Metabolite Signatures of Metabolic Risk Factors and their Longitudinal Changes

Xiaoyan Yin, PhD,[1,2], Subha Subramanian, BS,[1,3], Christine M. Willinger, BS,[1,3], George Chen, BS,[1,3], Peter Juhasz, PhD,[4], Paul Courchesne, MBA[1,3]; Brian H. Chen, PhD, MPH,[1,3], Xiaohang Li, PhD,[4], Shih-Jen Hwang, PhD[1,3], Caroline S. Fox, MD, MPH,[1,3,5]*, Christopher J. O'Donnell, MD, MPH[1,3,6], Pieter Muntendam, MD,[4], Valentin Fuster, MD[7,8], Ivana Bobeldijk-Pastorova, PhD,[9], Silvia C. Sookoian, MD, PhD,[10], Carlos J. Pirola, PD[10], Neal Gordon, MD,[4], Aram Adourian, PhD[4]*; Martin G. Larson, ScD[1,2]*; Daniel Levy, MD[1,3,11]*

[1] Framingham Heart Study, Framingham, MA; [2] Boston University Department of Mathematics and School of Public Health, Boston, MA; [3] Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD; [4] BG Medicine, Inc., Waltham, MA; [5] Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; [6] Department of Medicine, MA General Hospital, Harvard Medical School, Boston, MA; [7] Mt Sinai School of Medicine, NY, NY; [8] Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain; [9] TNO Triskelion BV, Inc., The Netherlands; [10] Institute of Medical Research A Lanari-IDIM, University of Buenos Aires - National Scientific and Technical Research Council (CONICET), Ciudad Autónoma de Buenos Aires, Argentina; [11] Boston University School of Medicine, Boston, MA; *Drs. Adourian, Larson, and Levy contributed equally to this manuscript.

Context: Metabolic dysregulation underlies key metabolic risk factors—obesity, dyslipidemia, and dysglycemia.

Objective: To uncover mechanistic links between metabolomic dysregulation and metabolic risk by testing metabolite associations with risk factors cross-sectionally and with risk factor changes over time.

Design: Cross-sectional—discovery samples (N=650; age=36–69 years) from the Framingham Heart Study (FHS) and replication samples (N=670; age=61–76 years) from the BioImage Study, both following a factorial design sampled from high versus low strata of body mass index, lipids, and glucose. Longitudinal—FHS participants (N=554) with 5–7 years of follow-up for risk factor changes.

Setting: Observational studies.

Participants: Cross-sectional samples with or without obesity, dysglycemia, and dyslipidemia, excluding prevalent cardiovascular disease and diabetes or dyslipidemia treatment. Age and sex matched by group.

Interventions: None.

Main Outcome Measure(s): Gas chromatography-mass spectrometry detected 119 plasma metabolites. Cross-sectional associations with obesity, dyslipidemia, and dysglycemia were tested in discovery, with external replication of 37 metabolites. Single- and multi-metabolite markers were tested for association with longitudinal changes in risk factors.

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received June 12, 2015. Accepted February 18, 2016.

Abbreviations: : Full Name Abbreviation; Alanine transaminase ALT; Aspartate transaminase AST; Body mass index BMI; Branched-chain amino acid BCAA; Cardiovascular disease CVD; False discovery rate FDR; Framingham Heart Study FHS; Gas chromatography-mass spectrometry GC-MS; HDL-cholesterol HDL-C; Triglyceride TG

Results: Cross-sectional metabolite associations were identified with obesity (n=26), dyslipidemia (n=21), and dysglycemia (n=11) in discovery. Glutamic acid, lactic acid, and sitosterol associated with all three risk factors in meta-analysis ($p<4.5x10-4$). Metabolites associated with longitudinal risk factor changes were enriched for bioactive lipids. Multi-metabolite panels explained 2.5–15.3% of longitudinal changes in metabolic traits.

Conclusions: Cross-sectional results implicated dysregulated glutamate cycling and amino acid metabolism in metabolic risk. Certain bioactive lipids were associated with risk factors cross-sectionally and over time, suggesting their upstream role in risk factor progression. Functional studies are needed to validate findings and facilitate translation into treatments or preventive measures.

Cardiovascular disease (CVD) is the foremost cause of death in the developed world and is emerging as a leading cause of death in the developing world (1, 2). The simultaneous presence of multiple CVD risk factors is used to diagnose metabolic syndrome (3–5). Metabolic syndrome is associated with cardiovascular morbidity and mortality, and current estimates place its prevalence in U.S. adults at nearly one-third (6, 7). Thanks to advances in mass spectrometry, population-based studies are now able to identify metabolite associations with metabolic risk factors in plasma and other biosamples (8–12). We hypothesized that metabolic pathway disturbances underlie the major components of metabolic syndrome. We aimed to identify plasma metabolite signatures of obesity, dyslipidemia, and dysglycemia cross-sectionally and of serial changes in the same metabolic traits.

To that end, we designed an efficient 2 × 2x2 factorial study to evaluate plasma metabolites as biomarkers of obesity, dyslipidemia, and dysglycemia. Metabolomic profiling was conducted using gas chromatography-mass spectrometry (GC-MS) on plasma samples from 650 Framingham Heart Study (FHS) participants. Independent external replication was performed on biosamples from 670 participants in the BioImage Study. We also evaluated associations of our complete metabolite set with longitudinal changes in body mass index (BMI), lipids [triglyceride (TG) and HDL-cholesterol (HDL-C)], and fasting glucose during 5–7 years of follow-up in FHS participants.

Materials and Methods

Study Samples. The FHS is an observational study that began recruitment of three generations of participants in 1948, 1971, and 2002, respectively (13–15). In discovery, we sampled attendees at Offspring cohort Examination 8 (2005–2008; n = 3,021) and Third Generation cohort Examination 1 (2002–2005; n = 4,095). Exclusion criteria included off-site examination (n = 107), no blood specimen (n = 126), lack of unrestricted consent (n = 335), prevalent CVD at baseline (n = 216), pharmacological treatment for diabetes (n = 257) or dyslipidemia (n = 1,201), missing BMI data (n = 4), smoking at baseline examination (n = 772), or age < 50 or > 79 years for the Offspring cohort (n = 123) or < 25 or > 59 years for the Third Generation cohort (n = 183). The eligible sample was 3792 individuals. Replication samples were selected from individuals enrolled in the BioImage Study, an ongoing prospective observational study of subjects free of prevalent CVD (ClinicalTrials.gov Identifier: NCT00738725) (16). In June 2009, enrollment was complete for 6822 subjects from the metropolitan areas of Chicago, IL, and Fort Lauderdale, FL. Study participants included men ages 55–80 and women ages 60–80; exclusion criteria included self-reported use of lipid- or glucose-lowering medications. Nonfasting plasma and serum samples from FHS and BioImage were collected according to standard operating procedures, aliquoted immediately after centrifugation, and stored at –80°C with < 3 freeze-thaw cycles prior to assay.

