, Nuts, Legumes was associated with lower prevalence of metabolic syndrome.(11) Despite these positive observational findings and promising preclinical data from animal studies, direct supplementation of betaine in humans during a randomized, controlled trial showed only minor improvements in fasting glucose, and no changes in dynamic measurements of Insulin sensitivity and intrahepatic triglycerides.(27) All together, this suggests that an

exploration of the choline-betaine metabolic pathways and downstream metabolites may yield insights into the pathogenesis of prediabetes and NAFLD.

Long and short-chain acylcarnitines have previously been associated with prevalent and incident diabetes.(28, 29) Short-chain acylcarnitines, specifically CAR(3:0) and CAR(5:0) acylcarnitines, are breakdown products of BCAA metabolism and are associated with insulin resistance.(30) In a previous assessment diet patterns and metabolites in the MASALA pilot study (N=150), a similar "Western/non-vegetarian" diet pattern was associated with short-chain acylcarnitines.(31) In our study, there was a direct association between increased consumption of the Animal Protein pattern and propionylcarnitine (CAR(3:0)). There have been conflicting associations between long-chain acylcarnitines and the presence of diabetes. Impaired fatty acid oxidation and oxidative stress due to peripheral insulin resistance may cause a buildup of long-chain acylcarnitines (29) resulting in a decrease in insulin synthesis and associations with prevalent diabetes. Conversely, several cohort studies, including the PREDIMED study and our prior work in the MASALA study, show inverse associations between long-chain acylcarnitine abundance and both prevalent diabetes and future glucose intolerance.(32, 33) The current investigation found a positive association between the Animal Protein pattern and CAR(18:0) and an inverse association with CAR (14:2), CAR (18:3), CAR(20:2) and CAR (20:3). In support of the association between this diet pattern and circulating CAR (18:0), a randomized trial of red meat intake also revealed positive associations with CAR(18:0). (34) Our prior work in the MASALA cohort identified a relationship between higher baseline CAR(18:0) and subsequent lower HbA1c at 5-year follow-up (33) in cohort members without baseline diabetes. These findings suggest that animal protein intake is associated with CAR(18:0), however further

associations with diabetes are varied in this population, and may depend on the prevalence of other diet components.

The Animal Protein pattern was also associated with a higher abundance of multiple ceramides and sphingomyelins, including Cer(d18:1/26:1) and SM(d18:1/18:0). Ceramides, which are bioactive sphingolipids, have strong ties with diabetes risk (35, 36). Both circulating ceramides and sphingomyelins have been associated with impaired glucose homeostasis. (37, 38) Ceramides are also associated with intake of saturated fats, and with non-alcoholic fatty liver disease.(39)

In our investigation, NAPEs were associated with consumption of the Animal Protein pattern, and were correlated with red meat, poultry, fish, eggs and coffee intake (Supplemental Table 2). NAPEs hydrolysis generates N-acylethanolamines that are precursors of endocannabinoids synthesized in phospholipid membranes. Endocannabinoids may be involved in signaling between gut microbiotia and adipose tissue, and have been implicated in metabolic disorders such as obesity and type 2 diabetes.(40) In some reports NAPEs have been shown to be increased in plasma after high-fat feeding and regulate food intake.(41) Phosphotidylethanolamines PE(O-16:1/20:4) and/or PE(P-16:0/20:4) were significantly and positively associated with intake of the Animal Protein pattern, and is an essential bioactive lipid abundant in mammalian cells.(42) One study has shown a potential link between this broader lipid species class and decreased odds of acute coronary syndrome,(43) however the particular risk conferred by the lipids found in our analysis are not known.

Several important polyunsaturated fatty acids differed between patterns, including lipids with docosahexaenoic acid (DHA) FA(22:6) moieties. These omega-3 fatty acids are correlated with the major non-vegetarian components, including red meat, poultry, eggs and fish

consumption in the Animal Protein pattern (Supplemental Table 2), negatively correlated with butter/ghee and legume intake in the Fried Snacks, Sweets, High-fat Dairy diet pattern (Supplemental Table 3) and have previously been linked with lower risk of cardiovascular disease.(44) Lipids with these moieties are lower in abundance with greater consumption of the Fried snacks, sweets and high-fat dairy pattern, suggesting that there may be lower consumption of these potentially beneficial fatty acids in this unhealthful vegetarian pattern. Omega-6 fatty acids found in lean meat, milk and eggs contain arachidonic acid FA (20:4), which is abundant in phospholipids, and important for cellular signaling in the brain and skeletal muscle, and was higher with consumption of the Animal Protein pattern. High levels of this fatty acid may be affected by oxidative stress and play a role in the pathogenesis of fatty liver and diabetes (45) and cardiovascular disease.(46)

