

Associations of network-derived metabolite clusters with prevalent type 2 diabetes among adults of Puerto Rican descent

Danielle E Haslam ^{1,2}, Liming Liang,^{3,4} Dong D Wang ^{1,2}, Rachel S Kelly,¹ Clemens Wittenbecher ², Cynthia M Pérez,⁵ Marijulie Martínez,⁶ Chih-Hao Lee ⁷, Clary B Clish ⁸, David T W Wong,⁹ Laurence D Parnell,¹⁰ Chao-Qiang Lai ¹⁰, José M Ordovás,^{11,12} JoAnn E Manson,^{1,4,13} Frank B Hu,^{1,2,4} Meir J Stampfer,^{1,2,4} Katherine L Tucker,¹⁴ Kaumudi J Joshipura,^{4,6} Shilpa N Bhupathiraju ^{1,2}

To cite: Haslam DE, Liang L, Wang DD, *et al.* Associations of network-derived metabolite clusters with prevalent type 2 diabetes among adults of Puerto Rican descent. *BMJ Open Diab Res Care* 2021;**9**:e002298. doi:10.1136/bmjdr-2021-002298

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjdr-2021-002298>).

Part of this study was presented as an abstract at the American Heart Association Scientific Sessions, November 13–17, 2020, a virtual experience.

Received 29 March 2021
Accepted 25 July 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Danielle E Haslam;
nhdah@channing.harvard.edu

ABSTRACT

Introduction We investigated whether network analysis revealed clusters of coregulated metabolites associated with prevalent type 2 diabetes (T2D) among Puerto Rican adults.

Research design and methods We used liquid chromatography-mass spectrometry to measure fasting plasma metabolites (>600) among participants aged 40–75 years in the Boston Puerto Rican Health Study (BPRHS; discovery) and San Juan Overweight Adult Longitudinal Study (SOALS; replication), with (n=357; n=77) and without (n=322; n=934) T2D, respectively. Among BPRHS participants, we used unsupervised partial correlation network-based methods to identify and calculate metabolite cluster scores. Logistic regression was used to assess cross-sectional associations between metabolite clusters and prevalent T2D at the baseline blood draw in the BPRHS, and significant associations were replicated in SOALS. Inverse-variance weighted random-effect meta-analysis was used to combine cohort-specific estimates.

Results Six metabolite clusters were significantly associated with prevalent T2D in the BPRHS and replicated in SOALS (false discovery rate (FDR) <0.05). In a meta-analysis of the two cohorts, the OR and 95% CI (per 1 SD increase in cluster score) for prevalent T2D were as follows for clusters characterized primarily by glucose transport (0.21 (0.16 to 0.30); FDR <0.0001), sphingolipids (0.40 (0.29 to 0.53); FDR <0.0001), acyl cholines (0.35 (0.22 to 0.56); FDR <0.0001), sugar metabolism (2.28 (1.68 to 3.09); FDR <0.0001), branched-chain and aromatic amino acids (2.22 (1.60 to 3.08); FDR <0.0001), and fatty acid biosynthesis (1.54 (1.29 to 1.85); FDR <0.0001). Three additional clusters characterized by amino acid metabolism, cell membrane components, and aromatic amino acid metabolism displayed significant associations with prevalent T2D in the BPRHS, but these associations were not replicated in SOALS.

Conclusions Among Puerto Rican adults, we identified several known and novel metabolite clusters that associated with prevalent T2D.

INTRODUCTION

The prevalence of type 2 diabetes (T2D) differs substantially by race and ethnicity.

Significance of this study

What is already known about this subject?

- Few studies have used network analysis approaches to identify plasma metabolites associated with type 2 diabetes (T2D), and none have been conducted among individuals of Puerto Rican descent, who are at high risk for T2D and related comorbidities.

What are the new findings?

- We have identified six metabolite clusters within a global metabolite network that are associated with prevalent T2D among adults of Puerto Rican descent.
- A novel cluster of acyl cholines were associated with lower odds of T2D.
- We replicated several previously observed associations and revealed novel coregulated metabolites in clusters, including branched-chain and aromatic amino acids, sphingolipids, and cell membrane components, along with metabolites related to glucose transport, fatty acid biosynthesis, and sugar metabolism.

How might these results change the focus of research or clinical practice?

- We identified several metabolite pathways that are perturbed among individuals of Puerto Rican descent with T2D.
- These metabolites should be explored further in prospective studies and studies examining lifestyle, genetic, or environmental factors that may influence these metabolites.

Individuals of Puerto Rican (PR) descent experience particularly high rates of T2D and death from related comorbidities.^{1–3} Plasma metabolomics may help us gain an in depth understanding of the underlying molecular processes contributing to T2D among this high-risk population.⁴ However, much of the

literature examining metabolomics and T2D focuses on non-Hispanic white populations.^{5–11}

Previous studies implicate several potential metabolites that positively associate with T2D, including 2-aminoadipic acid,⁵ acylcarnitines,^{9 12} glutamate,⁶ and metabolites related to aromatic and branched-chain amino acid (BCAA) metabolism,^{8 10 11} and glutamine negatively associates with T2D.⁶ However, none of these associations has been evaluated among PRs, who have a unique genetic background and unique lifestyle habits compared with non-Hispanic white populations.^{13 14} Further, most of these studies used candidate or single-metabolite approaches, with only a few leveraging the correlations between metabolites to identify pathways or networks that may be perturbed among those with T2D.^{15–17} Distinct subtypes of metabolic alterations often precede or accompany T2D and are defined by distinct patterns of diabetes-related risk factors, such as obesity, insulin resistance, or dyslipidemia.^{18–20} Examination of the associations between metabolites and diabetes-related risk factors may help identify novel mechanisms underlying these subtypes. Thus, integration of metabolic profiles of T2D with related risk factors among individuals of PR descent could reveal novel metabolic changes occurring after T2D diagnosis and biomarker candidates for early detection of T2D in this high-risk population.

In this study, we used a data-driven, network-based method to reduce 614 named metabolites to distinct clusters that may reflect underlying biological connections between metabolites among PR adults. Then, we used a cross-sectional design to examine whether these clusters were associated with prevalent T2D and diabetes-related risk factors among 679 participants of the Boston Puerto Rican Health Study (BPRHS) and 1011 participants of the San Juan Overweight Adult Longitudinal Study (SOALS). The goal was to identify potential metabolites or metabolic pathways that may be good candidate biomarkers for T2D in PR adults that could be explored in future studies.

RESEARCH DESIGN AND METHODS

Study participants

The BPRHS (discovery) is a well-characterized, population-based longitudinal cohort study of 1500 PR adults aged 45–75 years living in or near Boston, Massachusetts. A previous publication provides details about the design and data collection for the BPRHS.²¹ Participants were recruited primarily through door-to-door enumeration from areas of high Hispanic density, but also through advertisement in flyers, participation in community events, and referrals from participants. Bilingual interviewers visited participants' homes to collect data at baseline (2004) and follow-up visits.

The SOALS (replication) is a longitudinal, population-based cohort study of 1300 individuals aged 40–65 years who were residents of the San Juan municipality and its vicinity and were overweight or obese (body mass index

(BMI) ≥ 25 kg/m²). Additional details about the design and data collection are provided in previous publications.^{22–24} Participants were recruited through different types of mass media, and data collection began in 2011 (baseline).

In this study, we included all BPRHS participants with available metabolomic profiling (n=679). Of these, 322 participants free of T2D (controls) were included in the derivation of the metabolic networks and associations with diabetes-related risk factors. The 357 participants with T2D at the time of blood draw (ie, prevalent cases) were used to examine the associations between metabolic network clusters and prevalent T2D. In an important step to follow best practices in omics-based studies,²⁵ we replicated our findings in a total of 1011 SOALS participants with metabolomics profiling available. Of these, 934 participants were free of T2D and 77 met the criteria for T2D.

Blood sample collection and metabolomics profiling

All participants were asked to fast for 12 hours in the BPRHS and 10 hours in the SOALS before the blood draw. In the BPRHS, blood was drawn at participants' homes, and a portable centrifuge was used to separate the plasma. The samples were transported to the laboratory in coolers with dry ice, processed, and stored at -70°C . In SOALS, samples were drawn at the baseline visit, centrifuged, and stored at -80°C . Blinded specimens were assayed by Metabolon (Durham, North Carolina) in 2017 for the BPRHS and in 2019 for SOALS, as previously described.²⁶ Briefly, chemical peaks were identified using positive and negative ionization ultra-high-performance liquid chromatography-tandem mass spectrometry and gas chromatography-mass spectrometry. The software developed at Metabolon was used to identify 943 known compounds in the BPRHS and 1062 in SOALS, with differences due to the number of known peaks at the time of the analysis.

Internal quality control samples were included, and injection order was random with respect to case status. Several processing steps were implemented before analysis. Metabolites with a detection rate $<75\%$ and xenobiotics were removed. Among the remaining metabolites, undetectable values were imputed at a value equal to half the minimum of each measured metabolite. To reduce the impact of outliers and skewed distributions, we applied an inverse normal transformation. After processing, 614 named metabolites were available for analysis in both the BPRHS and SOALS cohorts.

T2D cases and related risk factors

Prevalent T2D at baseline was defined using the American Diabetes Association criteria: fasting plasma glucose concentration ≥ 126 mg/dL (7.0 mmol/L), glycosylated hemoglobin (HbA1c) $\geq 6.5\%$ (48 mmol/mol), or use of antidiabetic medication.²⁷ In SOALS, additional T2D cases were identified by 2-hour fasting glucose concentration ≥ 200 mg/dL (11.0 mmol/L) during an oral

glucose tolerance test (OGTT), a measure that was not available among BPRHS participants. In the BPRHS, glucose was measured using an enzymatic, kinetic reaction (OSCR6121; Olympus America, Melville, New York), where the intra-assay and interassay coefficients of variation (CV%) for serum glucose were 2% and 3.4%, respectively. In SOALS, glucose was assessed using a SIRRUS analyzer (intra-assay CV%: 1.21%; interassay: 3.06%) and HbA1c was assessed using a latex immunoagglutination inhibition method (intra-assay CV%: 2.89%; interassay: 1.88%).

Additional diabetes-related risk factors assayed among BPRHS and SOALS participants at baseline include insulin, Homeostatic Model Assessment for Insulin Resistance (HOMA-IR), HbA1c, high-density lipoprotein cholesterol (HDL-C), and triglycerides (TG). All analytes were measured using standardized protocols with intra-assay and interassay CV% ranging from 1.5% to <10%. Height and weight were measured in duplicate at baseline visits.^{21 22} BMI was calculated as weight (kg) divided by height (m²). Waist circumference (WC) was measured at the umbilical level and recorded to the nearest 0.1 cm.

Measurement of covariates

Validated questionnaires were used to collect information on age, education, household income, medication use, family history, and health behaviors, including medication use and history of smoking and alcohol consumption. A validated Food Frequency Questionnaire (FFQ) designed for the PR population was used to assess habitual food consumption and nutrient intakes in the BPRHS.²⁸ Reported food intakes were converted to food groups to assess adherence to the American Heart Association Diet Score (AHA-DS).²⁹ An FFQ was administered to only a subset of SOALS participants, so we did not include diet as a covariate in SOALS analyses. Physical activity was assessed using a modified version of the Paffenbarger questionnaire^{30 31} in BPRHS participants. A metabolic equivalent (MET) score was calculated based on a questionnaire that assessed the time and frequency of physical activities during a typical week among SOALS participants. We assessed stress using the Spanish version of the Perceived Stress Scale.³² In the BPRHS, acculturation was captured through the bidimensional Acculturation Scale for Hispanics, which yields a score of up to 100, indicating full acculturation with fluent use of the English language³³ (not applicable in SOALS).

Statistical methods

Among participants without T2D, normalized levels of all available metabolites were used to construct a global metabolite network (ie, network capturing an extensive range of biological processes) based on Spearman's rank partial correlations separately in the BPRHS and SOALS using the *igraph* R package (<https://cran.r-project.org>).³⁴ Each correlation coefficient for a pair of metabolites was calculated conditioning on the remaining metabolites. All edges and paths between metabolites with a

Benjamini-Hochberg false discovery rate (FDR) <0.05 in both the BPRHS and SOALS networks were retained to create a final global metabolite network. The structure and connectedness of this large metabolic network was examined using an algorithm developed to identify distinct clusters (greedy optimization algorithm³⁵) and calculate weights for metabolites within each cluster.³⁶ The metabolite weights were then used to calculate a weighted sum of metabolite levels among participants for each cluster (metabolite cluster score).

We first evaluated the relationships between all metabolite cluster scores (continuous) and prevalent T2D through logistic regression models in the discovery cohort (BPRHS). OR and 95% CI correspond to the odds of prevalent T2D with each 1 SD increase in the sum of weighted metabolite concentrations in each cluster. Second, clusters significantly associated with T2D in the BPRHS (FDR <0.05) were tested in the replication cohort (SOALS). For replication, weights derived from the final global network were used to compute continuous cluster scores in SOALS. Third, ORs in the discovery and replication cohorts for clusters found to be significantly associated with prevalent T2D in the BPRHS were combined through inverse-variance weighted, fixed-effects meta-analyses using the *meta* R package (<https://cran.r-project.org>). The I² statistic was used to assess heterogeneity among the cohorts. We built three multivariable models to assess the influence of covariates: model 1, adjusted for age and sex; model 2, adjusted for model 1 covariates plus smoking status (former, current, never), education (≤8th grade, 9th–12th grade or GED (General Educational Development Test), college/some graduate school), physical activity (continuous score (BPRHS) or METs (SOALS)), alcohol intake (non-drinker, moderate (women: 1 drink/day; men: 1–2 drinks/day), heavy (women: >1 drink/day; men: >2 drinks/day)), lipid-lowering medication (yes/no), hypertension medication (yes/no), income (<\$20 000/year, ≥\$20 000/year), acculturation (%; BPRHS only), Perceived Stress Score (continuous score), diet quality (AHA-DS continuous score; BPRHS only), mouthwash use (yes/no; SOALS only), and family history of T2D (yes/no; SOALS only); and model 3, adjusted for model 2 covariates plus WC and BMI. We subjectively assigned cluster names according to pathway membership of the majority of or heavily weighted metabolites in that cluster. Sensitivity analyses were conducted replacing the weighted metabolite cluster score with an unweighted score, along with a joint model including all clusters significantly associated with prevalent T2D in the BPRHS. To provide additional insight into which metabolites may be driving the overall cluster associations, we evaluated the associations between individual metabolites within clusters that significantly associated with prevalent T2D in both cohorts.

For the metabolite clusters significantly associated with prevalent T2D in the BPRHS (FDR <0.05), we examined cross-sectional associations between metabolite cluster scores and diabetes-related risk factors among participants in the

BPRHS and SOALS participants without T2D. Linear regression models adjusting for model 3 covariates (excluding BMI and WC for those outcomes) were used to quantify the associations between 1 SD increase in the weighted sum of metabolite clusters and glucose, insulin, HOMA-IR, HbA1c, HDL-C, TG, BMI, and WC. To increase interpretability, adjusted means by quartiles of metabolite cluster scores were used to present the results for diabetes-related risk factors. Sensitivity analyses were conducted among all participants, adjusting for T2D status.

All p values were corrected for multiple testing using the Benjamini-Hochberg procedure with targeted FDR <0.05.³⁷ All statistical analyses were completed using SAS (V.9.4) or R (V.3.6.0) statistical software, and individuals with missing covariate data were excluded from these statistical models.

RESULTS

In the BPRHS, participants with T2D tended to be older, have higher WC, BMI, and TG, have lower HDL-C, consume less alcohol, and were more likely to be prescribed lipid-lowering or hypertension medications (table 1). In SOALS, participants with T2D tended to have higher WC and TG and were less likely to be female or prescribed hypertension medications.

In the global metabolite network derived from the BPRHS and SOALS participants without T2D, 69 metabolite clusters containing ≥ 2 metabolites were identified. After FDR correction, a total of nine metabolite cluster scores were significantly associated with prevalent T2D among the BPRHS participants (discovery) in the fully adjusted model (table 2). Metabolites included in each of the nine clusters are provided in online supplemental table S1, network representations are provided in figure 1, and the results from multivariable models 1 and 2 are presented in online supplemental table S2. Four clusters that primarily contained metabolites related to glucose transport (cluster 33; figure 1A), sphingolipids (cluster 2; figure 1B), acyl cholines (cluster 22; figure 1F), and cell membrane components (cluster 26; figure 1H) were associated with lower odds of T2D, and five clusters characterized by sugar metabolism (cluster 53; figure 1C), BCAA and aromatic amino acids (cluster 21; figure 1D), fatty acid biosynthesis (cluster 40; figure 1E), amino acid metabolism (cluster 14; figure 1G), and aromatic amino acid metabolism (cluster 25; figure 1I) were associated with higher odds of T2D. Associations between all 69 of the identified metabolite clusters and prevalent T2D among BPRHS participants are presented in online supplemental table S3. Associations between individual metabolites and prevalent T2D in the statistically significant clusters are presented in online supplemental figure S1.