Clinical Measures. Metabolic syndrome risk factor definitions were similar to those of the National Cholesterol Education Program's Adult Treatment Panel III report (5). Obesity was defined as BMI ≥ 30 kg/m^2. Obesity was selected as a stratifying variable rather than waist circumference because it was available in both discovery and replication, and the correlation between BMI and waist circumference was extremely high (in pooled Offspring and Third Generation cohort participants, r = 0.92 [n = 7,001], adjusted for age, sex, and cohort). Dyslipidemia was defined as plasma TG ≥ 150 mg/dL or low plasma HDL-C (<40 mg/dL in men or < 50 mg/dL in women). Dysglycemia was defined as fasting blood glucose ≥ 100 mg/dL. Plasma aspartate transaminase (AST) and alanine transaminase (ALT) levels were measured in FHS and BioImage blood samples using the kinetic method (Beckman Liquid-Stat Reagent Kit).

Study Design. In discovery and replication, we implemented an efficient 2 × 2x2 factorial design by sampling for elevated vs nonelevated levels of BMI, lipids, and glucose (Table 1). Individuals were stratified to high vs. normal BMI, dyslipidemia vs. normal lipids, and high vs. normal fasting glucose. This stratification produced eight groups of equal size, each with a unique combination of presence and absence of the three metabolic risk factors, that permitted analyses of individual risk factor axes (ie, BMI, dysglycemia, or dyslipidemia) using data from all participants. The 2 × 2x2 factorial design ensured that when selecting any of the three metabolic risk factors as a discriminatory factor for analysis, the distribution of the other two factors remained balanced across the eight groups. We aimed to sample 80–84 individuals for each group, while balancing age and sex (and cohort in FHS). The final sample size for the FHS was 650 participants; for BioImage, the sample size was 670. The longitudinal sample included members of the discovery sample who attended the follow-up examination (FHS Offspring cohort Examination 9 [2011–2013] or Third Generation cohort Exami-

Table 1. Study Design and Clinical Characteristics of Discovery and Replication Samples

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Obesity*	-	-	-	-	+	+	+	+
Dysglycemia†	-	+	-	+	-	+	-	+
Dyslipidemia‡	-	-	+	+	-	-	+	+
Framingham Heart Study (Discovery)								
Sample size	84	83	84	84	82	84	74	83
Age (years)	52 ± 14	54 ± 13	52 ± 14	54 ± 15	53 ± 12	54 ± 11	50 ± 14	55 ± 13
Women (%)	50	49	50	50	50	50	57	49
BMI (kg/m^2)	24.5 ± 2.8	25.0 ± 2.7	25.9 ± 2.7	26.6 ± 2.3	32.4 ± 2.3	35.9 ± 5.1	35.4 ± 4.8	34.9 ± 4.9
Glucose (mg/dL)	92 ± 5	107 ± 15	92 ± 5	107 ± 7	93 ± 4	109 ± 7	93 ± 5	122 ± 37
HDL-C (mg/dL)	65 ± 18	62 ± 15	45 ± 11	44 ± 14	61 ± 13	58 ± 13	42 ± 10	42 ± 10
TG (mg/dL)	81 ± 26	84 ± 26	167 ± 115	175 ± 91	92 ± 28	95 ± 29	165 ± 77	202 ± 148
BioImage Study (Replication)								
Sample size (#)	83	84	83	84	84	84	84	84
Age (years)	69 ± 6	69 ± 6	69 ± 6	70 ± 6	67 ± 6	68 ± 6	68 ± 6	68 ± 6
Women (%)	49	50	51	50	50	50	50	50
BMI (kg/m^2)	25.6 ± 2.6	25.9 ± 2.6	25.8 ± 2.7	26.9 ± 2.3	33.6 ± 3.5	35.4 ± 4.8	33.8 ± 3.3	34.1 ± 3.8
Glucose (mg/dL)	89 ± 8	110 ± 13	90 ± 9	114 ± 19	90 ± 8	112 ± 12	92 ± 6	116 ± 19
HDL-C (mg/dL)	63 ± 14	64 ± 15	54 ± 17	50 ± 15	59 ± 13	59 ± 14	46 ± 11	48 ± 13
TG (mg/dL)	104 ± 27	106 ± 35	221 ± 119	238 ± 118	105 ± 33	110 ± 29	224 ± 92	223 ± 107

(+) and (-) represent the presence and absence, respectively, of obesity, dysglycemia, and dyslipidemia in each of the eight metabolic risk factor groups.

*Obesity was defined as BMI ≥30 kg/m^2.

†Dysglycemia was defined as a fasting blood glucose level ≥100 mg/dL.

‡Dyslipidemia was defined as either (1) TG level ≥150 mg/dL or (2) HDL-C level <40 mg/dL in men or <50 mg/dL in women.

nation 2 [2008–2010]). Participants without follow-up data for the metabolic traits of interest (BMI, TG, HDL-C, and glucose) were excluded. The final sample size for the longitudinal analysis was 554 participants. For individuals who reported treatment for dyslipidemia (n = 133) or diabetes (n = 30) at follow-up, values were imputed empirically to adjust for treatment effect (described below).

Metabolomic Profiling. GC-MS profiling of plasma samples was conducted at the TNO (Zeist, The Netherlands) with randomized cohorts and blinded analysis. A list of analyte targets was created from a preliminary profiling experiment before target peaks were measured in experimental samples. For GC-MS analysis, 100 μL samples were extracted with methanol and dried under a stream of nitrogen. For two-step derivatization, oximation was completed with ethoxamine in pyridine, and then trimethylsilylation was completed by adding 100 μL of N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA). To each sample, a cocktail of internal standards was added for normalization: d3-leucine, (13)C5-ribose, d3-glutamic acid, d5-phenylalanine, d4-alanine, d4-cholic acid, d4-citric acid, trifluoroacetylanthracene, and dicyclohexyl phthalate. The samples were injected into an Agilent 7890 N gas chromatograph connected to an Agilent 5975 mass spectrometer (Santa Clara, CA, USA). See Pellis et al for more detail (17). The analytical runs were arranged into 27 randomized batches, each containing 24 primary and three quality control (QC) samples. Quality control (QC) samples were generated by pooling plasma specimens from 15 healthy donors (7 male; 8 female) following standard collection protocols. Injections were made in duplicate for experimental and QC samples. ChemStation software (Version E02.00.493, Agilent Technologies) was used for data processing (18). A target table was constructed using an in-house library containing the mass spectra and retention times of > 600 reference metabolites (authentic standards), >100 annotated metabolites (spectral match with the NIST library or spectra of silylated compounds published in literature), and > 200 unknown metabolites found in serum or plasma in previous studies. Compounds were detected either underivatized, oximated and silylated, or silylated. Some lipid compounds, eg, sphingomyelins, were detected as a thermal degradation product by losing the phosphocholine group in the injector (19). Metabolites (known or unknown) specific for this study were added to the target table by searching the first three batches of samples for peaks present but not accounted for by the preselected target table. Peaks present in blanks and chemicals were excluded from the target table based on analysis of blank solvents and blank derivatizations. When one metabolite yielded ≥ 2 derivatization products, the most reliable was selected (based on specificity of ions, lack of coelution, and intensity) and the others excluded. A total of 175 targets were detected in this study and quantified in all samples by reconstructing an ion chromatogram of a specific mass from the mass spectrum of the target. The quantification for all targets was manually checked by visual control and peak integration was corrected as needed. Components detected in the GC-MS runs were quantified through peak areas normalized to an internal standard peak area. The internal standard was selected individually for each component: the standard yielding the smallest variability observed in QC samples. To compensate for different batch medians and intra-batch linear trends, batch and trend corrections were applied as needed. These methods are described in detail by van der Kloet et al (20) Normalized peak areas representing relative concentrations of the detected metabolites were used for statistical data analysis without being converting into absolute concentrations.