In conclusion, our findings suggest that prevalent diet patterns in the MASALA study are associated with groups of metabolites and lipids linked with cardiometabolic disease. The Fruits, Vegetables, Nuts, Legumes patterns associated with proline betaine which has been linked with reduced risk for diabetes. The Animal Protein pattern was associated with NAPEs, sphingomyelins and ceramides and long- and short-chain acylcarnitines. Furthermore, the Animal Protein and Fried snacks, Sweets, High-fat dairy patterns had opposite associations with long-chain omega 3 fatty acids, which have been linked with lower risk of cardiovascular disease. These conclusions are limited by the absence of data on intra-individual variability of metabolites. These findings support the next steps in investigation of diet and metabolites: the study of metabolites as biomarkers for measuring diet quality and to determine targeted dietary advice to reduce risk of cardiometabolic disease.

ACKNOWLEDGEMENTS

MDG is supported by 1K23DK119404-01A1 from the National Institute of Diabetes and Digestive and Kidney Diseases.

The MASALA study was supported by Grant Number R01HL093009 from the National Heart, Lung, And Blood Institute and the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through UCSF-CTSI Grant Number UL1RR024131. The metabolomic measurements were supported by an anonymous gift to Dr. Kanaya. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, And Blood Institute or the National Institutes of Health. The authors thank the other investigators, the staff, and the participants of the MASALA study for their valuable contributions.

This study was also supported by the Medical Research Council and National Institute for Health Research (NIHR) (grant NC-PC-12025, infrastructure support was provided by the NIHR Imperial Biomedical Research Centre (BRC).

DATA SHARING

Data described in the manuscript, code book, and analytic code will be made available upon request pending request to MASALA Study Steering Committee for reasons of participant confidentiality.

CONFLICT OF INTEREST

The authors do not have financial conflicts of interest to disclose.

REFERENCES CITED

1. Kanaya AM, Herrington D, Vittinghoff E, Ewing SK, Liu K, Blaha MJ, et al. Understanding the high prevalence of diabetes in U.S. south Asians compared with four racial/ethnic groups: the MASALA and MESA studies. Diabetes care. 2014;37(6):1621-8.
2. Narayan KMV, Kanaya AM. Why are South Asians prone to type 2 diabetes? A hypothesis based on underexplored pathways. Diabetologia. 2020;63(6):1103-9.
3. Anand SS, Yusuf S, Vuksan V, Devanesen S, Teo KK, Montague PA, et al. Differences in risk factors, atherosclerosis and cardiovascular disease between ethnic groups in Canada: the study of health assessment and risk in ethnic groups (SHARE). Indian heart journal. 2000;52(7 Suppl):35.
4. Cheng YJ, Kanaya AM, Araneta MRG, Saydah SH, Kahn HS, Gregg EW, et al. Prevalence of Diabetes by Race and Ethnicity in the United States, 2011-2016. JAMA. 2019;322(24):2389-98.
5. Volgman AS, Palaniappan LS, Aggarwal NT, Gupta M, Khandelwal A, Krishnan AV, et al. Atherosclerotic Cardiovascular Disease in South Asians in the United States: Epidemiology, Risk Factors, and Treatments: A Scientific Statement From the American Heart Association. Circulation. 2018;138(1):e1-e34.
6. Mohan V, Radhika G, Sathya RM, Tamil SR, Ganesan A, Sudha V. Dietary carbohydrates, glycaemic load, food groups and newly detected type 2 diabetes among urban Asian Indian population in Chennai, India (Chennai Urban Rural Epidemiology Study 59). The British journal of nutrition. 2009;102(10):1498-506.
7. Bhattacharya M. A historical exploration of Indian diets and a possible link to insulin resistance syndrome. Appetite. 2015;95:421-54.
8. Lesser IA, Gasevic D, Lear SA. The association between acculturation and dietary patterns of South Asian immigrants. PloS one. 2014;9(2):e88495.
9. Talegawkar SA, Kandula NR, Gadgil MD, Desai D, Kanaya AM. Dietary intakes among South Asian adults differ by length of residence in the USA. Public health nutrition. 2016;19(2):348-55.
10. Kanaya AM, Kandula N, Herrington D, Budoff MJ, Hulley S, Vittinghoff E, et al. Mediators of Atherosclerosis in South Asians Living in America (MASALA) Study: Objectives, Methods, and Cohort Description. Clinical cardiology. 2013.
11. Gadgil MD, Anderson CA, Kandula NR, Kanaya AM. Dietary patterns are associated with metabolic risk factors in South Asians living in the United States. The Journal of nutrition. 2015;145(6):1211-7.
12. Bancks MP, Bertoni AG, Carnethon M, Chen H, Cotch MF, Gujral UP, et al. Association of Diabetes Subgroups With Race/Ethnicity, Risk Factor Burden and Complications: The MASALA and MESA Studies. J Clin Endocrinol Metab. 2021.
13. Playdon MC, Moore SC, Derkach A, Reedy J, Subar AF, Sampson JN, et al. Identifying biomarkers of dietary patterns by using metabolomics. The American Journal of Clinical Nutrition. 2016.
14. Kelemen LE, Anand SS, Vuksan V, Yi Q, Teo KK, Devanesen S, et al. Development and evaluation of cultural food frequency questionnaires for South Asians, Chinese, and Europeans in North America. Journal of the American Dietetic Association. 2003;103(9):1178-84.