Six of the nine T2D-associated clusters from the BPRHS were replicated in fully adjusted models among SOALS participants (FDR <0.05: glucose transport, sphingolipids, sugar metabolism, BCAA and aromatic amino acids, fatty acid biosynthesis, and acyl cholines) (table 2). In a meta-analysis of BPRHS and SOALS results, three of

the replicated clusters associated with lower odds of T2D: glucose transport (OR (95% CI) 0.21 (0.16 to 0.30); FDR <0.0001), sphingolipids (OR (95% CI) 0.40 (0.29 to 0.53); FDR <0.0001), and acyl cholines (OR (95% CI) 0.35 (0.22 to 0.56); FDR <0.0001), and three associated with higher odds of T2D: sugar metabolism (OR (95% CI) 2.28 (1.68 to 3.09); FDR <0.0001), BCAA and aromatic amino acids (OR (95% CI) 2.22 (1.60 to 3.08); FDR <0.0001), and fatty acid biosynthesis (OR (95% CI) 1.54 (1.29 to 1.85); FDR <0.0001). The results were similar in sensitivity analyses using an unweighted metabolite cluster scoring scheme, although attenuated effect sizes were observed (online supplemental table S4). In a sensitivity analysis including all clusters significantly associated with T2D in BPRHS participants in one model, the glucose transport, sugar metabolism, BCAA and aromatic amino acids, and fatty acid biosynthesis clusters, remained significant (online supplemental table S5).

The six T2D-associated clusters that were replicated in SOALS were all significantly associated with various diabetes-related risk factors among participants without T2D in the BPRHS and/or SOALS cohorts (figure 2, online supplemental tables S6 and S7). The larger sample size resulted in more statistically significant findings in the SOALS cohort, but trends were similar in the BPRHS. SOALS participants in the highest versus lowest quartile of the glucose transport cluster had significantly higher TG (SOALS FDR <0.0001) and HDL-C (SOALS FDR=0.02). SOALS participants in the highest versus lowest quartile of the sphingolipids cluster had significantly higher HDL-C concentrations (SOALS FDR <0.0001). Participants in the highest versus lowest quartile for the acyl cholines cluster had significantly higher TG (SOALS FDR <0.0001) and HDL-C (BPRHS FDR=0.02; SOALS FDR <0.0001) concentrations, along with lower WC (SOALS FDR=0.05), glucose (SOALS FDR <0.0001), insulin (SOALS FDR <0.0001), and HOMA-IR (SOALS FDR <0.0001). SOALS participants in the highest versus lowest quartile of the sugar metabolism cluster had higher HbA1c (SOALS FDR=0.05) and TG (SOALS FDR <0.0001). Participants in the highest versus lowest quartile of the BCAA and aromatic amino acid cluster had higher TG concentrations in both cohorts (BPRHS FDR=0.02; SOALS FDR <0.0001), and higher glucose (SOALS FDR=0.03), insulin (SOALS FDR <0.0001), HOMA-IR (SOALS FDR <0.0001), and HbA1c (SOALS FDR=0.001) in the SOALS cohort. Participants in the highest versus lowest quartile of the fatty acid biosynthesis cluster had higher HDL-C concentrations (BPRHS FDR=0.002; SOALS FDR=0.01) in both cohorts. When examining the associations between metabolite clusters and diabetes-related risk factors including participants with T2D and adjusting for diabetes status, the associations were generally similar, with many trends reaching statistical significance in the BPRHS due to an increase in sample size (online supplemental tables S8 and S9).

Although the three other T2D-associated clusters in BPRHS (amino acid metabolism (cluster 14), cell

Table 1 Baseline characteristics of study participants in BPRHS and SOALS stratified by T2D status*

	Discovery: BPRHS			Replication: SOALS		
	Without T2D	With T2D	P value	Without T2D	With T2D	P value
	n=322	n=357		n=934	n=77	
Age, years	55.6 (7.2)	58.4 (7.4)	<0.001	50.6 (6.8)	51.8 (6.4)	0.12
Sex, % female	74	73	0.68	74	60	0.009
Waist circumference, cm	97.7 (13.9)	106 (14.9)	<0.001	106 (14.3)	109 (10.6)	0.19
Body mass index, kg/m ²	30.7 (6.3)	33.6 (6.8)	<0.001	33.3 (6.3)	33.3 (5.4)	0.72
HDL-C (mg/dL)	47.3 (14.1)	43.1 (10.5)	<0.001	48.3 (12.6)	46.2 (17.1)	0.33
Triglyceride (mg/dL)†	131 (85)	148 (103)	0.002	129 (79)	164 (107)	<0.001
Glucose (mg/dL)	95.7 (9.5)	140 (58.3)	<0.001	92.4 (8.9)	138 (50.2)	<0.001
Insulin (μIU/mL)†	11.4 (9.5)	15.9 (11.9)	<0.001	8.9 (6.7)	12.5 (8.6)	<0.001
HbA1c (%)	5.8 (0.4)	8.0 (1.8)	<0.001	5.7 (0.3)	7.0 (1.5)	<0.001
HOMA-IR	3.25 (2.20)	8.11 (12.1)	<0.001	2.45 (1.68)	4.55 (2.27)	<0.001
Income, %, <\$20 000/year	68	73	0.50	46	43	0.70
Perceived Stress Score	23.6 (9.6)	23.3 (9.7)	0.54	12.7 (6.9)	11.6 (6.3)	0.40
Acculturation, %	18.4 (6.3)	17.9 (6.7)	0.40	–	–	–
Physical activity score	32.1 (4.9)	30.9 (4.1)	0.05	22.0 (40.2)	21.1 (29.3)	0.40
Alcohol consumption, %						
None in the past year	53	66	0.002	55	52	0.66
Moderate	38	29		41	45	
Heavy	7	4		2	1	
Education level, %						
≤8th grade	47	51	0.55	11	16	0.07
9th–12th grade or equivalent	38	36		43	53	
College/some graduate school	15	13		46	31	
Smoking status, %						
Never	46	47	0.31	64	62	0.90
Former	29	32		18	19	
Current	25	20		18	17	
American Heart Association Diet Score	23.1 (9.6)	24.0 (9.6)	0.57	–	–	–
Mouthwash use	–	–	–	53	60	0.32
Family history of T2D	–	–	–	52	64	0.07
Lipid-lowering medication, %	25	59	<0.001	13	16	0.60
Hypertension medication, %	40	70	<0.001	12	0	<0.001

*Values are mean (SE) unless otherwise noted. P values are age-adjusted and sex-adjusted.

†Geometric mean and IQR.

BPRHS, Boston Puerto Rican Health Study; HbA1c, glycosylated hemoglobin; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, Homeostatic Model Assessment for Insulin Resistance; SOALS, San Juan Overweight Adult Longitudinal Study; T2D, type 2 diabetes.

membrane components (cluster 26), and aromatic amino acid metabolism (cluster 25)) were not significantly associated with prevalent T2D in SOALS, a congruent direction of association was observed for two of the clusters (table 2). Additionally, participants in the highest versus lowest quartile of the amino acid metabolism cluster had significantly higher insulin concentrations in both cohorts (BPRHS FDR=0.003; SOALS FDR=0.0001) and HOMA-IR in the SOALS cohort (SOALS FDR=0.0003). Participants in the highest versus lowest quartile of the cell membrane components cluster in both cohorts

had significantly higher TG concentrations (BPRHS FDR=0.002; SOALS FDR <0.0001). No associations between the aromatic amino acid metabolism cluster score and diabetes-related risk factors were observed in either cohort (all FDR >0.05).

DISCUSSION

We identified six metabolite clusters within a global metabolite network that were consistently associated with prevalent T2D in two independent cohorts of individuals

Table 2 OR (95% CI) for prevalent T2D per 1 SD difference in weighted sum of metabolite concentrations in top metabolite clusters among BPRHS (n=679) and SOALS (n=1011) participants*

Cluster number	Cluster name	BPRHS			SOALS			Meta-analysis		
		n _{metab} †	OR (95% CI)	FDR	OR (95% CI)	FDR	OR (95% CI)	P value	I ² (%)	
33	Glucose transport	6	0.22 (0.14 to 0.32)	<0.0001	0.21 (0.13 to 0.35)	<0.0001	0.21 (0.16 to 0.30)	<0.0001	0	
2	Sphingolipids	18	0.44 (0.30 to 0.64)	0.002	0.33 (0.20 to 0.54)	0.0001	0.40 (0.29 to 0.53)	<0.0001	0	
53	Sugar metabolism	5	2.41 (1.65 to 3.56)	0.0004	2.07 (1.26 to 3.44)	0.04	2.28 (1.68 to 3.09)	<0.0001	0	
21	BCAA and aromatic amino acids	13	1.96 (1.40 to 2.77)	0.007	2.79 (1.69 to 4.66)	0.0006	2.22 (1.60 to 3.08)	<0.0001	21	
40	Fatty acid biosynthesis	6	1.54 (1.22 to 1.94)	0.02	1.56 (1.15 to 2.11)	0.04	1.54 (1.29 to 1.85)	<0.0001	0	
22	Acyl cholines	9	0.43 (0.31 to 0.60)	<0.0001	0.27 (0.17 to 0.42)	<0.0001	0.35 (0.22 to 0.56)	<0.0001	62	
14	Amino acid metabolism	15	2.21 (1.45 to 3.43)	0.02	1.31 (0.72 to 2.38)	1.00	1.78 (1.07 to 2.95)	0.03	49	
26	Cell membrane components	8	0.46 (0.30 to 0.70)	0.02	0.85 (0.50 to 1.44)	1.00	0.61 (0.34 to 1.12)	0.11	69	
25	Aromatic amino acid metabolism	11	2.01 (1.43 to 2.84)	0.004	0.93 (0.58 to 1.48)	1.00	1.39 (0.65 to 2.96)	0.39	85	

Benjamini-Hochberg FDR for 69 (BPRHS) or 9 (SOALS and meta-analysis) metabolite clusters.

*Statistically significant cross-sectional associations between metabolite clusters and prevalent T2D in logistic regression models among BPRHS participants are presented here. Associations between additional metabolite clusters and prevalent T2D in the BPRHS are presented in online supplemental table S2. Meta-analysis represents inverse-variance weighted random-effect meta-analysis estimates. Model adjusts for age, sex, smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), acculturation (%; BPRHS only), Perceived Stress Score, AHA-DS (quartiles; BPRHS only), mouthwash use (yes/no; SOALS only), family history of T2D (yes/no; SOALS only), waist circumference (cm), and body mass index (kg/m²).

†Number of metabolites included in each cluster.

‡AHA-DS, American Heart Association Diet Score; BCAA, branched-chain amino acids; BPRHS, Boston Puerto Rican Health Study; FDR, false discovery rate; SOALS, San Juan Overweight Adult Longitudinal Study; T2D, type 2 diabetes.

of PR descent. To our knowledge, this is the first study to comprehensively compare over 600 plasma metabolites levels among individuals of PR descent with and without T2D. Participants were less likely to have T2D if they had higher concentrations of metabolites in three clusters characterized by sphingolipids, acyl cholines, or metabolites related to glucose transport. In contrast, participants with higher concentrations of metabolites within the three clusters characterized by sugar metabolism, BCAA and aromatic amino acid metabolism, and fatty acid biosynthesis had a higher likelihood of T2D. Additional metabolite clusters, including metabolites related to amino acid metabolism and components of cell membranes, also displayed suggestive associations with prevalent T2D. Several associations between metabolite clusters and known diabetes-related risk factors were also observed. These findings replicate known metabolites and identify novel metabolites associated with prevalent T2D, which may reflect biological changes occurring among individuals of PR descent with T2D.

This is the first population-based study to identify a cluster of acyl choline metabolites that were associated with lower odds of prevalent T2D. A previous small study (n=107) observed that arachidonylcholine (one component of the acyl choline cluster) levels were significantly lower among obese patients with T2D compared with obese insulin-sensitive patients.¹⁶ Consistently, we identified additional acyl cholines associated with lower odds of T2D, along with novel associations with higher HDL-C and lower WC, glucose, insulin, and HOMA-IR. The acyl choline cluster was also surprisingly associated with higher TG concentrations, suggesting that participants with high concentrations of metabolites in the acyl choline cluster represent a unique group of individuals that have higher TG concentrations but are less likely to have T2D. The glucose transport cluster that was associated with lower odds of T2D contains a well-established biomarker of short-term glycemic control, 1,5-anhydroglucitol (1,5-AG), which has been negatively correlated with plasma glucose concentrations.³⁸ However, we identified additional metabolites in the same network cluster as 1,5-AG that also associated with lower odds of T2D. Several of these metabolites have been potentially linked to glucose transport in animal and cell studies, such as N-acetyltaurine³⁹ and phosphatidylcholine,⁴⁰ but the biological connections between the metabolites in this cluster are unclear. The sphingomyelin cluster was associated with a lower prevalence of T2D, similar to previous studies.^{41–43} This includes one large study in a mainly non-Hispanic white population (n=3082)⁴² and a moderate-sized (n=1035) study in an ethnically diverse population.⁴³ We have expanded these findings to PR adults and identified the associations between sphingolipids and higher HDL-C concentrations. Previous studies have linked higher HDL-C concentrations to better pancreatic β cell function,^{44 45} suggesting that further exploration of how metabolites in this sphingolipid cluster influence pancreatic β cell function may be fruitful.

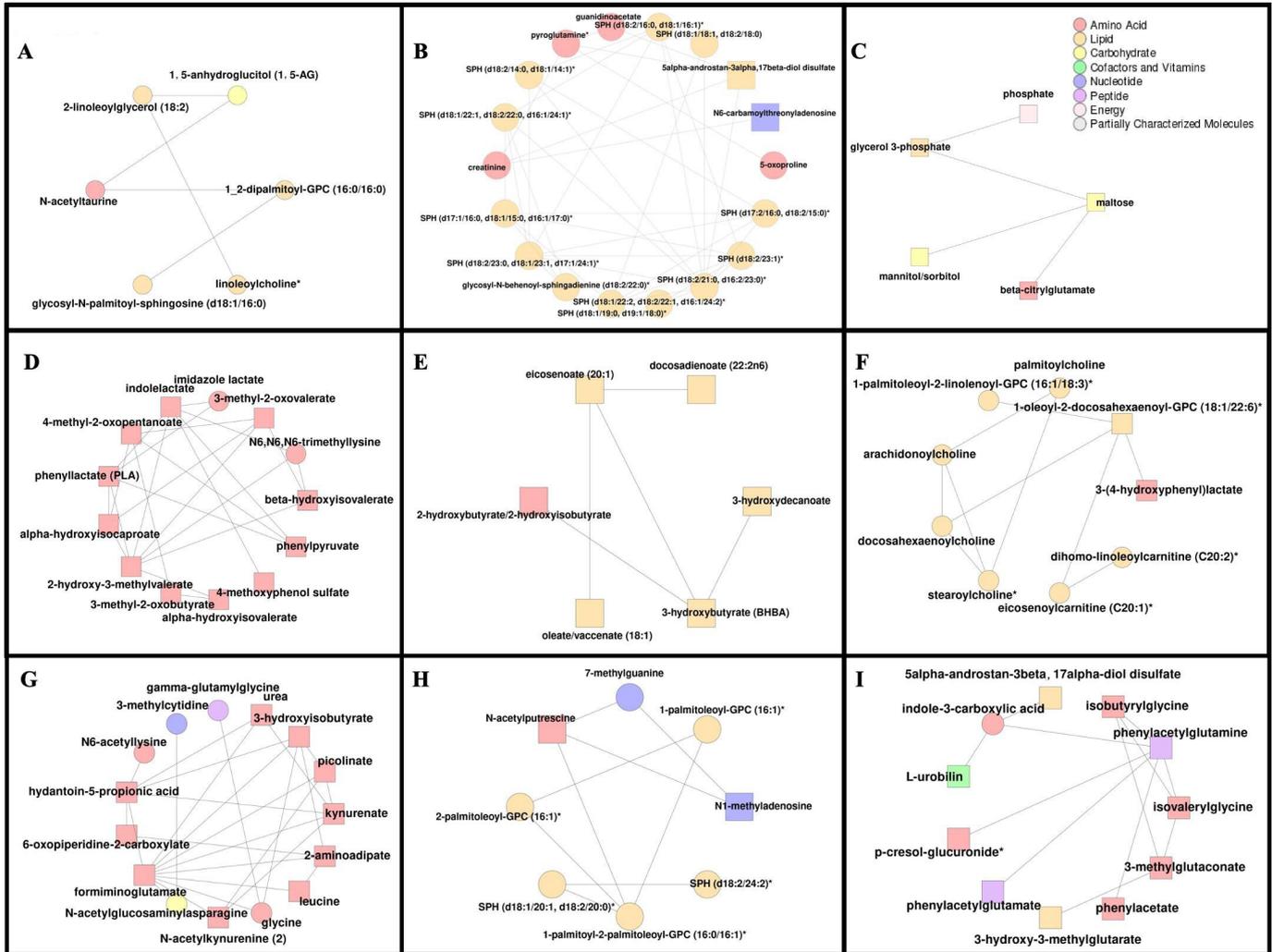


Figure 1 Metabolite clusters significantly associated with prevalent type 2 diabetes in Boston Puerto Rican Health Study participants. Nodes represent metabolites and edges represent partial correlation with $p < 0.05$ for metabolite pairs. Direction of association between metabolites and prevalent type 2 diabetes is indicated by node shape, where circles represent inverse and squares represent positive associations. Nodes are colored according to metabolite class. The following names have been given to each cluster based on the metabolites included: (A) cluster 33: glucose transport; (B) cluster 2: sphingolipids; (C) cluster 53: sugar metabolism; (D) cluster 21: BCAA and aromatic amino acids; (E) cluster 40: fatty acid biosynthesis; (F) cluster 22: acyl choline; (G) cluster 14: amino acid metabolism; (H) cluster 26: cell membrane components; and (I) cluster 25: aromatic amino acid metabolism. *Indicates a statistically significant association ($FDR < 0.05$) (blue=BPRHS; orange=SOALS). BCAA, branched-chain amino acids; GPC, glycerophosphocholine; SPH, sphingomyelin.