Statistical methods. Statistical analyses were performed using SAS software version 9.2 for Windows (SAS Institute Inc., Cary, NC, USA). We calculated rank normal scores of each metabolite using the Blom method (21). In discovery and replication, general linear mixed models were used to identify metabolites that differed across all eight metabolic groups or with individual risk factors (SAS procedure GLIMMIX). Metabolites were treated as outcome variables and metabolic group was a categorical predictor. Covariates included age, sex, sex*age, cohort, and batch effect, with additional adjustment in FHS for cohort, cohort*sex,

and cohort*age. The effect of each main risk factor was studied using models that adjusted for the other two risk factors. Analogous methods were applied in replication, except cohort adjustments were not necessary. False discovery rate (FDR) was used to select metabolites for replication (SAS procedure MULTTEST). Metabolites for replication were chosen as follows: (1) FDR ≤ 0.01 in the eight-group comparison and (2) top 10 metabolites from each main effect analysis, prioritized by discovery FDR. Metabolites that were highly correlated with a previously selected metabolite were excluded ($r \geq 0.7$). Meta-analyses were carried out using a fixed-effect approach with inverse-variance weighting (22). Specifically, the estimated combined effect was the inverse variance weighted average of those in discovery and replication. Associations of AST and ALT with metabolites were conducted using general linear mixed models adjusting for age, sex, alcohol consumption (average drinks per week), and metabolic risk factors (obesity, dyslipidemia, and dysglycemia), with additional adjustment in FHS for cohort, cohort*sex, and cohort*age. Plasma levels of ALT and AST were natural log transformed and standardized, and outlier values were excluded.

In our longitudinal study, we used general linear models (SAS procedure GLM). Each metabolite was treated as the predictor and change from baseline for each continuous metabolic trait (BMI, TG, HDL-C, and glucose) was the outcome. Covariates included age, sex, cohort, and baseline values of the metabolic traits. For participants on lipid or diabetes medications at follow-up, values were imputed empirically. Specifically, HDL-C was reduced by 10% and TG elevated by 20%, and glucose was increased by 10% or to 126 mg/dL (whichever was higher), respectively. As a sensitivity analysis, metabolite associations were retested with treated individuals excluded. Multimetabolite associations with longitudinal change were assessed using stepwise regression with forward selection. For each trait, candidate lipids were considered those with FDR < 0.25 in the single-marker analysis, and a p-value threshold corresponding to FDR < 0.25 was used for stepwise selection [at step i, $p<(0.25*i)/119$]. Aggregate partial r^2 was calculated for multimetabolite panels after forward selection was complete. Using this approach, highly correlated markers could not enter the model.

The Bonferroni method was used to correct for multiple testing ($P < .05$/# of measured metabolites). For discovery and longitudinal analyses, $P < 4.2 \times 10^{-4}$ was used (0.05/119). The replication threshold was $P < 1.4 \times 10^{-3}$ (0.05/37). For meta-analysis, a more stringent cut-off was implemented by including a term for the number of metabolic risk factors evaluated (n = 3): $P < .05/(37*3) = 4.5 \times 10^{-4}$.

Results

Study Sample Characteristics. Clinical characteristics of study participants in discovery and replication are shown in Table 1. Mean age and percentage of women were constant across groups. BioImage participants were slightly older than FHS participants (16). BMI, lipids, and glucose levels differed systematically across metabolic groups per study design. Supplemental Table 1 reports clinical characteristics of the longitudinal samples, and the distribution of changes from baseline for each metabolic trait is provided in Supplemental Figure 1A-D. The mean changes from baseline to follow-up were 0.6 ± 2.6 kg/m^2 for BMI, -0.5 ± 84.6 mg/dL for TG, 3.6 ± 10.7 mg/dL for HDL-C, and 0.5 ± 20.4 mg/dL for glucose.

Metabolite Associations in Discovery. In discovery, 149 metabolites were detected by GC-MS. Of these, 26 were unknown, one was a duplicate, and two were exogenous compounds; these were excluded prior to analysis. The median coefficient of variation in analytical replicates was 9.3% (25th percentile = 6.4%; 75th percentile = 10.9%). One analyte, cysteine, was excluded due to poor measurement quality. The final metabolite set consisted of 119 metabolites, consisting of alcohols (4.2%), amino acids (23.5%), lipids (38.7%), organic acids (13.4%), sterols (2.5%), sugars (9.2%), and others (8.4%). See Supplemental Table 2 for an annotated list of these metabolites.

We identified cross-sectional associations of metabolites from multiple classes with obesity (n = 26), dyslipidemia (n = 22), and dysglycemia (n = 11) at $P < 4.2 \times 10^{-4}$ (Tables 2-4). The top three metabolites associated with obesity were glutamic acid ($P = 2.1 \times 10^{-18}$), sitosterol ($P = 1.6 \times 10^{-10}$), and uric acid ($P = 3.0 \times 10^{-10}$) (Table 2). The strongest markers for dyslipidemia were sphingomyelins (SMs), including SM (d18:1/24:0) and SM (d18:1/24:1) ($P = 1.4 \times 10^{-27}$), SM (d18:2/24:0) and SM (d18:2/24:1) ($P = 4.5 \times 10^{-20}$), and SM (d18:1/16:0) ($P = 7.4 \times 10^{-11}$) (Table 3). For dysglycemia, the top three metabolites were glucose ($P = 1.4 \times 10^{-42}$), fructose ($P = 2.9 \times 10^{-7}$), and 2-hydroxybutanoic acid ($P = 5.5 \times 10^{-7}$) (Table 4). Glutamic acid and lactic acid were positively associated with all three risk factors. Several metabolites were associated with multiple risk factors, including direct associations of branched chain amino acids (BCAAs) and inverse associations of sitosterol and campesterol with obesity and dyslipidemia (Supplemental Table 3). Forty metabolites differed across the eight metabolic groups at $P < 4.2 \times 10^{-4}$ (Supplemental Table 2).

Metabolite Associations in Replication and Meta-Analysis. Thirty-eight metabolites met the replication criterion, of which 37 were successfully measured in BioImage samples. Replication results for metabolite associations with obesity (n = 10), dyslipidemia (n = 12), and dysglycemia (n = 7) at $P < 1.4 \times 10^{-3}$ are provided alongside their corresponding discovery results in Tables 2–4. We compared estimated effect sizes between discovery and replication phases for each metabolite and found high degrees of similarity (Supplemental Figure 2A-C). All metabolites in replication showed consistent directionality of associations. Correlations between metabolites in both discovery and replication phases are shown in Sup-

Table 2. Cross-Sectional Metabolite Associations with Obesity in Discovery, Replication, and Meta-Analysis