15. Gadgil MD, Anderson CAM, Kandula NR, Kanaya AM. Dietary Patterns and Associations with Metabolic Risk Factors in South Asians. Diabetes. 2014;63(Supplement 1):A343-A425.
16. Lewis MR, Pearce JT, Spagou K, Green M, Dona AC, Yuen AH, et al. Development and Application of Ultra-Performance Liquid Chromatography-TOF MS for Precision Large Scale Urinary Metabolic Phenotyping. Anal Chem. 2016;88(18):9004-13.
17. Izzi-Engbeaya C, Comninos AN, Clarke SA, Jomard A, Yang L, Jones S, et al. The effects of kisspeptin on beta-cell function, serum metabolites and appetite in humans. Diabetes Obes Metab. 2018;20(12):2800-10.
18. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, et al. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol. 2012;30(10):918-20.
19. Wolfer AM, Correia GDS, Sands CJ, Camuzeaux S, Yuen AHY, Chekmeneva E, et al. peakPantheR, an R package for large-scale targeted extraction and integration of annotated metabolic features in LC-MS profiling datasets. Bioinformatics. 2021.
20. Sands CJ, Wolfer AM, Correia GDS, Sadawi N, Ahmed A, Jimenez B, et al. The nPYc-Toolbox, a Python module for the pre-processing, quality-control and analysis of metabolic profiling datasets. Bioinformatics. 2019;35(24):5359-60.
21. Micha R, Penalvo JL, Cudhea F, Imamura F, Rehm CD, Mozaffarian D. Association Between Dietary Factors and Mortality From Heart Disease, Stroke, and Type 2 Diabetes in the United States. JAMA. 2017;317(9):912-24.
22. Walford GA, Ma Y, Clish C, Florez JC, Wang TJ, Gerszten RE, et al. Metabolite Profiles of Diabetes Incidence and Intervention Response in the Diabetes Prevention Program. Diabetes. 2016;65(5):1424-33.
23. Sun Y, Gao HY, Fan ZY, He Y, Yan YX. Metabolomics Signatures in Type 2 Diabetes: A Systematic Review and Integrative Analysis. J Clin Endocrinol Metab. 2020;105(4).
24. Craig SA. Betaine in Human Nutrition. Am J Clin Nutr. 2004;80:539-49.
25. Guertin KA, Moore SC, Sampson JN, Huang WY, Xiao Q, Stolzenberg-Solomon RZ, et al. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. The American Journal of Clinical Nutrition. 2014;100(1):208-17.
26. Morgan EKKMGNBCSMACTBSWFaTR. Betaine improves nonalcoholic fatty liver and associated insulin resistance: a potential mechanism for hepatoprotection by betaine. Am J Physiol Gastrointest Liver Physiol. 2010;299(5):G1068–G77.
27. Grizales AM PM, Lin AP, et al. Metabolic Effects of Betaine: A Randomized Clinical Trial of Betaine Supplementation in Prediabetes. . J Clin Endocrinol Metab.103(8):3038-49.
28. Mihalik SJ, Goodpaster BH, Kelley DE, Chace DH, Vockley J, Toledo FG, et al. Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipotoxicity. Obesity (Silver Spring). 2010;18(9):1695-700.
29. Sun L, Liang L, Gao X, Zhang H, Yao P, Hu Y, et al. Early Prediction of Developing Type 2 Diabetes by Plasma Acylcarnitines: A Population-Based Study. Diabetes care. 2016;39(9):1563-70.
30. Lee CC, Watkins SM, Lorenzo C, Wagenknecht LE, Il'yasova D, Chen YD, et al. Branched-Chain Amino Acids and Insulin Metabolism: The Insulin Resistance Atherosclerosis Study (IRAS). Diabetes care. 2016;39(4):582-8.