BCAA (isoleucine, leucine, and valine) and aromatic amino acid (tyrosine and phenylalanine) metabolites have been consistently associated with higher T2D risk and diabetes-related risk factors in prospective cohort studies among mainly non-Hispanic white populations, as well as cross-sectional and case-control studies in non-Hispanic white, Asian, and African American populations.⁸ Only one small prospective study included 50 participants of Hispanic ethnicity,⁴⁶ and with their limited statistical power, they did not observe a higher risk of T2D among Hispanic participants with elevated BCAA and aromatic amino acid metabolites. In the current study, we identified a cluster of BCAA and aromatic amino acid metabolites that associated with higher odds of prevalent T2D, higher concentrations of glucose, insulin, and TG, as well as higher HOMA-IR and HbA1c among PR adults.

We also replicated several associations of metabolites related to sugar metabolism (driven largely by mannitol and maltose) and fatty acid biosynthesis (including 2-hydroxybutyrate, 3-hydroxybutyrate, and several fatty acids) with high T2D risk.^{8 42 47 48} The clusters we identified include additional related metabolites that may point to mechanisms for how these metabolites could help characterize T2D phenotypes.

Additional clusters of metabolites that were significantly associated with prevalent T2D among BPRHS participants were not replicated among SOALS participants (amino acid metabolism, cell membrane components, and aromatic amino acid metabolites). However, some notable associations between these clusters and diabetes-related risk factors were observed. For example, a cluster of cell membrane components, including

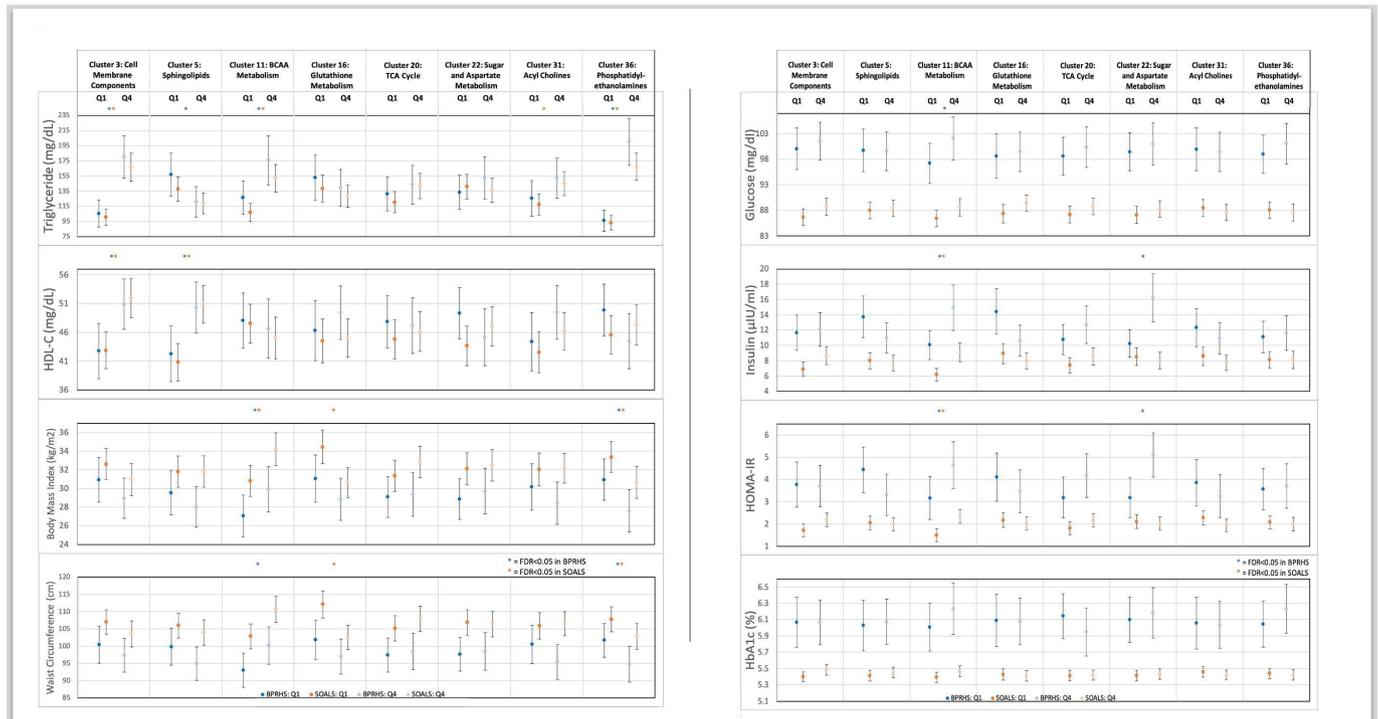


Figure 2 Adjusted means by quartile of weighted T2D-associated metabolite cluster scores among BPRHS (n=322) and SOALS (n=934) participants without T2D. Cross-sectional associations adjusted for age, sex, diet quality (BPRHS only), smoking status, education, physical activity, alcohol intake, lipid-lowering medication use, hypertension medication use, income, acculturation (BPRHS only), Perceived Stress Score, mouthwash use (SOALS only), family history of T2D (SOALS only), WC (except WC and BMI outcomes), and BMI (except WC and BMI outcomes). P values are corrected for multiple exposures and outcomes using the Benjamini-Hochberg FDR. BCAA, branched-chain amino acid; BMI, body mass index; BPRHS, Boston Puerto Rican Health Study; FDR, false discovery rate; HbA1c, glycosylated hemoglobin; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, Homeostatic Model Assessment of Insulin Resistance; SOALS, San Juan Overweight Adult Longitudinal Study; T2D, type 2 diabetes; WC, waist circumference.

phosphatidylcholines, lysophospholipids, and sphingolipids, displayed a significant association with higher TG concentrations in both cohorts. Although the cell membrane components cluster was associated with lower odds of T2D among BPRHS participants, it was not associated with T2D prevalence among SOALS participants. This inconsistent association is likely due to the difference in the direction of associations for the individual metabolites within the cell membrane components cluster in the two cohorts. Sphingomyelin (d18:2/24:2) was consistently associated with a lower prevalence of T2D in both cohorts, but the associations for other metabolites were inconsistent, which could be due to chance or differences in study methodology or population characteristics. Cell membrane structure likely influences T2D risk,^{8 49 50} but further research on how complex changes in cell membrane structure and fluidity influence diabetes-related metabolic changes is warranted. The amino acid metabolism cluster also contained several individual metabolites that were strongly associated with the prevalence of T2D in consistent directions. This included glycine and leucine, which have been consistently linked to lower and higher T2D risk in previous studies, respectively.⁸

Our study has several strengths and limitations that influence the interpretation of the results. Network analysis allows us to make inferences about a large number of metabolites from data sets of moderate sample size. The cross-sectional design of this study provides no information on the directionality of associations. We, therefore, derived our metabolite network and examined the associations between metabolite clusters and diabetes-related risk factors among individuals without T2D. We also strengthened our findings through replication in an independent population. Although lack of consistency in some observed associations could be due to differences in methodology between the BPRHS and SOALS cohorts, our ability to identify consistent differences in metabolites by T2D status despite these differences lends strength to our findings. During the BPRHS baseline visit, OGTTs were not conducted, and a portion of participants may have been misclassified as not having T2D when their OGTT may have indicated the presence of T2D. This could lead to attenuation of the association between metabolite clusters and T2D in the BPRHS. Residual confounding is always possible in observational studies, and we have minimized this by controlling for many potential confounders and covariates. Given these

weaknesses, findings should be interpreted as largely hypothesis-generating and warrant further examination using prospective study designs.

In summary, we used a data-driven, network-based method to identify several metabolic pathways that are perturbed among individuals of PR descent with T2D, providing deeper insight into pathways that should be further explored to understand the pathogenesis of T2D in this high-risk group. A novel cluster of acyl cholines was associated with a lower odds of T2D and should be investigated in other populations. We replicated several previously observed associations between metabolites and T2D and used network modules to reveal novel correlated metabolites that were also associated with the prevalence of T2D. Additional associations between the identified metabolite clusters and traditional diabetes-related risk factors provide insight into distinct T2D phenotypes that may be characterized by integrative alterations in these metabolites and traditional risk factors. More prospective studies are needed to assess which lifestyle, genetic, or environmental factors may influence these metabolites and whether alterations in these pathways start occurring before or whether they are a consequence of T2D.

Author affiliations

¹Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

²Nutrition, Harvard T H Chan School of Public Health, Boston, Massachusetts, USA

³Biostatistics, Harvard T H Chan School of Public Health, Boston, Massachusetts, USA

⁴Epidemiology, Harvard T H Chan School of Public Health, Boston, Massachusetts, USA

⁵Department of Biostatistics and Epidemiology, Graduate School of Public Health, University of Puerto Rico Medical Sciences Campus, San Juan, Puerto Rico

⁶Center for Clinical Research and Health Promotion, University of Puerto Rico Medical Sciences Campus, San Juan, Puerto Rico

⁷Molecular Metabolism, Harvard T H Chan School of Public Health, Boston, Massachusetts, USA

⁸Metabolomics Platform, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

⁹Center for Oral/Head and Neck Oncology Research, School of Dentistry, University of California Los Angeles, Los Angeles, California, USA

¹⁰Agricultural Research Service, Jean Mayer US Department of Agriculture Human Nutrition Research Center on Aging at Tufts University, Boston, Massachusetts, USA

¹¹IMDEA-Food Institute, CEI UAM+CSIC, Madrid, Spain

¹²Nutrition and Genomics, Jean Mayer US Department of Agriculture Human Nutrition Research Center on Aging at Tufts University, Boston, Massachusetts, USA

¹³Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

¹⁴Department of Biomedical and Nutritional Sciences and Center for Population Health, University of Massachusetts Lowell, Lowell, Massachusetts, USA

Acknowledgements The authors thank all BPRHS and SOALS participants and staff for their contribution to these studies.

Contributors DEH, LL, and DDW performed the statistical analyses. DEH and SNB drafted the manuscript. DEH, LL, DDW, RSK, CW, CMP, MM, C-HL, CBC, DTWW, LDP, C-QL, JMO, JEM, FBH, MJS, KLT, KJJ, and SNB contributed to interpretation of the data and revised the article critically for important intellectual content. All authors approved the final version of the manuscript. DEH and SNB are guarantors of the work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and accuracy of the data analysis.

Funding This work is supported by the National Institutes of Health (NIH): 2T32CA009001 (DEH), 1K01DK107804-01A1 (SNB), and 1R01DK120560-01 (DEH, LL, C-HL, DTWW, FBH, MJS, KLT, KJJ, and SNB).

Competing interests DTWW is consultant to Mars Wrigley and Colgate-Palmolive and has equity in Liquid Diagnostics.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Information on requesting data from the BPRHS (<https://www.uml.edu/Research/UML-CPH/Research/bprhs/>) and SOALS (<http://soals.rcm.upr.edu/>) can be found on their respective websites.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Danielle E Haslam <http://orcid.org/0000-0003-0144-3287>

Dong D Wang <http://orcid.org/0000-0002-0897-3048>

Clemens Wittenbecher <http://orcid.org/0000-0001-7792-877X>

Chih-Hao Lee <http://orcid.org/0000-0002-6090-0786>

Clary B Clish <http://orcid.org/0000-0001-8259-9245>

Chao-Qiang Lai <http://orcid.org/0000-0003-1107-8375>

Shilpa N Bhupathiraju <http://orcid.org/0000-0001-9910-6777>

REFERENCES

- Smith CAS, Barnett E. Diabetes-related mortality among Mexican Americans, Puerto Ricans, and Cuban Americans in the United States. *Rev Panam Salud Publica* 2005;18:381–7.
- Benjamin EJ, Virani SS, Callaway CW, *et al*. Heart disease and stroke statistics-2018 update: a report from the American heart association. *Circulation* 2018;137:e67–492.
- Colón-Ramos U, Rodríguez-Ayuso I, Gebrekristos HT, *et al*. Transnational mortality comparisons between archipelago and mainland Puerto Ricans. *J Immigrant Minority Health* 2017;19:1009–17.
- Laakso M. Biomarkers for type 2 diabetes. *Molecular Metabolism* 2019;27:S139–46.
- Wang TJ, Ngo D, Psychogios N, *et al*. 2-Aminoadipic acid is a biomarker for diabetes risk. *J Clin Invest* 2013;123:4309–17.
- Cheng S, Rhee EP, Larson MG, *et al*. Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation* 2012;125:2222–31.
- Rebholz CM, Yu B, Zheng Z, *et al*. Serum metabolomic profile of incident diabetes. *Diabetologia* 2018;61:1046–54.
- Guasch-Ferré M, Hruby A, Toledo E, *et al*. Metabolomics in prediabetes and diabetes: a systematic review and meta-analysis. *Diabetes Care* 2016;39:833–46.
- Mihalik SJ, Goodpaster BH, Kelley DE, *et al*. Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipototoxicity. *Obesity* 2010;18:1695–700.
- Newgard CB, An J, Bain JR, *et al*. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab* 2009;9:311–26.
- McCormack SE, Shaham O, McCarthy MA, *et al*. Circulating branched-chain amino acid concentrations are associated with obesity and future insulin resistance in children and adolescents. *Pediatr Obes* 2013;8:52–61.
- Adams SH, Hoppel CL, Lok KH, *et al*. Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid β -oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. *J Nutr* 2009;139:1073–81.