Metabolite	Discovery (FHS)			Replication (BioImage)			Meta-Analysis		
	β*	SE.	P-value†	β*	SE.	P-value‡	β*	SE.	P-value§
Glutamic acid	0.62	0.07	2.1×10^{-18}	0.36	0.07	2.3×10^{-7}	0.49	0.05	5.3×10^{-24}
Mannose	0.43	0.08	1.5×10^{-8}	0.51	0.07	2.5×10^{-12}	0.47	0.05	7.0×10^{-20}
Sitosterol	−0.48	0.07	1.6×10^{-10}	−0.34	0.07	2.9×10^{-6}	−0.41	0.05	2.2×10^{-15}
Uric acid	0.43	0.07	3.0×10^{-10}	0.31	0.07	1.8×10^{-5}	0.38	0.05	2.4×10^{-14}
Glycerol	0.36	0.06	1.1×10^{-8}	0.25	0.07	2.2×10^{-4}	0.31	0.05	1.3×10^{-11}
Glycine	−0.34	0.07	7.7×10^{-7}	−0.30	0.07	2.9×10^{-5}	−0.32	0.05	7.4×10^{-11}
2-Ketoglutaric acid	0.37	0.07	5.7×10^{-7}	0.30	0.07	8.2×10^{-5}	0.33	0.05	1.8×10^{-10}
Asparagine	−0.28	0.07	4.0×10^{-5}	−0.33	0.07	3.7×10^{-6}	−0.30	0.05	5.3×10^{-10}
Indole-3-propionic acid	−0.33	0.08	1.8×10^{-5}	−0.25	0.08	1.1×10^{-3}	−0.29	0.05	8.0×10^{-8}
Inositol	−0.30	0.07	2.2×10^{-5}	−0.23	0.07	1.9×10^{-3}	−0.27	0.05	1.5×10^{-7}
SM (d18:2/16:0)	0.30	0.07	4.8×10^{-6}	0.15	0.07	0.02	0.23	0.05	8.8×10^{-7}
4-Hydroxyglutamate semialdehyde	−0.21	0.07	4.6×10^{-3}	−0.29	0.07	7.2×10^{-5}	−0.25	0.05	1.3×10^{-6}
2-Hydroxybutanoic acid	0.31	0.08	3.2×10^{-5}	0.16	0.08	0.04	0.24	0.05	7.8×10^{-6}
LPA 16:0	−0.31	0.07	2.7×10^{-5}	−0.14	0.07	0.05	−0.23	0.05	1.3×10^{-5}
Leucine	0.30	0.07	1.5×10^{-5}	0.13	0.07	0.09	0.22	0.05	1.3×10^{-5}
Phenylalanine	0.32	0.07	1.9×10^{-5}	0.14	0.08	0.08	0.23	0.05	1.5×10^{-5}
Lactic acid	0.34	0.07	6.3×10^{-6}	0.09	0.07	0.21	0.22	0.05	4.0×10^{-5}
Aminomalonic acid	−0.29	0.07	4.6×10^{-5}	−0.10	0.07	0.12	−0.19	0.05	8.6×10^{-5}
Sn-Glycerol-3-phosphate	−0.25	0.07	6.0×10^{-4}	−0.15	0.07	0.04	−0.20	0.05	9.2×10^{-5}
Campesterol	−0.42	0.08	3.3×10^{-8}	-	-	-	-	-	-
Valine	0.35	0.07	1.2×10^{-6}	-	-	-	-	-	-
LPC 18:2	−0.34	0.07	1.7×10^{-6}	-	-	-	-	-	-
SM (d18:2/18:0)	0.32	0.07	7.5×10^{-6}	-	-	-	-	-	-
Pyruvic acid	0.34	0.08	8.7×10^{-6}	-	-	-	-	-	-
Creatinine	0.32	0.08	3.0×10^{-5}	-	-	-	-	-	-
SM (d16:1/18:0)	0.29	0.07	1.1×10^{-4}	-	-	-	-	-	-
Isoleucine	0.25	0.07	2.5×10^{-4}	-	-	-	-	-	-
Tyrosine	0.27	0.07	3.1×10^{-4}	-	-	-	-	-	-

(-) represents the absence of replication and meta-analysis data for a metabolite.

*Estimated β coefficients represent the mean differences in standardized metabolite measures between participants with and without obesity.

†The discovery p-value threshold was $P < 4.2\text{Xx}10^{-4}$ (0.05/119).

‡The replication p-value threshold was $P < 1.4 \times 10^{-3}$ (0.05/37).

§The meta-analysis p-value threshold was $P < 4.5 \times 10^{-4}$ [0.05/(37*3)]. All significant markers in meta-analysis are shown in order of ascending p-value.

Abbreviations: SM = sphingomyelin; LPA = lysophosphatidic acid; LPC = lysophosphatidylcholine. See Supplementary Table 2 for complete nomenclature.

plemental Figures 3–4. In meta-analysis of results from discovery and replication, 32 markers were associated with one or more risk factors, 16 were associated with two or more risk factors, and three (glutamic acid, lactic acid, and sitosterol) were associated with all three risk factors (Tables 2–4).

Transaminase Associations. Because the liver transaminases ALT and AST are robustly associated with metabolic syndrome and regulate several of our identified metabolites, (23–26) we examined metabolite associations with ALT and AST. Glutamic acid and 2-ketoglutaric acid were significantly associated with ALT and AST in discovery and replication (Supplemental Table 4).

Metabolite Associations with Longitudinal Changes in Risk Factors. We identified metabolites associated with serial change in BMI (n = 1), TG (n = 5), HDL-C (n = 2), and glucose (n = 1) at $P < 4.2 \times 10^{-4}$ (Table 5). All metabolites associated with longitudinal changes in risk factors were also associated with the same metabolic risk factor cross-sectionally ($P < 4.2 \times 10^{-4}$ in discovery), with consistent directionality of association. Following exclusion of treated individuals, the associations of two longitudinal markers were attenuated, but the effect sizes and directionalities were unchanged (Supplemental Table 5).

Multimetabolite panels were associated with longitudinal change in each metabolic trait (Table 6). For serial change in BMI, the sole marker was LPA 16:0, which explained 2.5% of change in BMI vs 2.1% explained by baseline BMI). Six metabolites were selected in the TG model: SM (d18:1/16:0), quinic acid, mannose, uric acid, phosphate, and 1-methylhistidine. In sum, these markers explained 15.3% of change in TG; baseline TG explained 23.2% of change. The metabolite panel for HDL-C consisted of SM (d18:2/20:0) and SM (d18:1/22:0), which explained nearly twice as much change in HDL-C as did baseline HDL-C (6.2% vs 3.8%). 1,2-diglyceride and 3-methyl-2-oxovaleric acid together accounted for 4.4% of change in glucose, relative to 21.6% explained by baseline glucose.

Discussion

We measured circulating levels of 119 metabolites and identified metabolomic signatures of key metabolic risk

Table 3. Cross-Sectional Metabolite Associations with Dyslipidemia in Discovery, Replication, and Meta-Analysis