31. Bhupathiraju SN, Guasch-Ferre M, Gadgil MD, Newgard CB, Bain JR, Muehlbauer MJ, et al. Dietary Patterns among Asian Indians Living in the United States Have Distinct Metabolomic Profiles That Are Associated with Cardiometabolic Risk. J Nutr. 2018;148(7):1150-9.
32. Guasch-Ferre M, Zheng Y, Ruiz-Canela M, Hruby A, Martinez-Gonzalez MA, Clish CB, et al. Plasma acylcarnitines and risk of cardiovascular disease: effect of Mediterranean diet interventions. The American Journal of Clinical Nutrition. 2016;103(6):1408-16.
33. Gadgil MD, Kanaya AM, Sands C, Lewis MR, Kandula NR, Herrington DM. Circulating metabolites and lipids are associated with glycaemic measures in South Asians. Diabetic Medicine. 2020;38(2).
34. Wedekind R, Kiss A, Keski-Rahkonen P, Viallon V, Rothwell JA, Cross AJ, et al. A metabolomic study of red and processed meat intake and acylcarnitine concentrations in human urine and blood. Am J Clin Nutr. 2020;112(2):381-8.
35. Chaurasia B, Summers SA. Ceramides - Lipotoxic Inducers of Metabolic Disorders. Trends Endocrinol Metab. 2015;26(10):538-50.
36. Chen GC, Chai JC, Yu B, Michelotti GA, Grove ML, Fretts AM, et al. Serum sphingolipids and incident diabetes in a US population with high diabetes burden: the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). Am J Clin Nutr. 2020;112(1):57-65.
37. Seah JYH, Chew WS, Torta F, Khoo CM, Wenk MR, Herr DR, et al. Dietary Fat and Protein Intake in Relation to Plasma Sphingolipids as Determined by a Large-Scale Lipidomic Analysis. Metabolites. 2021;11(2).
38. Rosqvist F, Iggman D, Kullberg J, Cedernaes J, Johansson HE, Larsson A, et al. Overfeeding polyunsaturated and saturated fat causes distinct effects on liver and visceral fat accumulation in humans. Diabetes. 2014;63(7):2356-68.
39. Summers SA. Ceramides: Nutrient Signals that Drive Hepatosteatosis. J Lipid Atheroscler. 2020;9(1):50-65.
40. Cani PD, Plovier H, Van Hul M, Geurts L, Delzenne NM, Druart C, et al. Endocannabinoids--at the crossroads between the gut microbiota and host metabolism. Nat Rev Endocrinol. 2016;12(3):133-43.
41. Gillum MP, Zhang D, Zhang XM, Erion DM, Jamison RA, Choi C, et al. N-acylphosphatidylethanolamine, a gut- derived circulating factor induced by fat ingestion, inhibits food intake. Cell. 2008;135(5):813-24.
42. Patel D, Witt SN. Ethanolamine and Phosphatidylethanolamine: Partners in Health and Disease. Oxid Med Cell Longev. 2017;2017:4829180.
43. Meikle PJ, Formosa MF, Mellett NA, Jayawardana KS, Giles C, Bertovic DA, et al. HDL Phospholipids, but Not Cholesterol Distinguish Acute Coronary Syndrome From Stable Coronary Artery Disease. J Am Heart Assoc. 2019;8(11):e011792.
44. Abdelhamid AS, Brown TJ, Brainard JS, Biswas P, Thorpe GC, Moore HJ, et al. Omega-3 fatty acids for the primary and secondary prevention of cardiovascular disease. Cochrane Database Syst Rev. 2020;3:CD003177.
45. Sonnweber T, Pizzini A, Nairz M, Weiss G, Tancevski I. Arachidonic Acid Metabolites in Cardiovascular and Metabolic Diseases. Int J Mol Sci. 2018;19(11).
46. Zhou Y, Khan H, Xiao J, Cheang WS. Effects of Arachidonic Acid Metabolites on Cardiovascular Health and Disease. Int J Mol Sci. 2021;22(21).