- 13 Qi Q, Stilp AM, Sofer T, et al. Genetics of type 2 diabetes in U.S. hispanic/Latino individuals: results from the hispanic community health study/study of Latinos (HCHS/SOL). *Diabetes* 2017;66:1419–25.
- 14 Aguayo-Mazzucato C, Diaque P, Hernandez S, et al. Understanding the growing epidemic of type 2 diabetes in the hispanic population living in the United States. *Diabetes Metab Res Rev* 2019;35:e3097.
- 15 Zhao J, Zhu Y, Hyun N, et al. Novel metabolic markers for the risk of diabetes development in American Indians. *Diabetes Care* 2015;38:220–7.
- 16 Al-Sulaiti H, Diboun I, Agha MV, et al. Metabolic signature of obesity-associated insulin resistance and type 2 diabetes. *J Transl Med* 2019;17 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6805293/>
- 17 Merino J, Leong A, Liu C-T, et al. Metabolomics insights into early type 2 diabetes pathogenesis and detection in individuals with normal fasting glucose. *Diabetologia* 2018;61:1315–24.
- 18 Ahlqvist E, Storm P, Käräjämäki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* 2018;6:361–9.
- 19 Udler MS, Kim J, von Grotthuss M, von GM, et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: a soft clustering analysis. *PLoS Med* 2018;15:e1002654.
- 20 Li L, Cheng W-Y, Glicksberg BS, et al. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. *Sci Transl Med* 2015;7:311ra174–311.
- 21 Tucker KL, Mattei J, Noel SE, et al. The Boston puerto Rican health study, a longitudinal cohort study on health disparities in puerto Rican adults: challenges and opportunities. *BMC Public Health* 2010;10:107.
- 22 Andriankaja OM, Jiménez JJ, Muñoz-Torres FJ, et al. Lipid-lowering agents use and systemic and oral inflammation in overweight or obese adult puerto ricans: the san juan overweight adults longitudinal study (SOALS). *J Clin Periodontol* 2015;42:1090–6.
- 23 Joshipura KJ, Muñoz-Torres FJ, Dye BA, et al. Longitudinal association between periodontitis and development of diabetes. *Diabetes Res Clin Pract* 2018;141:284–93.
- 24 Pérez CM, Muñoz F, Andriankaja OM, et al. Cross-sectional associations of impaired glucose metabolism measures with bleeding on probing and periodontitis. *J Clin Periodontol* 2017;44:142–9.
- 25 Ioannidis JPA, Khoury MJ. Improving validation practices in "omics" research. *Science* 2011;334:1230–2.
- 26 Evans AM, DeHaven CD, Barrett T, et al. 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;81:6656–67.
- 27 American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes—2021. *Diabetes Care* 2021;44:S15–33.
- 28 Tucker KL, Bianchi LA, Maras J, et al. Adaptation of a food frequency questionnaire to assess diets of puerto rican and non-hispanic adults. *Am J Epidemiol* 1998;148:507–18.
- 29 Bhupathiraju SN, Lichtenstein AH, Dawson-Hughes B, et al. Adherence index based on the AHA 2006 diet and lifestyle recommendations is associated with select cardiovascular disease risk factors in older Puerto ricans. *J Nutr* 2011;141:460–9.
- 30 Paffenbarger RS, Wing AL, Hyde RT. Physical activity as an index of heart attack risk in college alumni. *Am J Epidemiol* 1995;142:889–903.
- 31 Paffenbarger RS, Hyde RT, Wing AL, et al. The association of changes in physical-activity level and other lifestyle characteristics with mortality among men. *N Engl J Med* 1993;328:538–45.
- 32 Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983;24:385–96.
- 33 Marin G, Gamba RJ. A new measurement of acculturation for hispanics: the bidimensional acculturation scale for hispanics (BAS). *Hisp J Behav Sci* 2016.
- 34 Csardi G, Nepusz T. The igraph software package for complex network research [Internet], 2006. Available: <https://igraph.org>
- 35 Clauset A, Newman MEJ, Moore C. Finding community structure in very large networks. *Phys Rev E* 2004;70:066111.
- 36 Ravasz E, Somera AL, Mongru DA, et al. Hierarchical organization of modularity in metabolic networks. *Science* 2002;297:1551–5.
- 37 Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *J R Stat Soc Series B* 2001;57:289–300.
- 38 Dungan KM. 1,5-anhydroglucitol (GlycoMark™) as a marker of short-term glycemic control and glycemic excursions. *Expert Rev Mol Diagn* 2008;8:9–19.
- 39 Miyazaki T, Nakamura Y, Ebina K, et al. Increased N-acetyltaurine in the skeletal muscle after endurance exercise in rat. *Adv Exp Med Biol* 2017;975 Pt 1:403–11.
- 40 Rauch C, Loughna P. C2C12 skeletal muscle cells exposure to phosphatidylcholine triggers IGF-1 Like-Responses. *Cell Physiol Biochem* 2005;15:211–24.
- 41 Xu F, Tavintharan S, Sum CF, et al. Metabolic signature shift in type 2 diabetes mellitus revealed by mass spectrometry-based metabolomics. *J Clin Endocrinol Metab* 2013;98:E1060–5.
- 42 Floegel A, Stefan N, Yu Z, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* 2013;62:639–48.
- 43 Khan SR, Manialawy Y, Obersterescu A, et al. Diminished sphingolipid metabolism, a hallmark of future type 2 diabetes pathogenesis, is linked to pancreatic β cell dysfunction. *iScience* 2020;23:101566.
- 44 Brunham LR. Hdl and pancreatic β cells: a SMO-king gun? *J Lipid Res* 2020;61:468–9.
- 45 Drew BG, Duffy SJ, Formosa MF, et al. High-Density lipoprotein modulates glucose metabolism in patients with type 2 diabetes mellitus. *Circulation* 2009;119:2103–11.
- 46 Palmer ND, Stevens RD, Antinozzi PA, et al. Metabolomic profile associated with insulin resistance and conversion to diabetes in the insulin resistance atherosclerosis study. *J Clin Endocrinol Metab* 2015;100:E463–8.
- 47 Padberg I, Peter E, González-Maldonado S, et al. A new metabolomic signature in type-2 diabetes mellitus and its pathophysiology. *PLoS One* 2014;9:e85082 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3894948/>
- 48 Drogan D, Dunn WB, Lin W, et al. Untargeted metabolic profiling identifies altered serum metabolites of type 2 diabetes mellitus in a prospective, nested case control study. *Clin Chem* 2015;61:487–97.
- 49 Russo SB, Ross JS, Cowart LA. Sphingolipids in obesity, type 2 diabetes, and metabolic disease. *Handb Exp Pharmacol* 2013;216:373–401.
- 50 Haus JM, Kashyap SR, Kasumov T, et al. Plasma ceramides are elevated in obese subjects with type 2 diabetes and correlate with the severity of insulin resistance. *Diabetes* 2009;58:337–43.

Electronic Supplemental Material

Supplemental Table S1. Cluster membership and weights for metabolite clusters significantly associated with prevalent type 2 diabetes in Boston Puerto Rican Health Study participants.

Cluster	Biochemical	Super Pathway	Sub-pathway	Weight
Glucose Transport (Cluster 33)	1,2-dipalmitoyl-GPC (16:0/16:0)	Lipid	Phosphatidylcholine (PC)	0.18
	1,5-anhydroglucitol (1,5-AG)	Carbohydrate	Glycolysis, Gluconeogenesis, and Pyruvate Metabolism	0.18
	2-linoleoylglycerol (18:2)	Lipid	Monoacylglycerol	0.18
	N-acetyltaurine	Amino Acid	Methionine, Cysteine, SAM and Taurine Metabolism	0.18
	glycosyl-N-palmitoyl-sphingosine (d18:1/16:0)	Lipid	Ceramides	0.15
	linoleoylcholine*	Lipid	Fatty Acid Metabolism (Acyl Choline)	0.15
Sphingolipids (Cluster 2)	sphingomyelin (d18:2/21:0, d16:2/23:0)*	Lipid	Sphingolipid Metabolism	0.10
	sphingomyelin (d18:2/23:0, d18:1/23:1, d17:1/24:1)*	Lipid	Sphingolipid Metabolism	0.08
	sphingomyelin (d18:2/23:1)*	Lipid	Sphingolipid Metabolism	0.07
	sphingomyelin (d18:2/16:0, d18:1/16:1)*	Lipid	Sphingolipid Metabolism	0.07
	sphingomyelin (d18:2/14:0, d18:1/14:1)*	Lipid	Sphingolipid Metabolism	0.07
	sphingomyelin (d18:1/22:1, d18:2/22:0, d16:1/24:1)*	Lipid	Sphingolipid Metabolism	0.06
	sphingomyelin (d17:2/16:0_ d18:2/15:0)*	Lipid	Sphingolipid Metabolism	0.06
	sphingomyelin (d18:1/19:0_ d19:1/18:0)*	Lipid	Sphingolipid Metabolism	0.06
	sphingomyelin (d18:1/22:2_ d18:2/22:1, d16:1/24:2)*	Lipid	Sphingolipid Metabolism	0.06
	sphingomyelin (d18:1/18:1, d18:2/18:0)	Lipid	Sphingolipid Metabolism	0.06
	5alpha-androstan-3alpha,17beta-diol disulfate	Lipid	Androgenic Steroids	0.06
	sphingomyelin (d17:1/16:0, d18:1/15:0, d16:1/17:0)*	Lipid	Sphingolipid Metabolism	0.05
	glycosyl-N-behenoyl-sphingadienine (d18:2/22:0)*	Lipid	Ceramides	0.05
	pyroglutamine*	Amino Acid	Glutamate Metabolism	0.04
	creatinine	Amino Acid	Creatine Metabolism	0.04
	guanidinoacetate	Amino Acid	Creatine Metabolism	0.03
	5-oxoproline	Amino Acid	Glutathione Metabolism	0.02
N6-carbamoylthreonyladenosine	Nucleotide	Purine Metabolism, Adenine containing	0.02	
maltose	Carbohydrate	Glycogen Metabolism	0.27	
glycerol 3-phosphate	Lipid	Glycerolipid Metabolism	0.23	

Sugar Metabolism (Cluster 1)	mannitol/sorbitol	Carbohydrate	Fructose, Mannose and Galactose Metabolism	0.18
	beta-citrylglutamate phosphate	Amino Acid	Glutamate Metabolism	0.18
		Energy	Oxidative Phosphorylation	0.14
BCAA and Aromatic Amino Acids (Cluster 21)	2-hydroxy-3-methylvalerate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.12
	indolelactate	Amino Acid	Tryptophan Metabolism	0.10
	phenyllactate (PLA)	Amino Acid	Phenylalanine Metabolism	0.10
	alpha-hydroxyisocaproate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.09
	N6,N6,N6-trimethyllysine	Amino Acid	Lysine Metabolism	0.08
	3-methyl-2-oxovalerate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.08
	phenylpyruvate	Amino Acid	Phenylalanine Metabolism	0.08
	imidazole lactate	Amino Acid	Histidine Metabolism	0.07
	beta-hydroxyisovalerate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.07
	4-methoxyphenol sulfate	Amino Acid	Tyrosine Metabolism	0.06
	4-methyl-2-oxopentanoate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.06
	alpha-hydroxyisovalerate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.06
	3-methyl-2-oxobutyrate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.04
Fatty Acid Biosynthesis (Cluster 40)	eicosenoate (20:1)	Lipid	Long Chain Fatty Acid	0.23
	3-hydroxybutyrate (BHBA)	Lipid	Ketone Bodies	0.23
	3-hydroxydecanoate	Lipid	Fatty Acid, Monohydroxy	0.14
	docosadienoate (22:2n6)	Lipid	Polyunsaturated Fatty Acid (n3 and n6)	0.14
	2-hydroxybutyrate/2-hydroxyisobutyrate	Amino Acid	Glutathione Metabolism	0.14
	oleate/vaccenate (18:1)	Lipid	Long Chain Fatty Acid	0.14
Acyl Cholines (Cluster 22)	1-oleoyl-2-docosahexaenoyl-GPC (18:1/22:6)*	Lipid	Phosphatidylcholine (PC)	0.15
	docosahexaenoylcholine	Lipid	Fatty Acid Metabolism (Acyl Choline)	0.14
	arachidonoylcholine	Lipid	Fatty Acid Metabolism (Acyl Choline)	0.12
	stearoylcholine*	Lipid	Fatty Acid Metabolism (Acyl Choline)	0.12
	eicosenoylcarnitine (C20:1)*	Lipid	Fatty Acid Metabolism(Acyl Carnitine)	0.11

	palmitoylcholine	Lipid	Fatty Acid Metabolism (Acyl Choline)	0.11
	3-(4-hydroxyphenyl)lactate	Amino Acid	Tyrosine Metabolism	0.10
	1-palmitoleoyl-2-linolenoyl-GPC (16:1/18:3)*	Lipid	Phosphatidylcholine (PC)	0.10
	dihomo-linoleoylcarnitine (C20:2)*	Lipid	Fatty Acid Metabolism(Acyl Carnitine)	0.06
	formiminoglutamate	Amino Acid	Histidine Metabolism	0.13
	3-hydroxyisobutyrate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.09
	hydantoin-5-propionic acid	Amino Acid	Histidine Metabolism	0.09
	kynurenate	Amino Acid	Tryptophan Metabolism	0.08
	urea	Amino Acid	Urea cycle; Arginine and Proline Metabolism	0.08
	picolinate	Amino Acid	Tryptophan Metabolism	0.07
Amino Acid Metabolism (Cluster 14)	N-acetylkynurenine (2)	Amino Acid	Tryptophan Metabolism	0.07
	leucine	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.07
	2-aminoadipate	Amino Acid	Lysine Metabolism	0.07
	glycine	Amino Acid	Glycine, Serine and Threonine Metabolism	0.06
	N-acetylglucosaminylasparagine	Carbohydrate	Aminosugar Metabolism	0.06
	6-oxopiperidine-2-carboxylate	Amino Acid	Lysine Metabolism	0.05
	N6-acetyllysine	Amino Acid	Lysine Metabolism	0.04
	gamma-glutamylglycine	Peptide	Gamma-glutamyl Amino Acid	0.02
	3-methylcytidine	Nucleotide	Pyrimidine Metabolism, Cytidine containing	0.02
	1-palmitoyl-2-palmitoleoyl-GPC (16:0/16:1)*	Lipid	Phosphatidylcholine (PC)	0.19
	N-acetylputrescine	Amino Acid	Polyamine Metabolism	0.16
Cell Membrane Components (Cluster 26)	1-palmitoleoyl-GPC (16:1)*	Lipid	Lysophospholipid	0.13
	2-palmitoleoyl-GPC (16:1)*	Lipid	Lysophospholipid	0.13
	sphingomyelin (d18:1/20:1_d18:2/20:0)*	Lipid	Sphingolipid Metabolism	0.12
	N1-methyladenosine	Nucleotide	Purine Metabolism, Adenine containing	0.11
	7-methylguanine	Nucleotide	Purine Metabolism, Guanine containing	0.11
	sphingomyelin (d18:2/24:2)*	Lipid	Sphingolipid Metabolism	0.07
	phenylacetylglutamine	Peptide	Acetylated Peptides	0.16
Aromatic Amino Acid	isovalerylglycine	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.12

Metabolism (Cluster 25)	isobutyrylglycine	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.12
	indole-3-carboxylic acid	Amino Acid	Tryptophan Metabolism	0.11
	p-cresol-glucuronide*	Amino Acid	Tyrosine Metabolism	0.09
	phenylacetylglutamate	Peptide	Acetylated Peptides	0.09
	phenylacetate	Amino Acid	Phenylalanine Metabolism	0.09
	3-methylglutaconate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.07
	5alpha-androstan-3beta,17alpha-diol disulfate	Lipid	Androgenic Steroids	0.05
	L-urobilin	Cofactors and Vitamins	Hemoglobin and Porphyrin Metabolism	0.05
	3-hydroxy-3-methylglutarate	Lipid	Mevalonate Metabolism	0.05

Abbreviations: BCAA, branched-chain amino acids; GPC, glycerophosphocoline; SAM, S-adenosylmethionine; SOALS, San Juan Overweight Adults Longitudinal Study.

Supplemental Table S2. Odds ratios (95% confidence intervals) for prevalent type 2 diabetes per 1-standard deviation difference in sum of metabolite concentrations in top metabolite clusters among Boston Puerto Rican Health Study (BPRHS; n=679) and San Juan Overweight Adult (SOALS; n=1,011)*

Cluster #	Cluster Name	Model 1 [†]	FDR	Model 2 [‡]	FDR	Model 3 [§]	FDR
BPRHS							
33	Glucose Transport	0.18 (0.13, 0.26)	1.63E-19	0.20 (0.13, 0.29)	7.51E-14	0.22 (0.14, 0.32)	8.12E-12
2	Sphingolipids	0.41 (0.29, 0.57)	5.87E-06	0.43 (0.29, 0.62)	0.0006	0.44 (0.30, 0.64)	0.002
53	Sugar Metabolism	2.74 (1.99, 3.82)	1.03E-07	2.47 (1.71, 3.60)	0.0001	2.41 (1.65, 3.56)	0.0004
21	BCAA and Aromatic Amino Acids	1.91 (1.43, 2.58)	0.001	2.00 (1.44, 2.79)	0.003	1.96 (1.40, 2.77)	0.007
40	Fatty Acid Biosynthesis	1.62 (1.33, 1.97)	8.12E-05	1.61 (1.29, 2.01)	0.002	1.54 (1.22, 1.94)	0.02
22	Acyl Cholines	0.42 (0.31, 0.55)	1.94E-07	0.39 (0.28, 0.54)	1.65E-06	0.43 (0.31, 0.60)	7.95E-05
14	Amino Acid Metabolism	2.74 (1.90, 4.00)	7.65E-06	2.45 (1.61, 3.77)	0.003	2.21 (1.45, 3.43)	0.02
26	Cell Membrane Components	0.45 (0.32, 0.64)	0.001	0.49 (0.33, 0.73)	0.04	0.46 (0.30, 0.70)	0.02
25	Aromatic Amino Acid Metabolism	2.03 (1.52, 2.74)	0.0002	1.76 (1.27, 2.45)	0.05	2.01 (1.43, 2.84)	0.004
SOALS							
33	Glucose Transport	0.25 (0.16, 0.39)	1.16E-08	0.21 (0.13, 0.34)	3.95E-09	0.21 (0.13, 0.35)	4.54E-09
2	Sphingolipids	0.34 (0.21, 0.54)	5.96E-05	0.33 (0.20, 0.55)	0.0001	0.33 (0.20, 0.54)	0.0001
53	Sugar Metabolism	1.99 (1.23, 3.25)	0.05	2.06 (1.26, 3.43)	0.04	2.07 (1.26, 3.44)	0.04
21	BCAA and Aromatic Amino Acids	2.96 (1.84, 4.85)	9.92E-05	2.83 (1.72, 4.74)	0.0005	2.79 (1.69, 4.66)	0.0006
40	Fatty Acid Biosynthesis	1.59 (1.19, 2.13)	0.02	1.54 (1.14, 2.09)	0.04	1.56 (1.15, 2.11)	0.04
22	Acyl Cholines	0.28 (0.18, 0.43)	1.28E-07	0.26 (0.16, 0.41)	1.12E-07	0.27 (0.17, 0.42)	2.61E-07
14	Amino Acid Metabolism	1.28 (0.73, 2.24)	1.00	1.37 (0.76, 2.48)	1.00	1.31 (0.72, 2.38)	1.00
26	Cell Membrane Components	0.90 (0.55, 1.49)	1.00	0.87 (0.51, 1.45)	1.00	0.85 (0.50, 1.44)	1.00
25	Aromatic Amino Acid Metabolism	0.91 (0.58, 1.42)	1.00	0.91 (0.56, 1.45)	1.00	0.93 (0.58, 1.48)	1.00

* Statistically significant cross-sectional associations between metabolite clusters and prevalent type 2 diabetes in logistic regression models in BPRHS are presented here. Associations between additional metabolite clusters and prevalent type 2 diabetes among BPRHS participants are presented in Supplemental Table S2.