Metabolite	Discovery (FHS)			Replication (BioImage)			Meta-Analysis		
	β*	SE.	P-value†	β*	SE.	P-value‡	β*	SE.	P-value§
SM (d18:1/24:0) & SM(d18:1/24:1)	−0.73	0.06	1.4×10^{-27}	−0.44	0.07	4.2×10^{-11}	−0.59	0.05	6.2×10^{-38}
Glutamic acid	0.45	0.07	6.9×10^{-11}	0.50	0.07	6.9×10^{-13}	0.48	0.05	5.3×10^{-23}
Myo-inositol 1,2-cyclic phosphate	0.33	0.07	9.0×10^{-6}	0.53	0.07	2.7×10^{-15}	0.44	0.05	1.6×10^{-19}
SM (d18:1/17:0)	−0.46	0.07	4.9×10^{-10}	−0.33	0.07	4.0×10^{-6}	−0.40	0.05	9.2×10^{-15}
SM (d18:0/16:0)	−0.45	0.07	1.9×10^{-10}	−0.24	0.07	1.1×10^{-3}	−0.35	0.05	3.5×10^{-12}
Leucine	0.29	0.07	2.8×10^{-5}	0.41	0.07	2.6×10^{-8}	0.35	0.05	4.1×10^{-12}
Alanine	0.29	0.08	1.9×10^{-4}	0.44	0.07	4.3×10^{-9}	0.36	0.05	5.7×10^{-12}
Lactic acid	0.37	0.07	6.9×10^{-7}	0.33	0.07	8.2×10^{-6}	0.35	0.05	1.7×10^{-11}
Erythronic acid	0.22	0.06	5.2×10^{-4}	0.43	0.07	2.0×10^{-9}	0.32	0.05	2.4×10^{-11}
Sitosterol	−0.27	0.07	2.9×10^{-4}	−0.37	0.07	3.9×10^{-7}	−0.32	0.05	5.1×10^{-10}
Maltose	0.26	0.07	5.3×10^{-4}	0.29	0.08	1.7×10^{-4}	0.27	0.05	2.9×10^{-7}
Serine	−0.31	0.07	2.8×10^{-5}	−0.22	0.07	3.6×10^{-3}	−0.26	0.05	4.4×10^{-7}
Uric acid	0.24	0.07	4.5×10^{-4}	0.22	0.07	2.3×10^{-3}	0.23	0.05	3.0×10^{-6}
SM (d18:2/16:0)	−0.30	0.07	7.8×10^{-6}	−0.12	0.06	0.07	−0.21	0.05	7.5×10^{-6}
SM (d16:1/24:1)	−0.46	0.07	4.6×10^{-10}	0.00	0.07	0.98	−0.23	0.05	8.5×10^{-6}
3-Methyl-2-oxovaleric acid	0.20	0.07	1.6×10^{-3}	0.23	0.07	1.6×10^{-5}	0.21	0.05	1.6×10^{-5}
SM(d18:2/24:0) and SM(d18:2/24:1)	−0.67	0.07	4.5×10^{-20}	-	-	-	-	-	-
SM (d18:1/16:0)	−0.49	0.07	7.4×10^{-11}	-	-	-	-	-	-
SM (d18:2/20:0)	−0.35	0.07	1.5×10^{-7}	-	-	-	-	-	-
LPC 18:2	−0.34	0.07	1.1×10^{-6}	-	-	-	-	-	-
Isoleucine	0.33	0.07	2.1×10^{-6}	-	-	-	-	-	-
SM (d17:1/16:0)	−0.31	0.07	1.5×10^{-5}	-	-	-	-	-	-
Campesterol	−0.29	0.07	1.3×10^{-4}	-	-	-	-	-	-
SM (d18:1/23:0)	−0.28	0.07	1.4×10^{-4}	-	-	-	-	-	-
Valine	0.27	0.07	2.2×10^{-4}	-	-	-	-	-	-
Quinic acid	−0.27	0.08	4.1×10^{-4}	-	-	-	-	-	-

(-) represents the absence of replication and meta-analysis data for a metabolite.

*Estimated β coefficients represent the mean differences in standardized metabolite measures between participants with and without dyslipidemia.

†The discovery p-value threshold was $P < 4.2Xx10^{-4}$ (0.05/119).

‡The replication p-value threshold was $P < 1.4 \times 10^{-3}$ (0.05/37).

§The meta-analysis p-value threshold was $P < 4.5 \times 10^{-4}$ [0.05/(37*3)]. All significant markers in meta-analysis are shown in order of ascending p-value.

Abbreviations: SM = sphingomyelin; LPC = lysophosphatidylcholine. See Supplementary Table 2 for complete nomenclature.

Table 4. Cross-Sectional Metabolite Associations with Dysglycemia in Discovery, Replication, and Meta-Analysis

Metabolite	Discovery (FHS)			Replication (BioImage)			Meta-Analysis		
	β*	SE.	P-value†	β*	SE.	P-value‡	β*	SE.	P-value§
Glucose	0.95	0.06	1.4×10^{-42}	0.95	0.06	9.0×10^{-47}	0.95	0.04	2.1×10^{-102}
Mannose	0.25	0.08	8.5×10^{-4}	0.41	0.07	1.5×10^{-8}	0.33	0.05	9.6×10^{-11}
Lactic acid	0.31	0.07	4.3×10^{-5}	0.37	0.07	9.2×10^{-7}	0.34	0.05	1.4×10^{-10}
Glutamic acid	0.30	0.07	1.1×10^{-5}	0.31	0.07	8.2×10^{-6}	0.31	0.05	2.7×10^{-10}
2-Hydroxybutanoic acid	0.38	0.08	5.5×10^{-7}	0.23	0.08	3.2×10^{-3}	0.31	0.05	1.3×10^{-8}
3-Methyl-2-oxovaleric acid	0.24	0.07	4.8×10^{-4}	0.30	0.07	4.6×10^{-5}	0.27	0.05	7.6×10^{-8}
Alanine	0.26	0.08	8.8×10^{-4}	0.27	0.07	2.5×10^{-4}	0.26	0.05	6.6×10^{-7}
Fructose	0.39	0.07	2.9×10^{-7}	0.09	0.08	0.25	0.24	0.05	6.4×10^{-6}
Glycine	−0.23	0.07	7.9×10^{-4}	−0.20	0.07	4.4×10^{-3}	−0.22	0.05	1.0×10^{-5}
Aminomalonic acid	−0.35	0.07	1.1×10^{-6}	−0.09	0.07	0.18	−0.21	0.05	1.6×10^{-5}
2-Ketoglutaric acid	0.24	0.07	1.4×10^{-3}	0.21	0.07	5.4×10^{-3}	0.22	0.05	2.2×10^{-5}
SM (d18:1/24:0) & SM (d18:1/24:1)	−0.22	0.06	5.5×10^{-4}	−0.16	0.07	0.02	−0.19	0.05	2.9×10^{-5}
Malic acid	0.30	0.07	3.5×10^{-5}	0.13	0.07	0.09	0.22	0.05	3.3×10^{-5}
4-hydroxyglutamate semialdehyde	−0.28	0.07	1.1×10^{-4}	−0.14	0.07	0.05	−0.21	0.05	3.3×10^{-5}
SM (d18:1/17:0)	−0.26	0.07	4.2×10^{-4}	−0.16	0.07	0.03	−0.21	0.05	4.4×10^{-5}
Sitosterol	−0.16	0.07	0.03	−0.24	0.07	9.1×10^{-4}	−0.20	0.05	8.6×10^{-5}
Malic acid	0.30	0.07	3.5×10^{-5}	-	-	-	-	-	-
Pyruvic acid	0.30	0.08	8.9×10^{-5}	-	-	-	-	-	-
C18:0 FA (stearic acid)	0.28	0.08	3.2×10^{-4}	-	-	-	-	-	-
C16:0 FA (palmitic acid)	0.26	0.07	4.0×10^{-4}	-	-	-	-	-	-

(-) represents the absence of replication and meta-analysis data for a metabolite.

*Estimated β coefficients represent the mean differences in standardized metabolite measures between participants with and without dysglycemia.

†The discovery p-value threshold was $P < 4.2Xx10^{-4}$ (0.05/119).

‡The replication p-value threshold was $P < 1.4 \times 10^{-3}$ (0.05/37).