Table 1: Baseline characteristics of MASALA Study participants by tertile of diet pattern[1]

	Animal Protein			**Fried snacks, Sweets, High-fat dairy**			**Fruits, Vegetables, Nuts Legumes**		
	alcohol, coffee, eggs, fish, pasta, pizza, poultry, red meat, refined grains			added fat, butter/ghee, fried snacks, high-fat dairy, sugar-sweetened beverages, legumes, potatoes, refined grains, rice, whole grains			fruit, fruit juice, legumes, nuts, vegetable oil, vegetables, wholegrain		
	Tertile 1	Tertile 2	Tertile 3	Tertile 1	Tertile 2	Tertile 3	Tertile 1	Tertile 2	Tertile 3
	245	233	225	234	238	231	229	239	235
Women, (%)	141 (58)	105 (45)	70 (31)	122 (52)	112 (47)	82 (36)	97 (42)	115 (48)	104 (44)
Age, years	56 (9)	56 (9)	54 (9)	56 (10)	56 (10)	55(9)	54 (9)	55 (9)	57 (9)
BMI, (kg/m^2)	26 (4)	26 (4)	26 (4)	26 (4)	26 (4)	26 (4)	26 (4)	26 (4)	26 (4)
Smoker, (%)									
Never	228	15	2	85	83	81	182 (79)	202 (85)	199 (85)
Former	192	36	5	13	13	15	33 (14)	35 (15)	29 (12)
Current	163	46	16	3	4	3	14 (6)	2 (1)	7 (3)
Alcohol use, n (%)	33 (13)	74 (31)	133 (59)	86 (37)	78 (33)	76 (33)	82 (36)	77 (32)	81 (34)
Exercise[2], MET-min/wk	983 [1515]	1080 [1523]	945 [1300]	1170 [1845]	960 [1560]	735 [1542]	690 [1365]	1035 [1523]	1373 [1665]
Energy intake, kcal/d	1630 (461)	1613 (444)	1820 (555)	1366 (354)	1662 (413)	2035 (470)	1348 (342)	1664 (393)	2037 (480)
Diabetes, n (%)	77 (26)	75 (25)	72 (24)	64 (27)	60 (25)	50 (22)	65 (28)	60 (25)	49 (21)
Hypertension, n (%)	96 (39)	97 (42)	90 (40)	95 (41)	105 (44)	83 (36)	93 (41)	104 (44)	86 (37)

[1]Values are mean (SD), median [IQR], or frequency (percent)
[2]Metabolic-equivalent (MET) minutes per week

Table 2: Metabolites Associated with Animal Protein Diet Pattern, Elastic Net Regularized Regression