† Model 1 adjusts for age and sex.

‡ Model 2 adjusts for Model 1 covariates plus smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), acculturation (%; BPRHS only), perceived-stress score, AHA-DS (quartiles; BPRHS only), mouthwash use (yes/no; SOALS only), and family history of type 2 diabetes (yes/no; SOALS only).

§ Model 3 adjusts for Model 2 covariates plus waist circumference (cm) and body mass index (kg/m²).

Abbreviations: AHA-DS, American Heart Association Diet Score; BCAA, branched-chain amino acids; FDR, Benjamini-Hochberg false discovery rate for 69 (BPRHS) or 9 (SOALS) metabolite clusters; SE, standard error.

Supplemental Table S3. Odds ratios (95% confidence intervals) for prevalent type 2 diabetes per 1-standard deviation increase in sum of metabolite concentrations in metabolite clusters among Boston Puerto Rican Health Study participants (n=679)*

Cluster Number	Model 1 [†]	FDR-adjusted <i>p</i> -value	Model 2 [‡]	FDR-adjusted <i>p</i> -value	Model 3 [§]	FDR-adjusted <i>p</i> -value
33	0.18 (0.13, 0.26)	1.63E-19	0.20 (0.13, 0.29)	7.51E-14	0.22 (0.14, 0.32)	8.12E-12
22	0.42 (0.31, 0.55)	1.94E-07	0.39 (0.28, 0.54)	1.65E-06	0.43 (0.31, 0.60)	7.95E-05
53	2.74 (1.99, 3.82)	1.03E-07	2.47 (1.71, 3.60)	0.0001	2.41 (1.65, 3.56)	0.0004
2	0.41 (0.29, 0.57)	5.87E-06	0.43 (0.29, 0.62)	6.18E-04	0.44 (0.30, 0.64)	0.002
25	2.03 (1.52, 2.74)	0.0002	1.76 (1.27, 2.45)	0.05	2.01 (1.43, 2.84)	0.004
21	1.91 (1.43, 2.58)	0.001	2.00 (1.44, 2.79)	0.003	1.96 (1.40, 2.77)	0.007
40	1.62 (1.33, 1.97)	8.12E-05	1.61 (1.29, 2.01)	0.002	1.54 (1.22, 1.94)	0.02
26	0.45 (0.32, 0.64)	0.001	0.49 (0.33, 0.73)	0.04	0.46 (0.30, 0.70)	0.02
14	2.74 (1.90, 4.00)	7.65E-06	2.45 (1.61, 3.77)	0.003	2.21 (1.45, 3.43)	0.02
36	0.49 (0.35, 0.67)	0.001	0.54 (0.37, 0.79)	0.09	0.53 (0.36, 0.77)	0.08
61	1.68 (1.25, 2.28)	0.04	1.76 (1.25, 2.48)	0.09	1.71 (1.21, 2.45)	0.20
5	1.37 (1.05, 1.79)	1.00	1.54 (1.14, 2.08)	0.35	1.57 (1.15, 2.15)	0.31
57	1.83 (1.38, 2.44)	0.003	1.57 (1.14, 2.19)	0.44	1.62 (1.16, 2.27)	0.35
59	1.57 (1.22, 2.03)	0.03	1.66 (1.24, 2.23)	0.04	1.51 (1.12, 2.05)	0.47
37	0.67 (0.55, 0.82)	0.01	0.68 (0.54, 0.86)	0.12	0.73 (0.57, 0.93)	0.75
32	1.67 (1.26, 2.22)	0.03	1.57 (1.14, 2.17)	0.45	1.52 (1.09, 2.13)	0.93
48	0.70 (0.57, 0.85)	0.02	0.72 (0.58, 0.90)	0.31	0.75 (0.59, 0.94)	0.95
34	1.40 (1.13, 1.74)	0.16	1.29 (1.02, 1.65)	1.00	1.36 (1.06, 1.76)	1.00
44	1.35 (1.08, 1.69)	0.69	1.41 (1.09, 1.84)	0.62	1.38 (1.06, 1.81)	1.00
50	1.42 (1.13, 1.80)	0.21	1.37 (1.05, 1.80)	1.00	1.37 (1.04, 1.80)	1.00
23	0.65 (0.51, 0.83)	0.05	0.73 (0.55, 0.96)	1.00	0.73 (0.54, 0.96)	1.00
16	0.65 (0.52, 0.81)	0.01	0.72 (0.56, 0.93)	0.90	0.75 (0.58, 0.97)	1.00
60	0.89 (0.73, 1.07)	1.00	0.80 (0.64, 1.00)	1.00	0.77 (0.61, 0.98)	1.00
43	0.83 (0.60, 1.13)	1.00	0.69 (0.48, 0.99)	1.00	0.67 (0.46, 0.98)	1.00
41	1.44 (1.17, 1.78)	0.04	1.28 (1.01, 1.62)	1.00	1.28 (1.01, 1.64)	1.00
11	1.99 (1.39, 2.87)	0.01	1.57 (1.04, 2.38)	1.00	1.54 (1.01, 2.36)	1.00
6	0.39 (0.27, 0.56)	2.62E-05	0.63 (0.41, 0.95)	1.00	0.65 (0.42, 0.99)	1.00

3	1.34 (1.00, 1.81)	1.00	1.40 (1.00, 1.98)	1.00	1.43 (1.01, 2.04)	1.00
8	1.33 (0.99, 1.80)	1.00	1.47 (1.05, 2.07)	1.00	1.43 (1.01, 2.04)	1.00
35	1.81 (1.38, 2.40)	0.002	1.59 (1.17, 2.18)	0.25	1.35 (0.97, 1.87)	1.00
10	1.40 (1.14, 1.73)	0.10	1.31 (1.03, 1.67)	1.00	1.25 (0.98, 1.60)	1.00
30	1.62 (1.27, 2.09)	0.01	1.33 (1.01, 1.77)	1.00	1.29 (0.97, 1.73)	1.00
1	1.47 (1.09, 2.01)	0.92	1.32 (0.93, 1.88)	1.00	1.32 (0.93, 1.90)	1.00
29	0.64 (0.45, 0.91)	0.97	0.75 (0.50, 1.11)	1.00	0.72 (0.47, 1.09)	1.00
28	1.04 (0.84, 1.28)	1.00	1.11 (0.87, 1.42)	1.00	1.20 (0.93, 1.54)	1.00
38	0.85 (0.65, 1.11)	1.00	0.73 (0.54, 0.99)	1.00	0.80 (0.59, 1.10)	1.00
7	0.62 (0.45, 0.85)	0.24	0.71 (0.50, 1.02)	1.00	0.78 (0.54, 1.12)	1.00
42	1.27 (1.05, 1.55)	1.00	1.14 (0.91, 1.42)	1.00	1.17 (0.93, 1.46)	1.00
13	0.88 (0.70, 1.10)	1.00	0.79 (0.61, 1.03)	1.00	0.84 (0.64, 1.10)	1.00
68	0.87 (0.70, 1.10)	1.00	0.78 (0.60, 1.01)	1.00	0.84 (0.64, 1.10)	1.00
19	0.74 (0.52, 1.06)	1.00	0.79 (0.53, 1.17)	1.00	0.77 (0.51, 1.15)	1.00
51	1.15 (0.90, 1.47)	1.00	1.14 (0.86, 1.51)	1.00	1.20 (0.90, 1.60)	1.00
65	1.53 (1.24, 1.90)	0.01	1.25 (0.98, 1.61)	1.00	1.17 (0.91, 1.51)	1.00
47	1.15 (0.89, 1.50)	1.00	1.16 (0.86, 1.58)	1.00	1.22 (0.89, 1.67)	1.00
69	1.11 (0.91, 1.37)	1.00	1.20 (0.95, 1.51)	1.00	1.16 (0.91, 1.47)	1.00
49	1.22 (0.95, 1.58)	1.00	1.24 (0.92, 1.67)	1.00	1.20 (0.88, 1.63)	1.00
45	1.15 (0.88, 1.51)	1.00	1.24 (0.90, 1.71)	1.00	1.21 (0.87, 1.69)	1.00
12	1.16 (0.88, 1.52)	1.00	1.01 (0.73, 1.39)	1.00	1.19 (0.86, 1.67)	1.00
27	1.20 (0.87, 1.67)	1.00	1.08 (0.75, 1.56)	1.00	1.23 (0.84, 1.80)	1.00
15	0.97 (0.73, 1.31)	1.00	0.86 (0.61, 1.20)	1.00	0.84 (0.59, 1.19)	1.00
17	0.84 (0.62, 1.14)	1.00	0.78 (0.55, 1.10)	1.00	0.83 (0.58, 1.20)	1.00
18	1.34 (1.01, 1.79)	1.00	1.22 (0.88, 1.70)	1.00	1.17 (0.83, 1.64)	1.00
4	1.38 (0.93, 2.05)	1.00	1.41 (0.89, 2.23)	1.00	1.24 (0.77, 1.98)	1.00
54	1.14 (0.90, 1.45)	1.00	1.09 (0.83, 1.43)	1.00	1.13 (0.85, 1.49)	1.00
67	0.99 (0.83, 1.18)	1.00	0.99 (0.81, 1.21)	1.00	1.08 (0.87, 1.33)	1.00
58	1.27 (1.02, 1.59)	1.00	1.11 (0.86, 1.43)	1.00	1.08 (0.83, 1.41)	1.00
66	1.04 (0.86, 1.26)	1.00	1.03 (0.83, 1.28)	1.00	1.06 (0.85, 1.33)	1.00
31	0.99 (0.79, 1.24)	1.00	1.06 (0.82, 1.38)	1.00	1.07 (0.82, 1.39)	1.00
63	1.02 (0.83, 1.24)	1.00	0.94 (0.75, 1.18)	1.00	0.94 (0.75, 1.20)	1.00

39	1.19 (0.88, 1.63)	1.00	0.97 (0.68, 1.39)	1.00	1.09 (0.75, 1.58)	1.00
46	1.09 (0.85, 1.40)	1.00	0.95 (0.71, 1.26)	1.00	1.07 (0.79, 1.44)	1.00
9	0.85 (0.64, 1.13)	1.00	0.93 (0.67, 1.29)	1.00	1.07 (0.76, 1.50)	1.00
24	1.22 (1.00, 1.49)	1.00	1.16 (0.93, 1.46)	1.00	1.04 (0.82, 1.32)	1.00
55	0.96 (0.74, 1.25)	1.00	0.85 (0.63, 1.14)	1.00	0.95 (0.69, 1.30)	1.00
52	1.08 (0.91, 1.29)	1.00	1.04 (0.85, 1.28)	1.00	1.03 (0.84, 1.27)	1.00
64	1.07 (0.85, 1.34)	1.00	1.02 (0.78, 1.32)	1.00	1.01 (0.77, 1.33)	1.00
6	0.94 (0.75, 1.16)	1.00	1.00 (0.79, 1.28)	1.00	1.01 (0.78, 1.29)	1.00
56	0.81 (0.64, 1.03)	1.00	0.90 (0.68, 1.19)	1.00	1.00 (0.75, 1.34)	1.00
20	0.90 (0.68, 1.20)	1.00	1.02 (0.74, 1.41)	1.00	1.00 (0.72, 1.40)	1.00

* All cross-sectional associations between metabolite clusters and prevalent type 2 diabetes in logistic regression models presented here.

†Model 1 adjusts for age and sex.

‡ Model 2 adjusts for Model 1 covariates plus smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), acculturation (%), perceived-stress score, and AHA-DS (quartiles; BPRHS only).

§ Model 3 adjusts for Model 2 covariates plus waist circumference (cm) and body mass index (kg/m²).

Abbreviations: AHA-DS, American Heart Association Diet Score; FDR, Benjamini-Hochberg false discovery rate for 69 metabolite clusters; SE, standard error.

Supplemental Table S4. Odds ratios (95% confidence intervals) for prevalent type 2 diabetes per 1-standard deviation difference in unweighted sum of metabolite concentrations in top metabolite clusters among Boston Puerto Rican Health Study (BPRHS) (n=679) and San Juan Overweight Adults Longitudinal Study (n=1,011) participants*

Cluster #	Cluster Name	n _{metab} [†]	BPRHS		SOALS	
			OR (95% CI)	FDR	OR (95% CI)	FDR
33	Glucose Transport	6	0.78 (0.72, 0.83)	9.33E-12	0.78 (0.72, 0.84)	9.66E-08
2	Sphingolipids	18	0.96 (0.94, 0.98)	0.004	0.94 (0.91, 0.96)	0.0006
53	Sugar Metabolism	5	1.20 (1.12, 1.30)	0.0002	1.16 (1.05, 1.28)	0.33
21	BCAA and Aromatic Amino Acids	13	1.05 (1.03, 1.08)	0.005	1.08 (1.04, 1.13)	0.005
40	Fatty Acid Biosynthesis	6	1.08 (1.04, 1.12)	0.01	1.08 (1.03, 1.14)	0.25
22	Acyl Cholines	9	0.91 (0.88, 0.95)	0.0001	0.86 (0.81, 0.90)	6.66E-08
14	Amino Acid Metabolism	15	1.05 (1.02, 1.08)	0.09	1.01 (0.97, 1.05)	1.00
26	Cell Membrane Components	8	0.91 (0.86, 0.96)	0.03	0.97 (0.90, 1.04)	1.00
25	Aromatic Amino Acid Metabolism	11	1.07 (1.04, 1.10)	0.003	1.00 (0.95, 1.04)	1.00

* Only results from statistically significant cross-sectional associations between weighted metabolite clusters and prevalent type 2 diabetes in logistic regression models among BPRHS participants are presented here. Model adjusts for age, sex, smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), acculturation, perceived-stress score, AHA-DS (quartiles), waist circumference (cm) and body mass index (kg/m²).

[†] Number of metabolites included in each cluster

Abbreviations: AHA-DS, American Heart Association Diet Score; BCAA, branched-chain amino acids; FDR, Benjamini-Hochberg false discovery rate for 69 metabolite clusters; SE, standard error.

Supplemental Table S5. Odds ratios (95% confidence intervals) for prevalent type 2 diabetes per 1-standard deviation difference in weighted sum of metabolite concentrations among Boston Puerto Rican Health Study (BPRHS) (n=679) participants*

Cluster #	Cluster Name	OR (95% CI)	<i>p</i>
33	Glucose Transport	0.23 (0.13, 0.39)	2.00E-07
2	Sphingolipids	0.89 (0.55, 1.45)	0.64
53	Sugar Metabolism	2.31 (1.47, 3.67)	0.0004
21	BCAA and Aromatic Amino Acids	1.63 (1.05, 2.56)	0.03
40	Fatty Acid Biosynthesis	1.41 (1.08, 1.86)	0.01
22	Acyl Cholines	1.01 (0.62, 1.64)	0.97
14	Amino Acid Metabolism	1.42 (0.80, 2.52)	0.23
26	Cell Membrane Components	0.66 (0.38, 1.12)	0.12
25	Aromatic Amino Acid Metabolism	1.48 (1.00, 2.20)	0.05

* Odds ratios from a joint model including all metabolite clusters significantly associated with prevalent type 2 diabetes are included in a single logistic regression model among BPRHS. Model adjusts for age, sex, smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), acculturation (%), perceived-stress score, AHA-DS, waist circumference (cm) and body mass index (kg/m²).

Abbreviations: AHA-DS, American Heart Association Diet Score; BCAA, branched-chain amino acids; CI, confidence interval.