§The meta-analysis p-value threshold was $P < 4.5 \times 10^{-4}$ [0.05/(37*3)]. All significant markers in meta-analysis are shown in order of ascending p-value.

Abbreviations: SM = sphingomyelin; FA = fatty acid. See Supplementary Table 2 for complete nomenclature.

factors at baseline and of their longitudinal changes during 5–7 years of follow-up. Many top cross-sectional findings

Table 5. Metabolite Associations with Longitudinal Changes in Metabolic Traits

Change in Metabolic Trait	Metabolite	Estimated β*	SE.	P-value†	Partial R^2
Body mass index	LPA 16:0	−0.48	0.13	2.8×10^{-4}	0.025
Triglyceride	SM (d18:1/16:0)	−21.38	4.94	1.8×10^{-5}	0.035
	SM (d18:1/17:0)	−17.61	4.10	2.1×10^{-5}	0.035
	Quinic acid	−13.74	3.48	9.1×10^{-5}	0.030
	Sitosterol	−14.46	3.69	1.0×10^{-4}	0.029
	SM (d18:1/24:0) & SM (d18:1/24:1)	−17.83	4.92	3.2×10^{-4}	0.025
HDL cholesterol	SM (d18:2/20:0)	2.59	0.67	1.2×10^{-4}	0.029
	SM (d18:1/24:0) & SM (d18:1/24:1)	2.62	0.68	1.4×10^{-4}	0.028
Glucose	Glucose	5.31	1.05	5.2×10^{-7}	0.048

*Estimated β coefficients represent the mean differences in longitudinal change (native units) in the stated metabolic trait per one standard deviation difference in metabolite measures.

†The p-value threshold for significance ($P < 4.2 \times 10^{-4}$) was determined by the Bonferroni method (0.05/119 metabolites).

Abbreviations: LPA = lysophosphatidic acid; SM = sphingomyelin. See Supplementary Table 2 for complete nomenclature.

All markers of longitudinal change were also associated with the corresponding risk factor in cross-sectional analyses.

Table 6. Multi-Metabolite Associations with Longitudinal Changes in Metabolic Traits

Change in Metabolic Trait	Parameter	Estimated β*	SE.	P-value†	Partial R^2
Body mass index	Baseline BMI	−0.07	0.02	8×10^{-4}	0.021
	Baseline Covariates‡				0.070
	LPA 16:0	−0.48	0.13	3×10^{-4}	0.025
	Multi-Metabolite Panel				0.025
Triglyceride	Baseline TG	−0.57	0.05	$<1 \times 10^{-4}$	0.232
	Baseline Covariates‡				0.399
	SM (d18:1/16:0)	−25.25	5.18	$<1 \times 10^{-4}$	0.046
	Quinic acid	−12.02	3.34	$<4 \times 10^{-4}$	0.025
	Mannose	11.33	3.58	2×10^{-3}	0.020
	Uric acid	16.53	4.10	$<1 \times 10^{-4}$	0.032
	Phosphate	−10.71	3.80	5×10^{-3}	0.016
	1-Methylhistidine	−9.45	3.41	6×10^{-3}	0.015
	Multi-Metabolite Panel				0.153
HDL cholesterol	Baseline HDL-C	−0.16	0.04	$<1 \times 10^{-4}$	0.038
	Baseline Covariates‡				0.127
	SM (d18:2/20:0)	3.11	0.68	$<1 \times 10^{-4}$	0.040
	SM (d18:1/22:0)	−1.96	0.56	6×10^{-4}	0.023
	Multi-Metabolite Panel				0.062
Glucose	Baseline Glucose	−0.52	0.04	$<1 \times 10^{-4}$	0.216
	Baseline Covariates‡				0.275
	1,2-Diglyceride	−3.22	0.92	6×10^{-4}	0.023
	3-Methyl-2-oxovaleric acid	3.05	0.92	1×10^{-3}	0.021
	Multi-Metabolite Panel				0.044

*Estimated β coefficients represent the mean differences in longitudinal change (native units) in the stated metabolic trait per one standard deviation difference in metabolite measures.

†A dynamic p-value threshold for entrance into the model was defined based on FDR<0.25 [at step i, p<(0.25*i)/119].

‡Baseline covariates were as follows: age; sex; cohort; batch; and baseline values for total cholesterol, HDL cholesterol, triglyceride, glucose, and body mass index.

Abbreviations: LPA = lysophosphatidic acid; SM = sphingomyelin. See Supplementary Table 2 for complete nomenclature.

Metabolites also associated cross-sectionally with the corresponding metabolic risk factor are in **bold**.

in discovery were replicated in an independent external cohort, and meta-analysis identified three metabolites (glutamic acid, lactic acid, and sitosterol) associated with all three risk factors cross-sectionally. Our cross-sectional results identified risk factor associations with glutamic acid and other major amino acids—particularly branched chain and aromatic amino acids—that may contribute to metabolic risk through their involvement in transaminase-mediated pathways. Bioactive lipids also emerged as important markers of metabolic risk factors cross-sectionally and of their changes over time. Multimetabolite analyses demonstrated the predictive power of the top lipid markers, suggesting an upstream role for lipid dysregulation in metabolic risk factor progression.

Amino Acids. The top cross-sectional amino acid biomarker was glutamic acid (glutamate), which was directly associated with all three risk factors in discovery and replication. Glutamic acid is one of the most abundant amino acids and a key player in amino acid and carbohydrate metabolism. Glutamic acid has been linked to metabolic dysfunction, with increasing plasma levels identified in obesity, insulin resistance, and progression to type 2 diabetes in humans (8, 9, 27). The BCAAs—leucine, valine, and isoleucine—were directly associated with obesity and dyslipidemia in discovery. Aromatic amino acids phenylalanine and tyrosine were directly associated with obesity in discovery, while glycine and asparagine were inversely associated. Elevated levels of BCAAs and the same aro-

matic amino acids (as well as decreased levels of glycine) have been identified in recent cross-sectional studies as biomarkers of obesity, early insulin resistance, and risk of diabetes in humans (8, 10, 12, 27–29). A number of previous intervention studies, however, found that higher dietary BCAA intake reduces body fat and body weight in both humans and animal models, indicating that cross-sectional associations are not necessarily predictive of the results of intervention studies (30–33). We also examined levels of liver transaminases ALT and AST in association with our metabolites, and found strong associations of ALT and AST with glutamic acid and 2-ketoglutaric acid (its cognate α-keto acid) in discovery and replication. Transaminase-mediated reactions are key to amino acid catabolism and to the regulation of insulin secretion by pancreatic beta cells (34). We focused on ALT and AST because increased levels have been linked to obesity, diabetes, and CVD in large-scale epidemiological studies (23–26). Taken together, our results implicate dysregulated glutamate cycling via transaminase-mediated reactions, likely during metabolism of BCAAs and other amino acids, in the metabolic dysregulation underlying metabolic syndrome.