	Model 1, adjusted for age, sex, body mass index, energy intake [2,3]				**Model 2, fully adjusted[4]**			
	ß[2]	95% CI (Lower)	95% CI (Upper)	p-value	ß[3]	95% CI (Lower)	95% CI (Upper)	p-value
Creatine	0.08	0.06	0.1	1.21×10^{-15}	0.08	0.06	0.10	1.47×10^{-14}
Carnitine	0.02	0.01	0.03	2.79×10^{-07}	0.02	0.01	0.03	9.13×10^{-07}
CAR(3:0) (Propionylcarnitine)	0.03	0.02	0.05	1.89×10^{-05}	0.04	0.02	0.06	8.99×10^{-06}
Proline betaine	-0.11	-0.17	-0.06	1.41 x 10-05	-0.12	-0.17	-0.06	3.34×10^{-05}
Pipecolate/N-methyl proline	-0.07	-0.1	-0.03	1.34 x 10-04	-0.07	-0.10	-0.03	3.67 x 10-04
Taurine	0.03	0.02	0.05	3.28×10^{-05}	0.04	0.02	0.05	8.82×10^{-05}
Trigonelline	-0.16	-0.19	-0.12	1.89×10^{-17}	-0.17	-0.21	-0.13	7.93×10^{-18}
Trimethylaminoacetone	0.04	0.03	0.06	1.55×10^{-06}	0.04	0.02	0.06	5.51×10^{-05}
Methyladenosine	-0.01	-0.02	-0.01	1.79×10^{-04}	-0.01	-0.02	-0.00	1.60×10^{-03}
CAR14:2	-0.05	-0.07	-0.03	4.91×10^{-05}	-0.04	-0.06	-0.01	6.62×10^{-03}
CAR18:0	0.03	0.02	0.05	1.47×10^{-06}	0.03	0.02	0.05	4.82×10^{-05}
CAR18:3	-0.04	-0.06	-0.02	1.03×10^{-04}	-0.03	-0.05	-0.01	0.01
CAR20:1	-0.04	-0.05	-0.02	3.33×10^{-05}				
CAR20:2	-0.06	-0.07	-0.04	4.72×10^{-11}	-0.05	-0.07	-0.04	4.47×10^{-09}
CAR20:3	-0.05	-0.07	-0.03	3.56×10^{-08}	-0.05	-0.07	-0.03	3.13×10^{-07}
CE20:4	0.03	0.02	0.04	1.79×10^{-06}	0.03	0.01	0.04	4.42×10^{-05}
CE20:5	0.1	0.07	0.12	4.08×10^{-15}	0.08	0.06	0.11	3.20×10^{-10}
Cerd18:1/26:1	0.07	0.05	0.09	4.51×10^{-12}	0.06	0.04	0.08	1.17×10^{-08}

Cerd20:1/24:0	0.07	0.05	0.09	8.36×10^{-10}	0.07	0.05	0.09	3.33×10^{-08}
SulfoHexCerd18:1/18:0	0.05	0.03	0.06	1.25×10^{-08}	0.05	0.03	0.06	1.51×10^{-07}
SulfoHexCerd18:1/22:0OH	0.05	0.03	0.07	2.48×10^{-09}	0.05	0.03	0.07	2.19×10^{-08}
SulfoHexCerd18:1/24:0OH	0.03	0.01	0.05	1.40×10^{-04}	0.03	0.01	0.05	3.57×10^{-04}
CERd18:2/18:0	0.05	0.03	0.07	3.62×10^{-07}				
LPC17:1/0:0	0.03	0.01	0.04	6.36×10^{-05}	0.03	0.01	0.04	5.00×10^{-04}
LPC20:0/0:0	-0.04	-0.06	-0.02	1.23×10^{-06}	-0.03	-0.05	-0.02	1.19×10^{-04}
LPC20:1/0:0	-0.04	-0.06	-0.03	1.35×10^{-06}	0.04	-0.06	0.02	3.63×10^{-05}
LPC22:6/0:0	0.1	0.08	0.12	4.78×10^{-26}	0.09	0.07	0.11	5.76×10^{-19}
LPE18:1/0:0	-0.04	-0.06	-0.02	1.73×10^{-06}	-0.04	-0.06	-0.02	4.43×10^{-06}
LPE18:2/0:0	-0.05	-0.07	-0.04	1.23×10^{-10}	-0.05	-0.06	0.03	3.76×10^{-08}
PE16:0/20:3 and PE18:1/18:2 and PE18:0/18:3	-0.07	-0.09	-0.05	2.57×10^{-12}	0.07	-0.09	-0.05	6.63×10^{-10}
PC16:0/18:3	-0.05	-0.07	-0.03	2.92×10^{-06}	0.05	-0.07	-0.03	6.42×10^{-06}
PC16:0/20:4_2	0.02	0.01	0.04	1.21×10^{-05}	0.02	0.01	0.03	3.21×10^{-03}
PC16:0/22:4	-0.03	-0.05	-0.02	1.15×10^{-04}	-0.03	-0.05	-0.01	3.55×10^{-04}
PC18:1/20:3	-0.06	-0.07	-0.04	9.17×10^{-14}	-0.05	-0.07	-0.04	1.20×10^{-11}
PC34:0 PC18:0/16:0 and PC16:0/18:0	0.02	0.01	0.04	5.01×10^{-05}	0.02	0.01	0.03	1.92×10^{-04}
PE16:0/18:1	-0.05	-0.07	-0.03	2.13×10^{-04}	-0.06	-0.08	-0.04	1.51×10^{-06}
PE18:1/18:2	-0.07	-0.09	-0.05	1.21×10^{-10}	-0.07	-0.09	-0.04	8.29×10^{-09}
PE20:1/20:4	-0.07	-0.09	-0.05	1.29×10^{-04}	-0.08	-0.10	-0.06	3.90×10^{-11}
LPCP-18:0/0:0	0.07	0.06	0.09	5.58×10^{-21}	0.07	0.05	0.09	3.85×10^{-17}