Supplemental Table S6. Mean difference in cardiometabolic risk factors per 1-SD difference in sum of weighted metabolite concentrations in each cluster among Boston Puerto Rican Health Study participants without type 2 diabetes (n=322)*

Outcome	HDL-C (mg/dl)		TG (ln-mg/dl)		WC (cm)		BMI (kg/m2)	
	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR
Cluster 2: Sphingolipids	3.7 (1.6)	1.00	-0.10 (0.06)	1.00	-0.42 (1.67)	1.00	-0.005 (0.73)	1.00
Cluster 14: Amino Acid Metabolism	-2.0 (1.8)	1.00	0.13 (0.07)	1.00	5.01 (1.79)	0.38	2.50 (0.78)	0.10
Cluster 21: BCAA and Aromatic Amino Acids	1.9 (1.5)	1.00	0.20 (0.06)	0.02	4.79 (1.45)	0.07	1.55 (0.64)	1.00
Cluster 22: Acyl Cholines	5.2 (1.4)	0.02	0.13 (0.05)	1.00	-3.18 (1.44)	1.00	-1.36 (0.63)	1.00
Cluster 25: Aromatic Amino Acid Metabolism	-2.1 (1.4)	1.00	-0.03 (0.06)	1.00	-2.07 (1.48)	1.00	-0.76 (0.65)	1.00
Cluster 26: Cell Membrane Components	3.6 (1.8)	1.00	0.28 (0.07)	0.002	1.58 (1.82)	1.00	0.48 (0.80)	1.00
Cluster 33: Glucose Transport	5.4 (1.6)	0.07	0.18 (0.06)	0.30	-2.95 (1.67)	1.00	-1.24 (0.73)	1.00
Cluster 40: Fatty Acid Biosynthesis	3.8 (0.9)	0.002	-0.004 (0.04)	1.00	0.98 (0.97)	1.00	0.55 (0.43)	1.00
Cluster 53: Sugar Metabolism	0.2 (1.6)	1.00	0.17 (0.06)	0.24	1.21 (1.60)	1.00	-0.03 (0.71)	1.00
	Glucose		Insulin (ln- μ IU/mL)		Glyc-Hb, %		HOMA-IR	
	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR
Cluster 2: Sphingolipids	-0.55 (1.17)	1.00	-0.06 (0.07)	1.00	0.06 (0.05)	1.00	-0.32 (0.25)	1.00
Cluster 14: Amino Acid Metabolism	2.17 (1.30)	1.00	0.31 (0.08)	0.003	0.03 (0.06)	1.00	0.80 (0.27)	0.22
Cluster 21: BCAA and Aromatic Amino Acids	1.82 (1.06)	1.00	0.12 (0.06)	1.00	0.04 (0.05)	1.00	0.31 (0.22)	1.00
Cluster 22: Acyl Cholines	-1.05 (1.03)	1.00	-0.12 (0.06)	1.00	-0.07 (0.05)	1.00	-0.39 (0.22)	1.00
Cluster 25: Aromatic Amino Acid Metabolism	0.09 (1.05)	1.00	-0.04 (0.06)	1.00	-0.04 (0.05)	1.00	0.04 (0.22)	1.00
Cluster 26: Cell Membrane Components	0.07 (1.28)	1.00	0.04 (0.08)	1.00	-0.08 (0.06)	1.00	0.03 (0.27)	1.00
Cluster 33: Glucose Transport	-0.90 (1.19)	1.00	-0.04 (0.07)	1.00	-0.05 (0.05)	1.00	-0.08 (0.25)	1.00
Cluster 40: Fatty Acid Biosynthesis	0.68 (0.68)	1.00	-0.06 (0.04)	1.00	0.07 (0.03)	1.00	-0.23 (0.14)	1.00
Cluster 53: Sugar Metabolism	0.63 (1.13)	1.00	0.11 (0.07)	1.00	-0.02 (0.05)	1.00	0.41 (0.24)	1.00

* Cross-sectional associations adjusted for age, sex, AHA diet score, smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), acculturation (%), perceived-stress score, WC (cm) and BMI (kg/m2) (except for WC and BMI outcomes). P-values are corrected for multiple exposures and outcomes using the Benjamani-Hochberg false discovery rate.

Abbreviations: BP, blood pressure; BMI, body mass index; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostatic model assessment of insulin resistance; LDL-C, low-density lipoprotein cholesterol; TG, triglyceride; WC, waist circumference.

Supplemental Table S7. Mean difference in cardiometabolic risk factors per 1-SD difference in sum of weighted metabolite concentrations in each cluster among San Juan Overweight Adult Longitudinal Study participants without type 2 diabetes (n=934)*

Outcome	HDL-C (mg/dl)		TG (ln-mg/dl)		WC (cm)		BMI (kg/m ²)	
	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR
Cluster 2: Sphingolipids	6.1 (0.7)	2.21E-14	0.02 (0.03)	1.00	1.63 (0.89)	1.00	1.16 (0.40)	0.25
Cluster 14: Amino Acid Metabolism	-0.8 (0.9)	1.00	0.09 (0.04)	0.96	3.39 (1.08)	0.12	1.42 (0.48)	0.24
Cluster 21: BCAA and Aromatic Amino Acids	-1.3 (0.8)	1.00	0.24 (0.03)	1.19E-13	2.19 (0.91)	1.00	1.08 (0.41)	0.58
Cluster 22: Acyl Cholines	4.0 (0.7)	1.07E-06	0.22 (0.03)	2.51E-13	-2.81 (0.83)	0.05	-0.96 (0.37)	0.67
Cluster 25: Aromatic Amino Acid Metabolism	-0.4 (0.8)	1.00	-0.05 (0.03)	1.00	-1.44 (0.87)	1.00	-0.74 (0.39)	1.00
Cluster 26: Cell Membrane Components	2.8 (0.8)	0.06	0.27 (0.03)	2.02E-14	2.73 (0.98)	0.38	1.40 (0.44)	0.09
Cluster 33: Glucose Transport	2.8 (0.7)	0.02	0.22 (0.03)	9.44E-13	0.70 (0.87)	1.00	0.43 (0.39)	1.00
Cluster 40: Fatty Acid Biosynthesis	1.9 (0.5)	0.01	0.02 (0.02)	1.00	0.47 (0.57)	1.00	0.51 (0.26)	1.00
Cluster 53: Sugar Metabolism	-1.0 (0.8)	1.00	0.19 (0.03)	7.44E-08	0.37 (0.94)	1.00	0.30 (0.42)	1.00

Outcome	Glucose		Insulin (ln- μ IU/mL)		Glyc-Hb, %		HOMA-IR	
	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR
Cluster 2: Sphingolipids	-0.65 (0.57)	1.00	-0.07 (0.03)	1.00	0.02 (0.02)	1.00	-0.13 (0.10)	1.00
Cluster 14: Amino Acid Metabolism	1.96 (0.69)	0.30	0.20 (0.04)	0.0001	0.08 (0.03)	0.13	0.56 (0.12)	0.0003
Cluster 21: BCAA and Aromatic Amino Acids	2.05 (0.57)	0.03	0.27 (0.03)	2.55E-14	0.09 (0.02)	0.001	0.73 (0.10)	1.60E-11
Cluster 22: Acyl Cholines	-3.34 (0.52)	8.04E-09	-0.16 (0.03)	1.37E-05	-0.06 (0.02)	0.27	-0.48 (0.09)	1.53E-05
Cluster 25: Aromatic Amino Acid Metabolism	0.13 (0.55)	1.00	-0.02 (0.03)	1.00	0.03 (0.02)	1.00	0.01 (0.10)	1.00
Cluster 26: Cell Membrane Components	-0.78 (0.62)	1.00	0.03 (0.04)	1.00	-0.02 (0.02)	1.00	0.10 (0.11)	1.00
Cluster 33: Glucose Transport	-0.36 (0.55)	1.00	-0.02 (0.03)	1.00	0.003 (0.02)	1.00	-0.01 (0.10)	1.00
Cluster 40: Fatty Acid Biosynthesis	0.41 (0.36)	1.00	-0.02 (0.02)	1.00	-0.01 (0.01)	1.00	0.01 (0.06)	1.00
Cluster 53: Sugar Metabolism	1.75 (0.59)	0.23	0.11 (0.04)	0.18	0.07 (0.02)	0.05	0.32 (0.10)	0.13

* Cross-sectional associations adjusted for age, sex, smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), perceived-stress score, mouthwash use (yes/no), family history of type 2 diabetes (yes/no), WC (cm) and BMI (kg/m²) (except for BMI and WC outcomes). P-values are corrected for multiple exposures and outcomes using the Benjamani-Hochberg false discovery rate.

Abbreviations: BP, blood pressure; BMI, body mass index; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostatic model assessment of insulin resistance; LDL-C, low-density lipoprotein cholesterol; TG, triglyceride; WC, waist circumference.

Supplemental Table S8. Mean difference in cardiometabolic risk factors per 1-SD difference in sum of weighted metabolite concentrations in each cluster among all Boston Puerto Rican Health Study participants (n=679).*

Outcome	HDL-C (mg/dl)		TG (ln-mg/dl)		WC (cm)		BMI (kg/m ²)	
	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR
Cluster 2: Sphingolipids	4.2 (0.9)	0.001	-0.05 (0.04)	1.00	-0.64 (1.17)	1.00	0.56 (0.51)	1.00
Cluster 14: Amino Acid Metabolism	-2.6 (1.1)	1.00	0.11 (0.05)	1.00	2.56 (1.31)	1.00	1.00 (0.58)	1.00
Cluster 21: BCAA and Aromatic Amino Acids	-0.1 (0.8)	1.00	0.23 (0.04)	1.64E-08	0.53 (1.04)	1.00	0.07 (0.46)	1.00
Cluster 22: Acyl Cholines	3.3 (0.8)	0.01	0.17 (0.04)	1.68E-04	-3.14 (1.01)	0.14	-1.05 (0.45)	1.00
Cluster 25: Aromatic Amino Acid Metabolism	-0.9 (0.8)	1.00	-0.02 (0.04)	1.00	-2.50 (1.03)	1.00	-1.48 (0.45)	0.07
Cluster 26: Cell Membrane Components	3.8 (1.0)	0.02	0.30 (0.04)	1.70E-09	1.70 (1.28)	1.00	0.96 (0.56)	1.00
Cluster 33: Glucose Transport	2.3 (0.9)	0.83	0.25 (0.04)	1.65E-08	-3.46 (1.11)	0.13	-0.87 (0.49)	1.00
Cluster 40: Fatty Acid Biosynthesis	1.6 (0.6)	0.28	0.01 (0.03)	1.00	1.63 (0.69)	1.00	0.89 (0.30)	0.22
Cluster 53: Sugar Metabolism	-1.9 (0.9)	1.00	0.25 (0.04)	9.51E-08	1.46 (1.15)	1.00	0.30 (0.51)	1.00

Outcome	Glucose		Insulin (ln- μ IU/mL)		Glyc-Hb, %		HOMA-IR	
	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR
Cluster 2: Sphingolipids	-8.96 (3.6)	0.93	-0.06 (0.06)	1.00	-0.13 (0.12)	1.00	-1.50 (0.77)	1.00
Cluster 14: Amino Acid Metabolism	4.87 (4.05)	1.00	0.35 (0.06)	1.35E-06	-0.14 (0.13)	1.00	1.54 (0.86)	1.00
Cluster 21: BCAA and Aromatic Amino Acids	9.76 (3.17)	0.15	0.25 (0.05)	1.72E-05	-0.17 (0.10)	1.00	1.52 (0.68)	1.00
Cluster 22: Acyl Cholines	-18.9 (3.06)	4.54E-08	-0.13 (0.05)	0.73	-0.35 (0.10)	0.04	-2.43 (0.66)	0.02
Cluster 25: Aromatic Amino Acid Metabolism	2.71 (3.2)	1.00	0.04 (0.05)	1.00	0.15 (0.10)	1.00	0.97 (0.68)	1.00
Cluster 26: Cell Membrane Components	-9.52 (3.92)	1.00	0.03 (0.06)	1.00	-0.45 (0.12)	0.02	-1.00 (0.83)	1.00
Cluster 33: Glucose Transport	-30.0 (3.23)	1.17E-18	-0.06 (0.05)	1.00	-0.86 (0.10)	1.90E-14	-2.51 (0.73)	0.04
Cluster 40: Fatty Acid Biosynthesis	7.99 (2.11)	0.01	-0.10 (0.03)	0.24	0.17 (0.07)	0.94	-0.91 (0.45)	1.00
Cluster 53: Sugar Metabolism	6.62 (3.53)	1.00	0.18 (0.05)	0.08	0.04 (0.11)	1.00	1.55 (0.75)	1.00

* Cross-sectional associations adjusted for age, sex, AHA diet score, smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), acculturation (%), perceived-stress score, WC (cm) and BMI (kg/m²) (except for WC and BMI outcomes). P-values are corrected for multiple exposures and outcomes using the Benjamani-Hochberge false discovery rate.

Abbreviations: BP, blood pressure; BMI, body mass index; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostatic model assessment of insulin resistance; LDL-C, low-density lipoprotein cholesterol; TG, triglyceride; WC, waist circumference.

Supplemental Table S9. Mean difference in cardiometabolic risk factors per 1-SD difference in sum of weighted metabolite concentrations in each cluster among all San Juan Overweight Adult Longitudinal Study participants (n=1,011).*

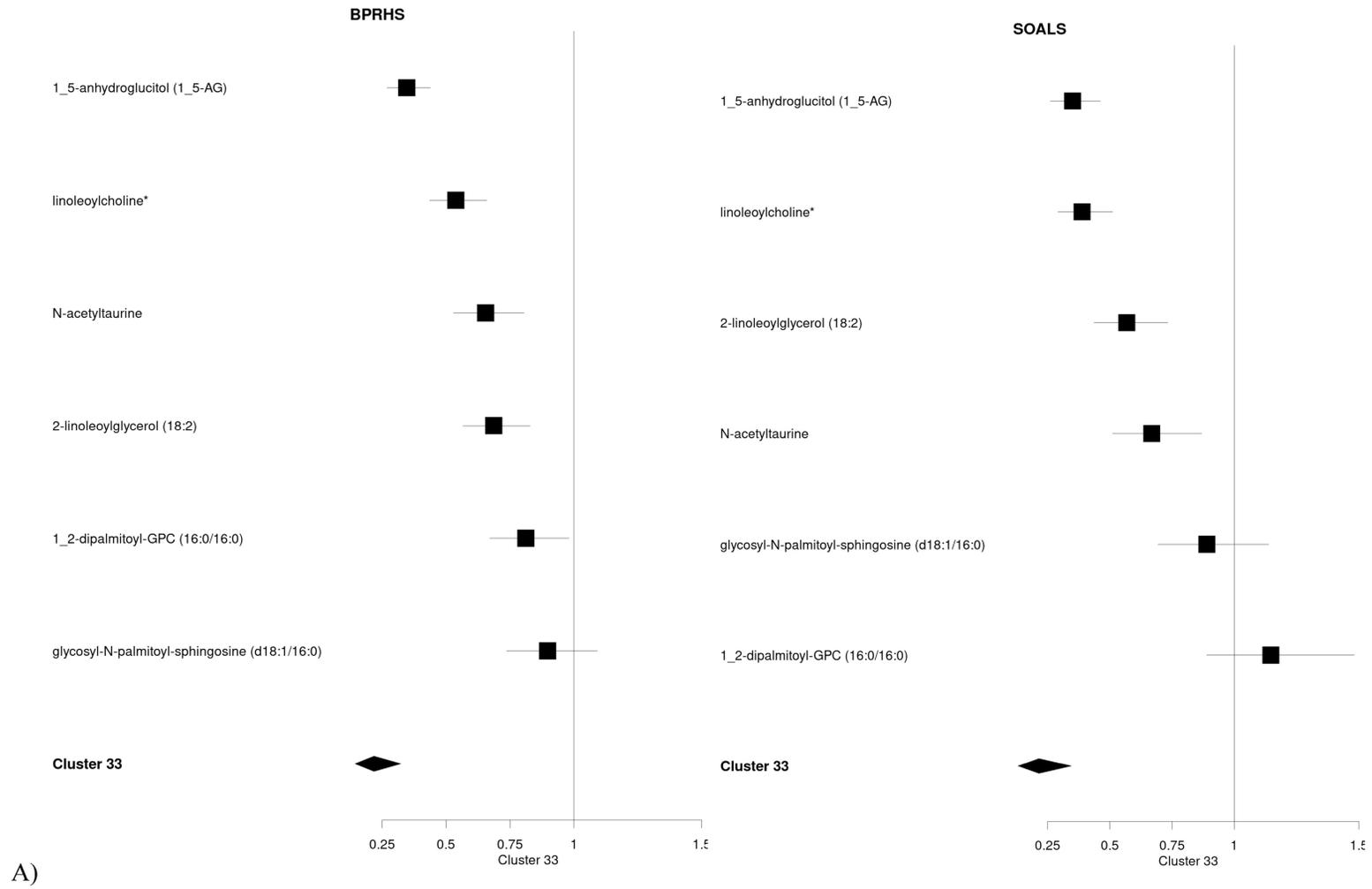
Outcome	HDL-C (mg/dl)		TG (ln-mg/dl)		WC (cm)		BMI (kg/m ²)	
	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR
Cluster 2: Sphingolipids	5.7 (0.8)	2.21E-14	0.01 (0.03)	1.00	1.31 (0.85)	1.00	1.05 (0.38)	0.25
Cluster 14: Amino Acid Metabolism	-0.8 (1.0)	1.00	0.08 (0.04)	0.96	3.54 (1.03)	0.12	1.49 (0.46)	0.24
Cluster 21: BCAA and Aromatic Amino Acids	-0.9 (0.8)	1.00	0.24 (0.03)	1.19E-13	2.29 (0.86)	1.00	1.08 (0.38)	0.58
Cluster 22: Acyl Cholines	3.9 (0.7)	1.07E-06	0.18 (0.03)	2.51E-13	-2.98 (0.77)	0.05	-1.06 (0.35)	0.67
Cluster 25: Aromatic Amino Acid Metabolism	-0.5 (0.8)	1.00	-0.05 (0.03)	1.00	-1.61 (0.83)	1.00	-0.82 (0.37)	1.00
Cluster 26: Cell Membrane Components	3.0 (0.9)	0.06	0.25 (0.03)	2.02E-14	2.75 (0.93)	0.38	1.41 (0.42)	0.09
Cluster 33: Glucose Transport	2.7 (0.7)	0.02	0.17 (0.03)	9.44E-13	0.14 (0.80)	1.00	0.22 (0.36)	1.00
Cluster 40: Fatty Acid Biosynthesis	2.5 (0.5)	0.01	0.02 (0.02)	1.00	0.56 (0.54)	1.00	0.53 (0.24)	1.00
Cluster 53: Sugar Metabolism	-0.9 (0.8)	1.00	0.19 (0.03)	7.44E-08	0.65 (0.89)	1.00	0.32 (0.40)	1.00