Lipids. One of the top cross-sectional lipid biomarkers was the lysophosphatidylcholine species 1-linoleoyl-sn-glycero-3-phosphocholine (LPC 18:2). LPC 18:2 showed inverse associations with obesity and dyslipidemia in discovery. Lysophosphatidylcholines are low-abundance lipid derivatives involved in signaling processes; they are theorized to act through G-protein-coupled receptors to trigger glucose-dependent insulin secretion and activate glucose uptake by adipocytes (35, 36). LPC 18:2 has recently emerged as a biomarker of metabolic dysfunction, showing robust inverse associations with insulin resistance, dysglycemia, type 2 diabetes, and obesity (12, 37, 38). LPC 18:2 was also reported as a predictive biomarker of incident coronary heart disease (CHD) (39). These findings suggest that reduced levels of LPC 18:2 may contribute to the metabolic dysfunction that promotes CVD. Lysophosphatidic acid, a lipid derivative of lysophosphatidylcholines, was our top longitudinal marker for BMI. LPA 16:0 (1-palmitoyl-L-α-lysophosphatidic acid) was inversely associated with BMI at baseline and with change in BMI in single marker and multimetabolite analyses, explaining 2.5% of longitudinal change in BMI. LPA 16:0 is the most common circulating form of lysophosphatidic acid, a lipid mediator with roles in essential cellular processes. One signaling pathway in which LPA is generated involves the action of autotaxin (ATX), which is highly expressed in adipocytes (40). ATX is elevated in the visceral fat of obese and insulin-resistant individuals (41, 42). ATX-knockout mice exhibit enhanced adiposity and reduced plasma LPA, suggesting that disruption of ATX/LPA signaling in adipose tissue contributes to obesity and that lowered plasma LPA is a biomarker of this process (43). The detailed mechanism by which ATX mediates LPA signaling in humans, however, remains unclear, and other LPA-producing pathways that are independent of ATX also exist (44, 45). Complicating this picture further, some studies have reported direct associations between LPA and obesity in mice (46).

Our results point to sphingomyelins as biomarkers of dyslipidemia. Tightly linked to sterol metabolism, sphingomyelins are bioactive sphingolipids that regulate cell growth and immune functions (47, 48).

Ten sphingomyelin species were inversely associated with dyslipidemia in discovery, and five of these were inversely associated with longitudinal change in TG and/or directly associated with change in HDL-C. In multimetabolite analyses, top sphingomyelin biomarkers explained 4.6% and 4.0% of longitudinal change in TG and HDL-C, respectively.

Two phytosterols, sitosterol and campesterol, were inversely associated with obesity and dysglycemia in discovery and replication, and sitosterol showed an inverse relation to TG in longitudinal analysis. These natural plant sterols, structurally similar to cholesterol, lower cholesterol absorption by displacing cholesterol from micelles in the intestinal tract, among other mechanisms (49). Increased intake of plant sterols is associated with cholesterol-lowering activity, (50, 51) and reduced levels of plasma phytosterols may serve as a proxy for altered cholesterol homeostasis (52).

Organic Acids. Lactic acid (lactate) was our top organic acid biomarker of all three risk factors in discovery. Lactic acid is an indicator of oxidative capacity, and hyperlactaemia is broadly characteristic of acute illness. In the clinical setting, lactic acid levels serve as a prognostic biomarker of a number of metabolic disorders (53). For example, increased plasma lactic acid has been found in the context of insulin resistance, diabetic ketoacidosis, and type 2 diabetes (54–56). Lactic acid has also been linked to incident diabetes in a large case-control study, suggesting a causal role of oxidative stress in metabolic disease (57). Particularly in the pediatric population, elevated lactate is associated with inborn errors of metabolism that affect key metabolic pathways involving pyruvate dehydrogenase, gluconeogenesis, and the tricarboxylic acid cycle (58). Multiple organic acids were directly associated with dysglycemia, including 2-hydroxybutanoic acid, malic acid, and pyruvic acid. Quinic acid was inversely associated with dyslipidemia cross-sec-

tionally and with change in TG in single and multimetabolite analyses. Quinic acid is a naturally occurring plant cyclic polyol found as chlorogenic acid in common dietary constituents, including tea, coffee, blueberries, apples, grapes, and cereals (59–61). A functional role for quinic acid as a prometabolite that enhances antioxidant production and DNA repair enhancement has been described (62). While human studies on chlorogenic acid have been inconsistent, this compound has been shown to improve lipid metabolism, BMI, and glycemic traits in mice (63–65).

Strengths and Limitations. Although our discovery and replication studies were modest in size (650 and 670, respectively), the cost-effective factorial design of independent discovery and replication phases enabled the identification of a robust set of metabolites associated with metabolic risk factors. Sample selection and statistical analysis for the replication phase were highly consistent with discovery, allowing us to validate many top markers. The use of an independent external, comparably-sized cohort for replication provided superior power to identify robust signals. The availability of longitudinal data enabled us to identify metabolites associated with risk factors at baseline and their serial changes and to estimate their predictive power in multimetabolite panels. It should be noted, however, that baseline covariates including age, sex, and baseline metabolic risk factor values explained a greater proportion of longitudinal changes in metabolic risk factors than multimetabolite panels alone.

Our study is limited by its use of predominantly middle-aged individuals of European ancestry in discovery and longitudinal analysis. We cannot generalize these findings to other ethnic groups or younger ages. The replication sample, however, was multiethnic (69% Caucasian; 13% Hispanic; 12% African American; 4% Asian; 2% other), and results were highly consistent with discovery. We were also limited by the number of detectable metabolites; newer platforms offer more comprehensive metabolomic screening. Of note, a recent population-based study showed that body weight change is associated with various lipoprotein classes, but our results lacked data on associations between lipoproteins and metabolic traits (66). In replication, restriction to uncorrelated metabolites limited our ability to analyze complete pathways.

Conclusion

We employed a factorial study design to identify cross-sectional metabolomic signatures of obesity, dyslipidemia, and dysglycemia, with independent external replication of top findings. Using follow-up data from the discovery sample, we identified metabolites associated with longitudinal changes in metabolic traits. Amino acid biomarkers of obesity and dysglycemia revealed a glutamic acid-enhanced signature of metabolic risk that may represent a systemic metabolic derangement involving BCAA metabolism and transaminase-mediated pathways. Organic acid intermediates were most highly associated with dyslipidemia, and may represent shifts in cellular metabolism in response to oxidative stress. Longitudinal results highlighted alterations in bioactive lipids, with lysophospholipid derivatives as biomarkers of BMI and its longitudinal change and sphingomyelins associated with dyslipidemia and serial changes in TG and HDL-C. Our multimetabolite analyses demonstrated a number of longitudinal associations, as multimetabolite panels explained a substantial proportion of longitudinal change in each metabolic trait compared with the corresponding baseline trait value. Our findings shed light on the profound breadth and depth of metabolomic derangements associated with metabolic syndrome risk factors. Follow-up studies guided by these results may help in the targeting of therapies to treat metabolic risk factors or prevent their emergence.

Acknowledgments

We would like to acknowledge the following additional people. From BG Medicine: Moira Lynch, Manuel Paniagua, and Ming Kuo; from TNO Triskelion BV: Leo van Stee and Bas Muilwijk.

Address all correspondence and requests for reprints to: Aram Adourian, PhD, BG Medicine, 610 Lincoln Street, Waltham, MA 02451 email: aadourian@bg-medicine.com, phone: 781–890-1199, fax: 781–895-1119; Martin G. Larson, ScD, Framingham Heart Study, 73 Mount Wayte Avenue, Framingham, MA 01702, email: mlarson@bu.edu phone: 508–935-3458, Daniel Levy, MD, Framingham Heart Study, 73 Mount Wayte Avenue, Framingham, MA 01702, email: levyd@nih.gov, phone: 508–935-3458, fax: 508–626-1262.