PCO-16:0/18:2	0.07	0.06	0.09	8.72×10^{-21}	0.08	0.06	0.09	1.96×10^{-24}
PE18:0/16:0	-0.05	-0.07	-0.03	1.15×10^{-08}	-0.06	-0.07	0.04	4.25×10^{-09}
SMd18:1/18:0	0.03	0.02	0.04	5.00×10^{-09}	0.03	0.02	0.04	2.72×10^{-07}
SMd35:1	0.03	0.02	0.04	3.96×10^{-06}	0.03	0.02	0.05	6.53×10^{-07}
TG52:4	-0.03	-0.05	-0.02	1.72×10^{-04}	-0.02	-0.04	-0.01	0.01
FA18:4	0.05	0.03	0.08	3.28×10^{-05}	0.05	0.03	0.08	1.50×10^{-04}
FA20:4	0.03	0.01	0.04	1.47×10^{-04}				
FA22:5_2	0.09	0.06	0.11	2.36×10^{-13}	0.08	0.05	0.10	3.68×10^{-10}
FA22:6	0.15	0.13	0.17	1.41×10^{-32}	0.14	0.11	0.16	8.90×10^{-25}
FA24:5	0.07	0.05	0.1	3.86×10^{-07}	0.07	0.04	0.10	8.25×10^{-06}
LPI16:1/0:0	0.1	0.07	0.13	1.07×10^{-09}	0.09	0.06	0.13	1.19×10^{-07}
LPI20:4/0:0	0.04	0.03	0.06	9.51×10^{-08}	0.04	0.02	0.05	1.98×10^{-06}
LPI22:6/0:0	0.14	0.12	0.17	1.14×10^{-31}	0.14	0.11	0.16	6.41×10^{-25}
PEO-16:1/20:4 and/or PEP-16:0/20:4	0.13	0.11	0.14	3.35×10^{-63}	0.13	0.11	0.14	2.44×10^{-56}
PAO-18:1/20:4 and/or PAP-18:0/20:4	0.13	0.11	0.15	3.11×10^{-36}	0.13	0.11	0.15	2.32×10^{-30}
NAPEO-18:1/18:1/16:0 and/or NAPEP-18:0/18:1/16:0	0.03	0.02	0.05	4.60×10^{-07}				
PEO-18:2/20:4 and/or PEP-18:1/20:4	0.1	0.09	0.11	8.18×10^{-40}	0.1	0.09	0.12	6.93×10^{-36}
PEO-16:0/22:6	0.11	0.09	0.12	8.78×10^{-41}	0.11	0.09	0.12	1.23×10^{-35}
PEO-18:1/18:2 and/or PEP-18:0/18:2	0.04	0.03	0.06	6.30×10^{-09}	0.05	0.03	0.06	1.51×10^{-09}

NAPEO-18:1/20:4/16:0 and/or NAPEP-18:0/20:4/16:0	0.1	0.09	0.11	6.58×10^{-43}	0.1	0.08	0.11	3.95 x 10-35
PEO-18:1/18:1 and/or PE-P-18:0/18:1	0.09	0.07	0.1	9.05×10^{-25}	-0.05	-0.07	-0.04	1.20×10^{-11}
NAPEO-18:1/20:4/18:0 and/or NAPEP-18:0/20:4/18:0	0.13	0.11	0.14	4.63×10^{-47}	0.12	0.11	0.14	9.35×10^{-39}

[1] All metabolites significant at $p<0.0002$

[2] Adjusted for age, sex, body mass index, energy intake

[3] Increase in log-metabolites per 1-point increase in dietary pattern scores

[4] Further adjusted for diabetes, normoglycemic prediabetes or diabetes, hypertension, statin use, physical activity (MET-min/wk), alcohol use $\geq$ drink/wk, smoking