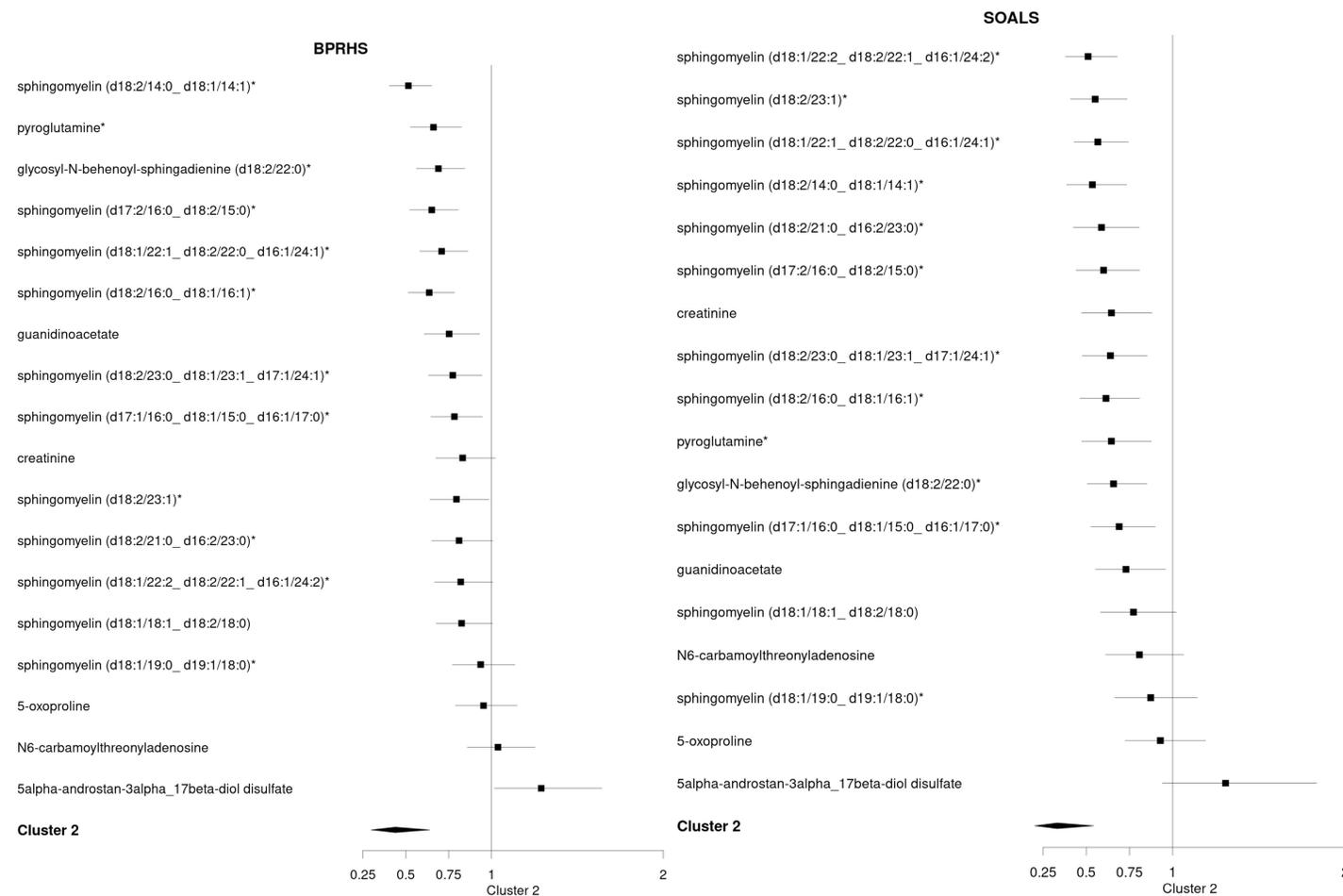
Outcome	Glucose		Insulin (ln- μ IU/mL)		Glyc-Hb, %		HOMA-IR	
	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR	β (SE)	FDR
Cluster 2: Sphingolipids	-5.52 (1.25)	1.00	-0.10 (0.03)	1.00	-0.10 (0.04)	1.00	-0.36 (0.10)	1.00
Cluster 14: Amino Acid Metabolism	2.92 (1.55)	0.30	0.18 (0.04)	0.0001	0.12 (0.05)	0.13	0.53 (0.13)	0.0003
Cluster 21: BCAA and Aromatic Amino Acids	6.08 (1.26)	0.03	0.28 (0.03)	2.55E-14	0.18 (0.04)	0.001	0.86 (0.10)	1.60E-11
Cluster 22: Acyl Cholines	-9.69 (1.12)	8.04E-09	-0.18 (0.03)	1.37E-05	-0.24 (0.03)	0.27	-0.70 (0.09)	1.53E-05
Cluster 25: Aromatic Amino Acid Metabolism	-0.57 (1.24)	1.00	-0.01 (0.03)	1.00	-0.01 (0.04)	1.00	0.01 (0.10)	1.00
Cluster 26: Cell Membrane Components	-2.51 (1.40)	1.00	0.02 (0.04)	1.00	-0.09 (0.04)	1.00	0.05 (0.12)	1.00
Cluster 33: Glucose Transport	-8.89 (1.15)	1.00	-0.05 (0.03)	1.00	-0.26 (0.04)	1.00	-0.34 (0.10)	1.00
Cluster 40: Fatty Acid Biosynthesis	3.16 (0.80)	1.00	-0.01 (0.02)	1.00	0.06 (0.02)	1.00	0.09 (0.07)	1.00
Cluster 53: Sugar Metabolism	5.19 (1.31)	0.23	0.11 (0.03)	0.18	0.19 (0.04)	0.05	0.44 (0.11)	0.13

* Cross-sectional associations adjusted for age, sex, smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), acculturation (%), perceived-stress score, mouthwash use (yes/no), and family history of type 2 diabetes (yes/no), WC (cm) and BMI (kg/m²) (except for WC and BMI outcomes). P-values are corrected for multiple exposures and outcomes using the Benjamani-Hochberge false discovery rate.

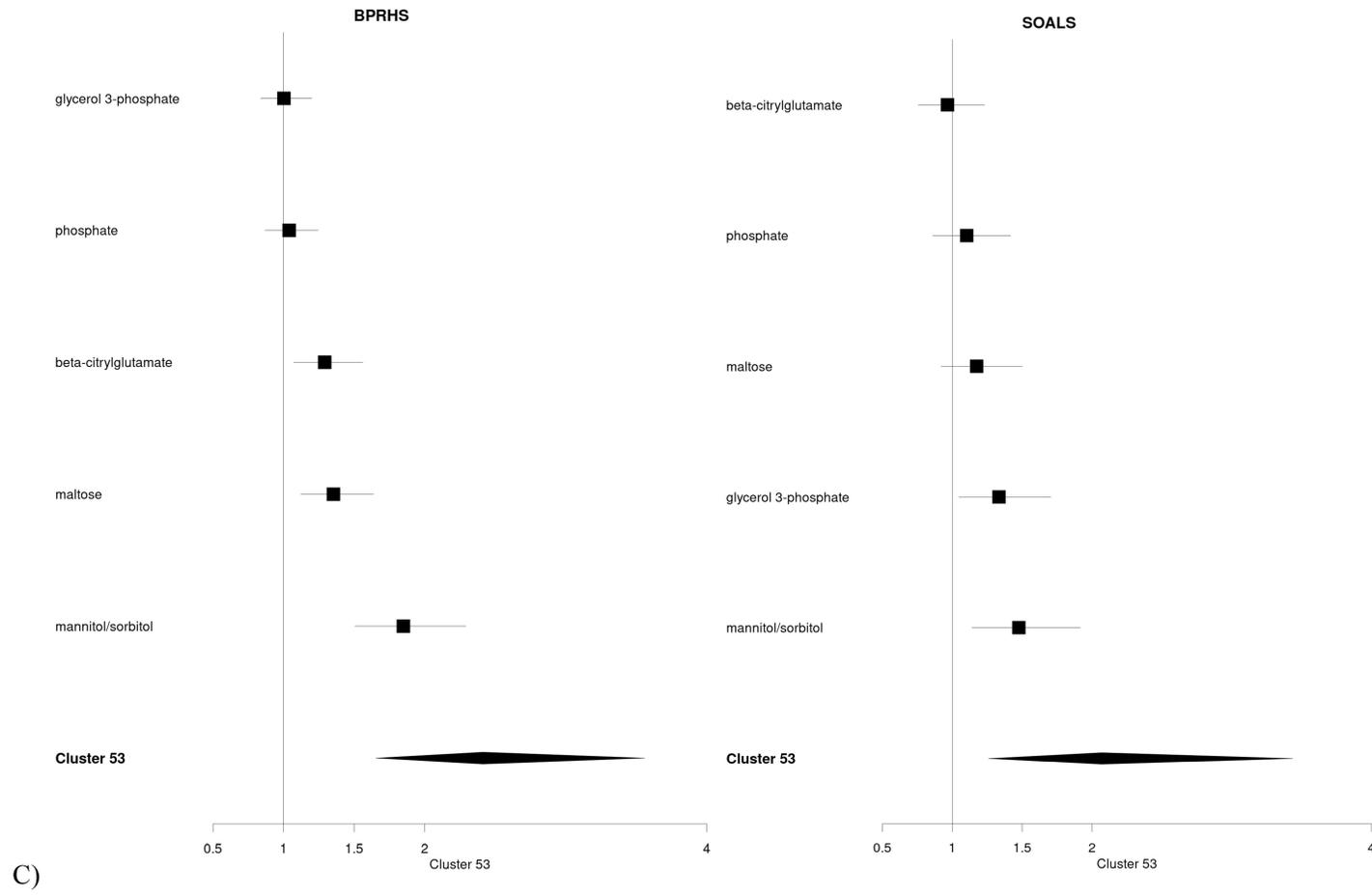
Abbreviations: BP, blood pressure; BMI, body mass index; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostatic model assessment of insulin resistance; LDL-C, low-density lipoprotein cholesterol; TG, triglyceride; WC, waist circumference.

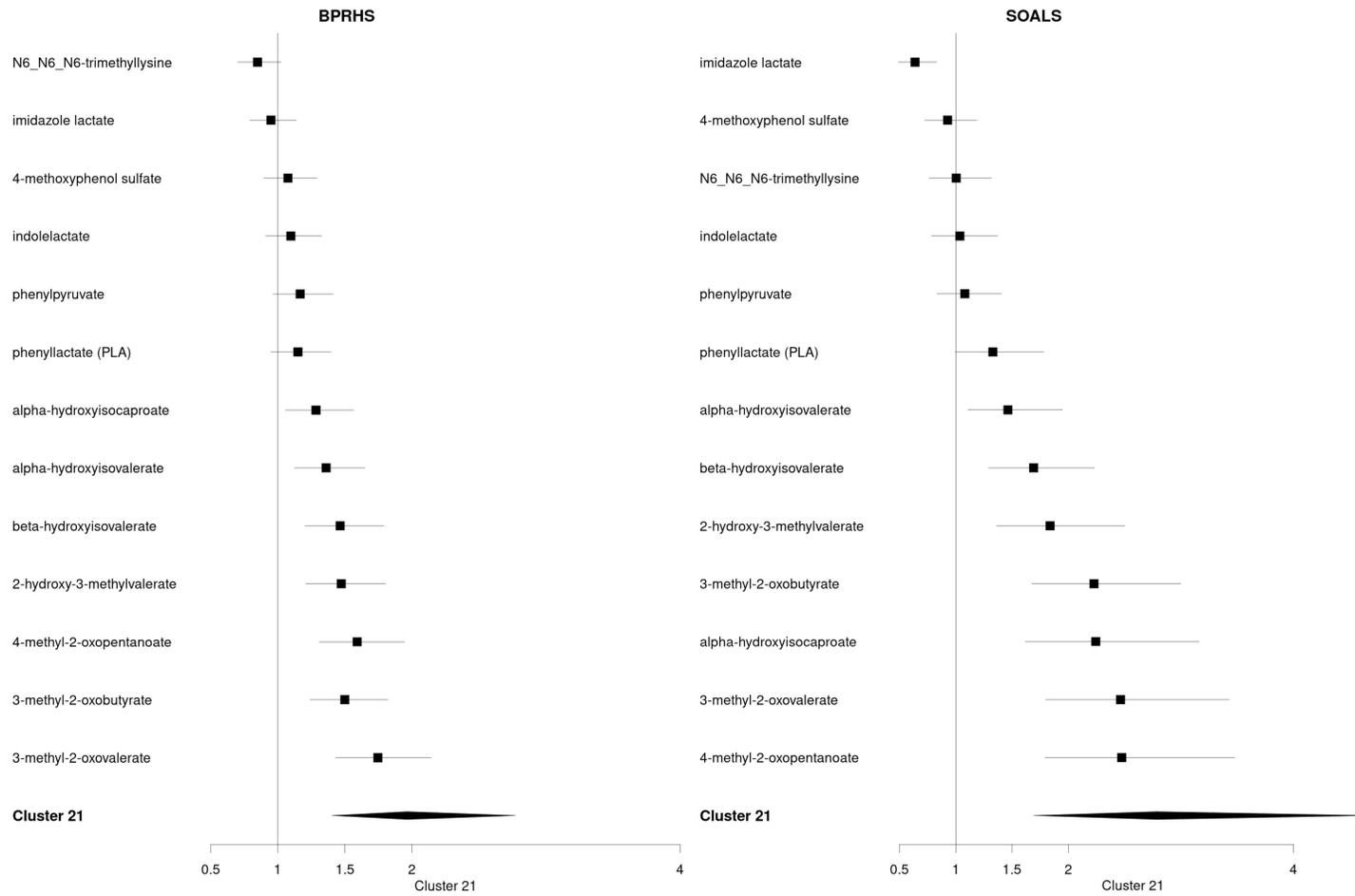
Supplemental Figure S1. Odds ratios (95% confidence intervals) for prevalent type 2 diabetes per 1-standard deviation difference in individual metabolite concentrations in network clusters significantly associated with prevalent type 2 diabetes in the Boston Puerto Rican Health Study (n=679) and SOALS (n=1,011). Logistic regression models adjust for age, sex, smoking status (former, current, never), education, physical activity, alcohol intake, lipid-lowering medication (yes/no), hypertension medication (yes/no), income (\$/year), acculturation (%), perceived-stress score, American Heart Association diet score (quartiles; BPRHS only), mouthwash use (yes/no; SOALS only), family history of type 2 diabetes (yes/no; SOALS only), waist circumference (cm) and body mass index (kg/m²). A) Cluster 33: Glucose Transport; B) Cluster 2: Sphingolipids; C) Cluster 53: Sugar Metabolism; D) Cluster 21: BCAA and Aromatic Amino Acids; E) Cluster 40: Fatty Acid Biosynthesis; F) Cluster 22: Acyl Cholines; G) Cluster 14: Amino Acid Metabolism; H) Cluster 26: Cell Membrane Components; and I) Cluster 25: Aromatic Amino Acid Metabolism.