Metabolic-Equivalent minutes per week = MET-min/wk

Table 3: Metabolites Associated with Fried snacks Sweets High-fat Dairy Diet Pattern Elastic Net Regularized Regression adjusted for age sex body mass index energy intake[1]

[1]

	Model 1, adjusted for age, sex, body mass index, energy intake [2,3]				**Model 2, fully adjusted[4]**			
	ß[2]	95% CI Lower	95% CI Upper	p-value	ß[3]	95% CI Lower	95% CI Upper	p-value
PC16:0/22:6	-0.09	-0.11	-0.06	1.16×10^{-08}	-0.07	-0.10	-0.04	1.30×10^{-05}
LPE18:2/0:0	0.05	0.02	0.07	4.37×10^{-05}	0.03	0.01	0.06	4.13×10^{-03}
LPI22:6/0:0	-0.08	-0.12	-0.04	1.00×10^{-05}	-0.07	-0.11	-0.03	3.08×10^{-04}
PC16:0/22:4	0.07	0.04	0.09	1.73×10^{-08}	0.05	0.03	0.08	1.48×10^{-05}
PC18:0/22:4	0.08	0.05	0.11	1.71×10^{-07}	0.07	0.04	0.10	6.49×10^{-06}
PC16:0/20:5	-0.09	-0.13	-0.06	1.65×10^{-08}	-0.07	-0.11	-0.04	8.54×10^{-06}
PC16:0/22:6	-0.08	-0.09	-0.06	1.71×10^{-16}	-0.06	-0.08	-0.05	1.22×10^{-11}
PA16:0/18:1	-0.04	-0.06	-0.02	8.83×10^{-05}	-0.04	-0.06	-0.01	2.59×10^{-03}
FA22:6	-0.14	-0.17	-0.1	1.96×10^{-13}	-0.11	-0.15	-0.07	7.22×10^{-09}
PC18:0/22:5	0.07	0.04	0.11	1.52×10^{-05}	0.07	0.03	0.10	9.11×10^{-05}
SulfoHexCerd18:1/24:0-OH	-0.05	-0.07	-0.02	6.15×10^{-05}	-0.05	-0.07	-0.02	5.74×10^{-05}
SulfoHexCerd18:1/24:1-OH	-0.05	-0.07	-0.03	3.55×10^{-05}	-0.04	-0.07	-0.02	2.72×10^{-04}

All metabolites significant at $p<0.0002$

[2] Adjusted for age, sex, body mass index, energy intake

[3] Increase in log-metabolites per 1-point increase in dietary pattern scores

[4] Further adjusted for diabetes, normoglycemic prediabetes or diabetes, hypertension, statin use, physical activity (MET-min/wk), alcohol use ≥ drink/wk, smoking

Metabolic-Equivalent minutes per week = MET-min/wk

Table 4: Metabolites Associated with Fruits Vegetables Nuts Legumes Diet Pattern Elasticnet Regularized Regression adjusted for age sex body mass index energy intake[1]

	Model 1, adjusted for age, sex, body mass index, energy intake [2,3]				**Model 2, fully adjusted[4]**			
	ß[2]	95% CI Lower	95% CI Upper	p-value	ß[3]	95% CI Lower	95% CI Upper	p-value
Proline betaine	0.17	0.09	0.25	1.0×10^{-04}	0.18	0.09	0.26	3.65×10^{-05}
LPC22:4/0:0	-0.08	-0.11	-0.04	5.86×10^{-06}	-0.07	-0.10	-0.03	1.17×10^{-04}
PC18:0/22:4	-0.06	-0.09	-0.03	1.0×10^{-04}	-0.05	-0.09	-0.20	1.67×10^{-03}
SMd19:1/16:0	-0.07	-0.1	-0.04	7.40×10^{-05}	-0.06	-0.09	-0.03	1.42×10^{-05}
LPE22:4/0:0	-0.12	-0.18	-0.07	1.36×10^{-05}	-0.11	-0.17	-0.05	1.43×10^{-04}

[1] All metabolites significant at $p<0.0002$

[2] Adjusted for age, sex, body mass index, energy intake

[3] Increase in log-metabolites per 1-point increase in dietary pattern scores

[4] Further adjusted for diabetes, normoglycemic prediabetes or diabetes, hypertension, statin use, physical activity (MET-min/wk), alcohol use ≥ drink/wk, smoking

Metabolic-Equivalent minutes per week = MET-min/wk