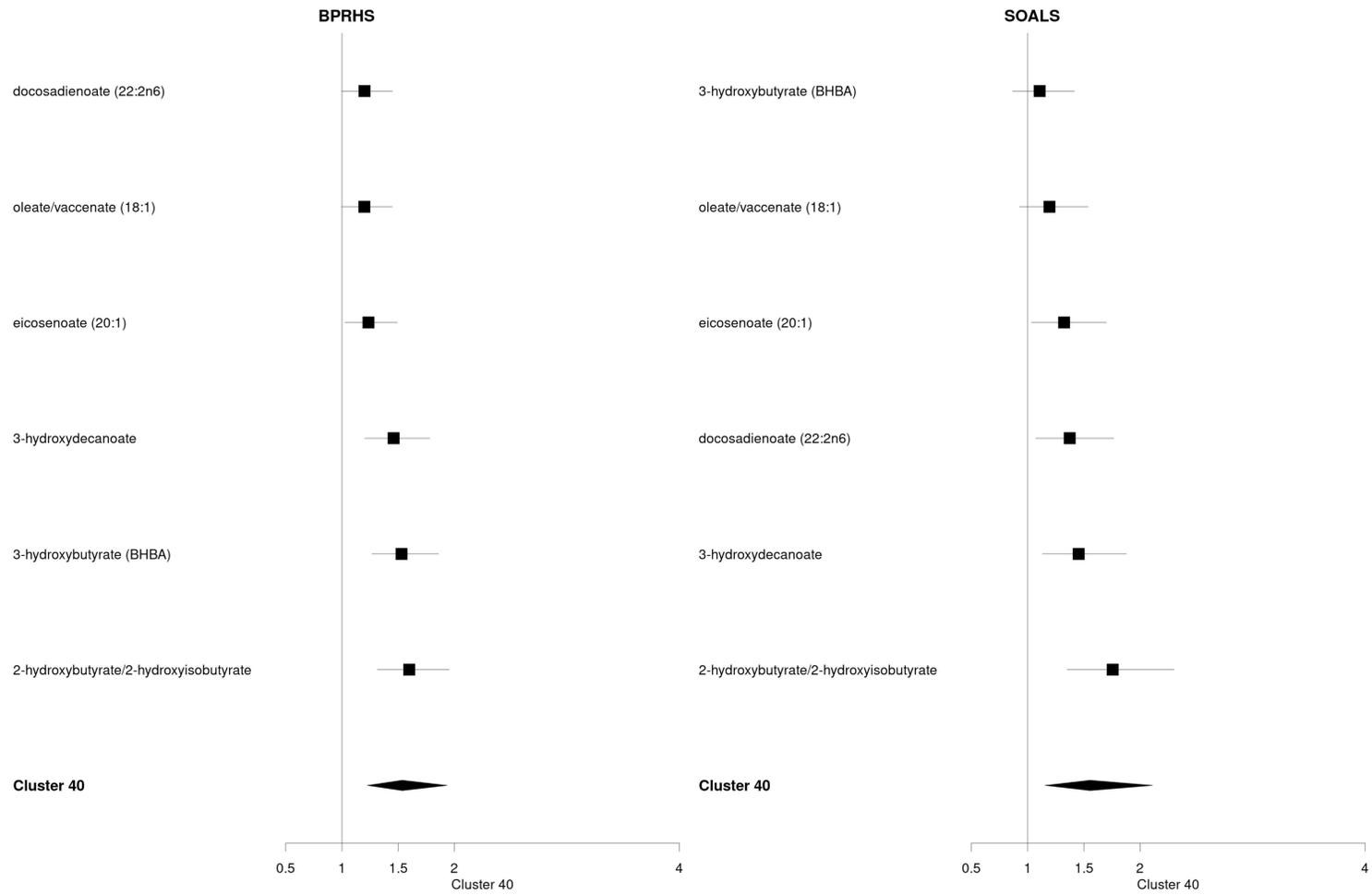


B)

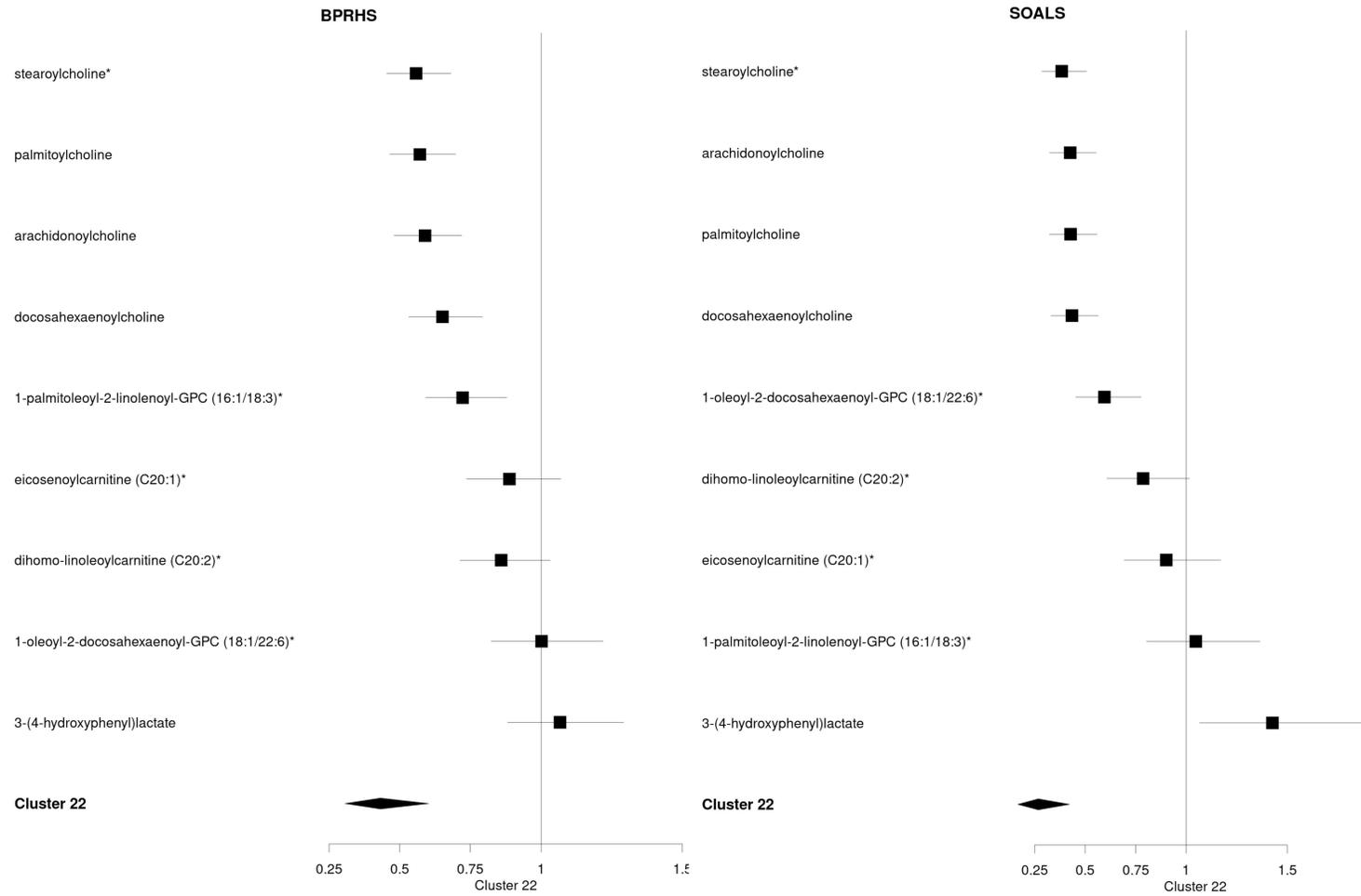




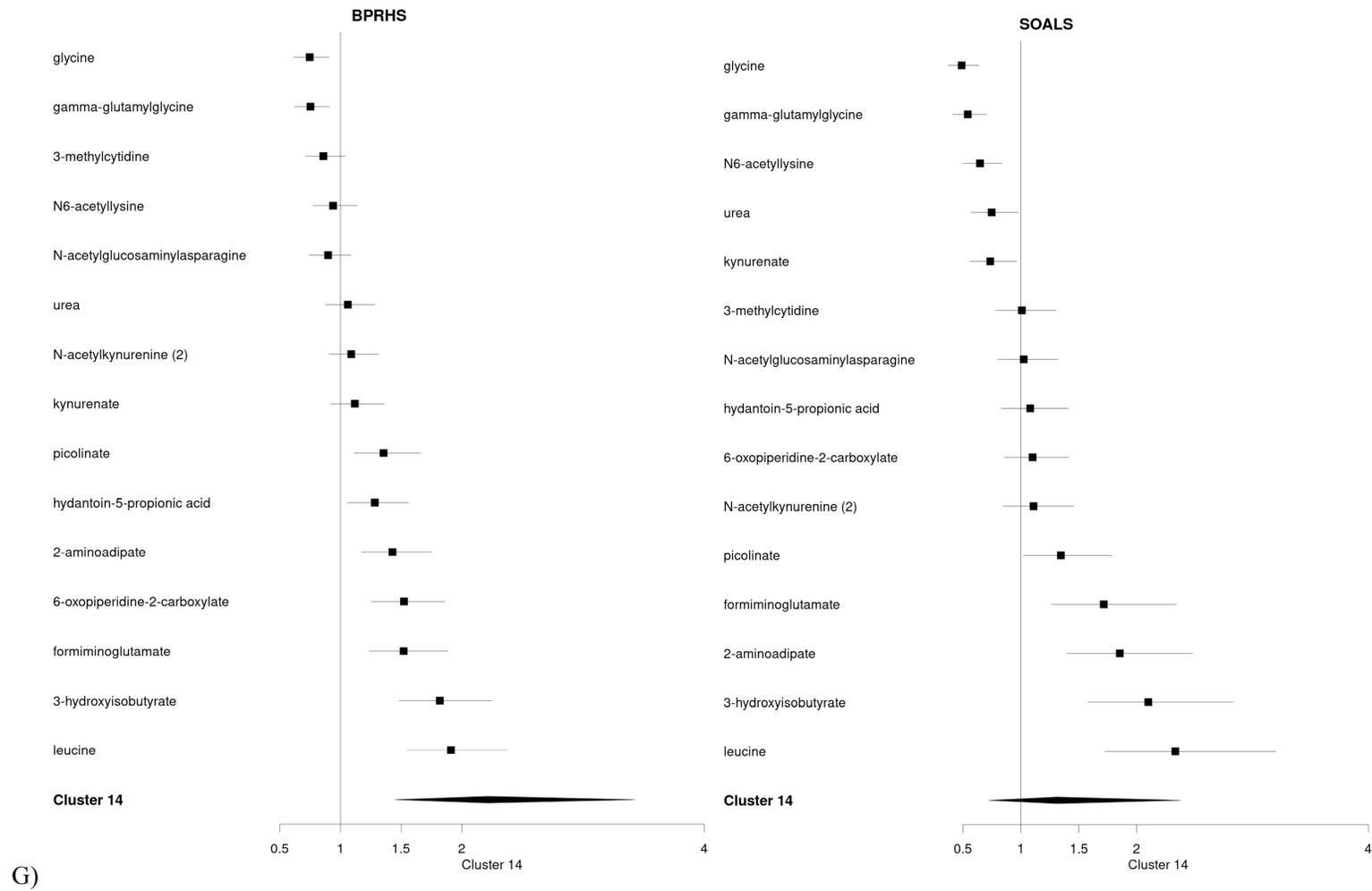
D)

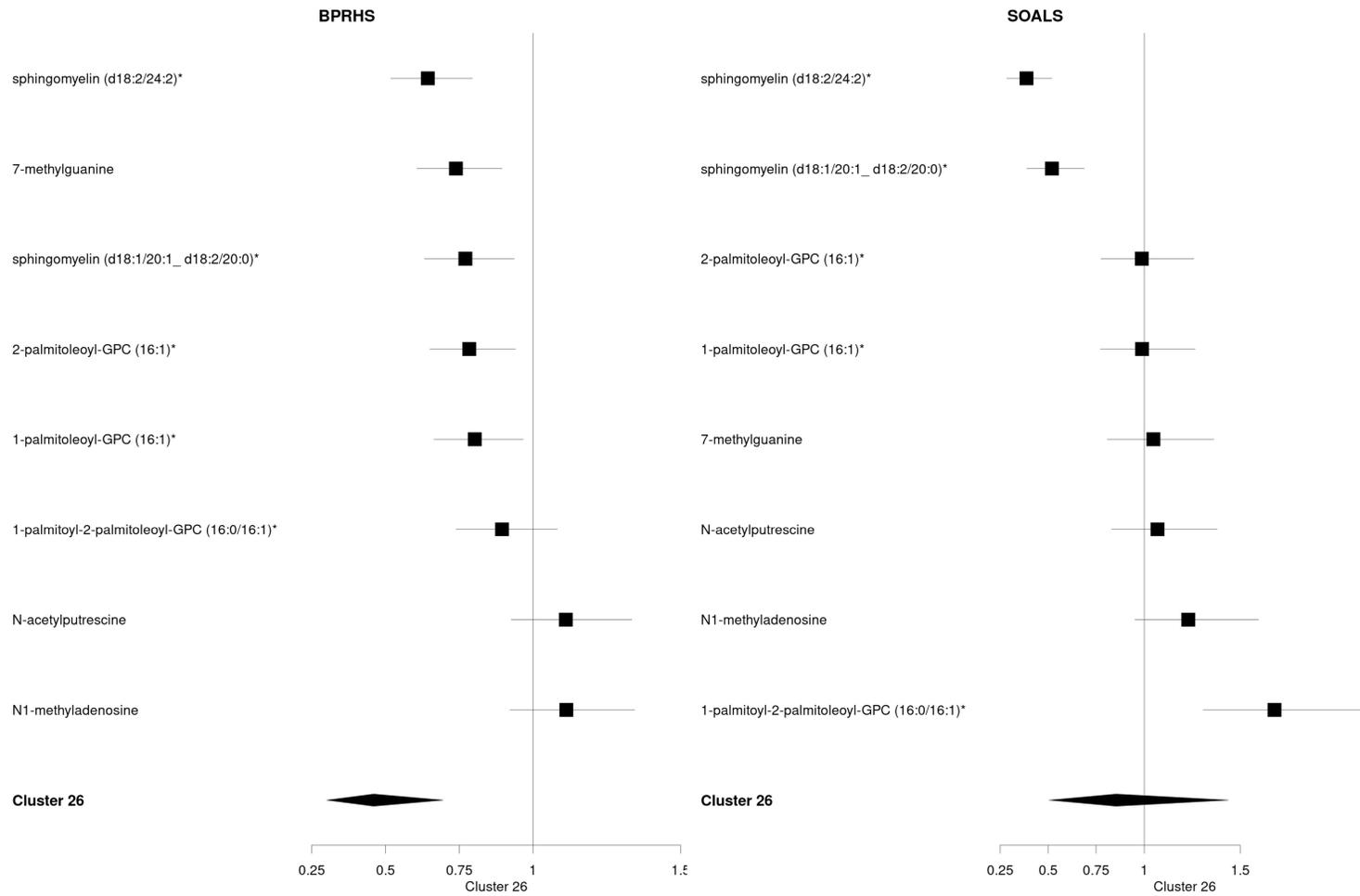


E)

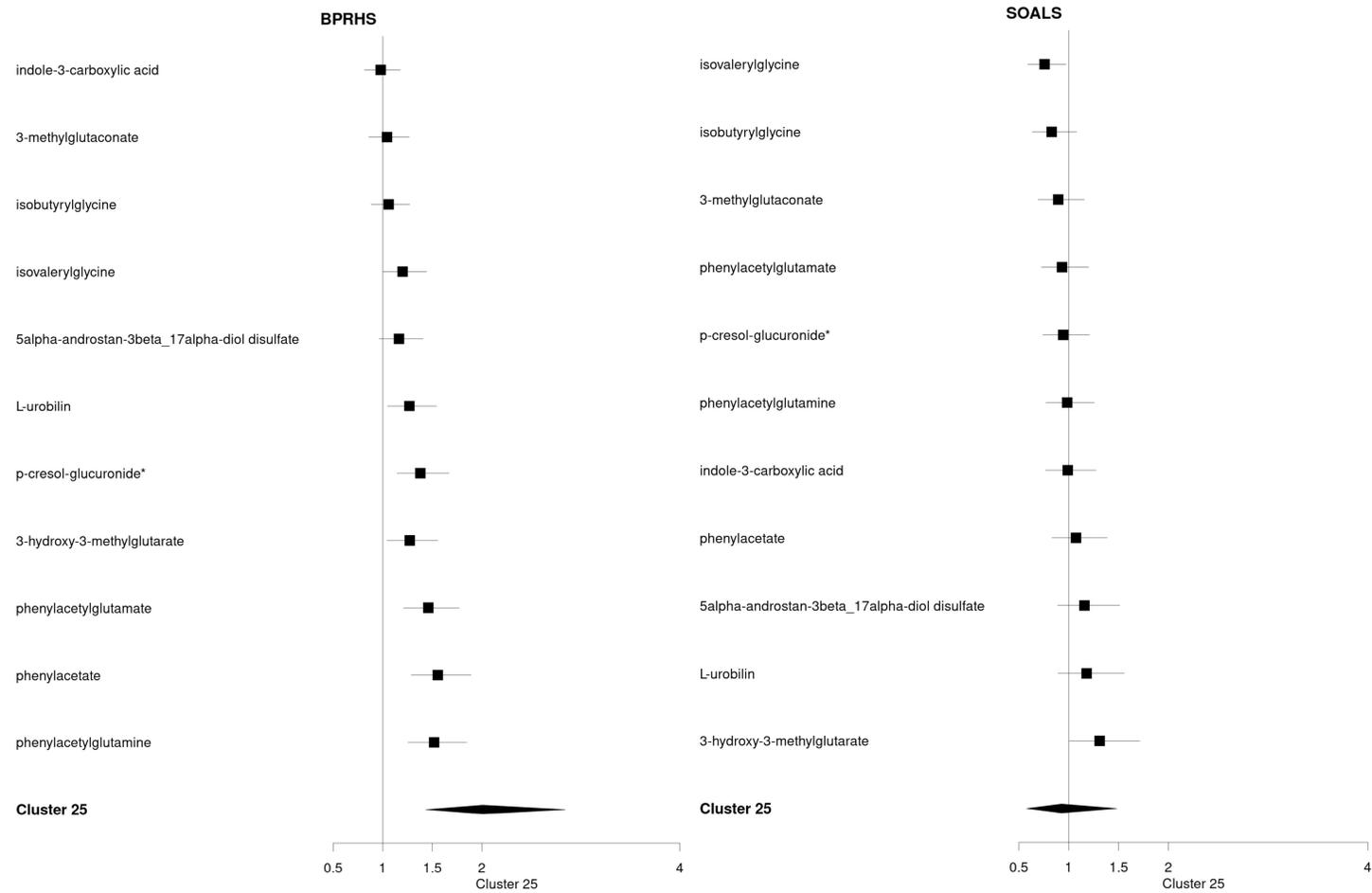


F)





H)



I)