This work was supported by *Sources of Funding*: The Framingham Heart Study is funded by NIH contract N01-HC-25195; this research was also supported by the Division of Intramural Research of the National Heart, Lung, and Blood Institute. This study was made possible by a CRADA between BG Medicine, Inc., Boston University, and the National Heart Lung, and Blood Institute.

Disclosure Statement: P.J., P.M., N.G., and A.A. contributed to this project while affiliated with BG Medicine, Inc., a biomarker discovery company. I.B.P. is an employee of TNO Triskelion, Inc. C.P. serves as an advisor to and is a grant recipient from Merck, Sharp, and Dohme. All other authors have nothing to disclose.

References

1. **NHLBI. Fact Book for Fiscal Year 2002. US Dept of Health and Human Services, Public Health Service, NIH.**
2. **Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ.** Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet*. 2006;367:1747–57.
3. **Khot UN, Khot MB, Bajzer CT et al.** Prevalence of conventional risk factors in patients with coronary heart disease. *JAMA*. 2003;290: 898–904.
4. **Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB.** Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837–47.
5. **III ATP. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Circulation.**
6. **Isomaa B, Almgren P, Tuomi T et al.** Cardiovascular morbidity and mortality associated with the metabolic syndrome. *Diabetes Care*. 2001;24:683–9.
7. **Aguilar M, Bhuket T, Torres S, Liu B, Wong RJ.** Prevalence of the metabolic syndrome in the United States, 2003–2012. *Jama*. 2015; 313:1973–4.
8. **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.
9. **Cheng S, Rhee E, Larson M et al.** Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation*. 2012;125:2222–2231.
10. **Wang TJ, Larson MG, Vasan RS et al.** Metabolite profiles and the risk of developing diabetes. *Nat Med*. 2011;17:448–53.
11. **Menni C, Fauman E, Erte I et al.** Biomarkers for type 2 diabetes and impaired fasting glucose using a nontargeted metabolomics approach. *Diabetes*. 2013;62:4270–6.
12. **Szymanska E, Bouwman J, Strassburg K et al.** Gender-dependent associations of metabolite profiles and body fat distribution in a healthy population with central obesity: towards metabolomics diagnostics. *Omics*. 2012;16:652–67.
13. **Dawber TR, Meadors GF, Moore FE, Jr.** Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health*. 1951;41:279–81.
14. **Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP.** The Framingham Offspring Study. Design and preliminary data. *Prev Med*. 1975;4:518–25.
15. **Splansky GL, Corey D, Yang Q et al.** The Third Generation Cohort of the National Heart, *Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. Am J Epidemiol*. 2007;165:1328–35.
16. **Muntendam P, McCall C, Sanz J, Falk E, Fuster V. The BioImage Study: novel approaches to risk assessment in the primary prevention of atherosclerotic cardiovascular disease–study design and objectives. Am Heart J;160:49–57 e1.**
17. **Pellis L, van Erk M, van Ommen B et al.** Plasma metabolomics and proteomics profiling after a postprandial challenge reveal subtle diet effects on human metabolic status. *Metabolomics : Official journal of the Metabolomic Society*. 2012;8:347–359.
18. **Koek MM, van der Kloet FM, Kleemann R, Kooistra T, Verheij ER, Hankemeier T.** Semi-automated non-target processing in GC x GC-MS metabolomics analysis: applicability for biomedical studies. *Metabolomics*. 2011;7:1–14.
19. **Vieu C, Terce F, Chevy F et al.** Coupled assay of sphingomyelin and ceramide molecular species by gas liquid chromatography. *J Lipid Res*. 2002;43:510–22.
20. **van der Kloet FM, Bobeldijk I, Verheij ER, Jellema RH.** Analytical error reduction using single point calibration for accurate and precise metabolomic phenotyping. *J Proteome Res*. 2009;8:5132–41.
21. **Blom G. Statistical estimates adn transformed beta-variables. Wiley, 1958:176.**
22. **Greenland S RK, editor Meta-Analysis. 3rd ed: Lippincott Williams and Wilkins, 2008.**
23. **Goessling W, Massaro J, Vasan R, D'Agostino R, Ellison R, Fox C.** Aminotransferase levels and 20-year risk of metabolic syndrome, diabetes, and cardiovascular disease. *Gastroenterology*. 2008;135: 1935.
24. **Schindhelm R, Dekker J, Nijpels G et al.** Alanine aminotransferase and the 6-year risk of the metabolic syndrome in Caucasian men and women: the Hoorn Study. *Diabetic medicine : a journal of the British Diabetic Association*. 2007;24:430–435.
25. **Olynyk J, Knuiman M, Divitini M, Davis T, Beilby J, Hung J.** Serum alanine aminotransferase, metabolic syndrome, and cardiovascular disease in an Australian population. *The American journal of gastroenterology*. 2009;104:1715–1722.
26. **Lee K, Yang JH.** Which liver enzymes are better indicators of metabolic syndrome in adolescents: the fifth Korea national health and nutrition examination survey, 2010. *Metabolic syndrome and related disorders*. 2013;11:229–235.
27. **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;100:2014–2357.
28. **Kochhar S, Jacobs DM, Ramadan Z, Berruex F, Fuerholz A, Fay LB.** Probing gender-specific metabolism differences in humans by nuclear magnetic resonance-based metabonomics. *Analytical Biochemistry*. 2006;352:274–281.
29. **Batch BC, Shah SH, Newgard CB et al.** Branched chain amino acids are novel biomarkers for discrimination of metabolic wellness. *Metabolism*. 2013;62:961–9.
30. **Mourier A, Bigard AX, De kerviler E, Roger B, Legrand H, Guezennec CY.** Combined effects of caloric restriction and branched-chain amino acid supplementation on body composition and exercise performance in elite wrestlers. *Int J Sports Med*. 1997;18(1): 47–55.
31. **Layman DK. The role of leucine in weight loss diets and glucose homeostasis. J Nutr. 2003;133(1):261S–267S.**
32. **Layman DK, Shiue H, Sather C, Erickson DJ, Baum J.** Increased dietary protein modifies glucose and insulin homeostasis in adult women during weight loss. *J Nutr*. 2003;133(2):405–10.
33. **Qin LQ, Xun P, Bujnowski D, et al.** Higher branched-chain amino acid intake is associated with a lower prevalence of being overweight or obese in middle-aged East Asian and Western adults. *J Nutr*. 2011;141(2):249–54.
34. **Brosnan JT, Brosnan ME.** Glutamate: a truly functional amino acid. *Amino Acids*. 2013;45:413–8.
35. **Soga T, Ohishi T, Matsui T et al.** Lysophosphatidylcholine enhances glucose-dependent insulin secretion via an orphan G-protein-coupled receptor. *Biochemical and Biophysical Research Communications*. 2005;326:744–751.
36. **Yea K, Kim J, Yoon JH et al.** Lysophosphatidylcholine activates adipocyte glucose uptake and lowers blood glucose levels in murine models of diabetes. *J Biol Chem*. 2009;284:33833–40.
37. **Ferrannini E, Natali A, Camastra S et al.** Early metabolic markers of the development of dysglycemia and type 2 diabetes and their physiological significance. *Diabetes*. 2013;62:1730–7.
38. **Wang-Sattler R, Yu Z, Herder C et al.** Novel biomarkers for pre-diabetes identified by metabolomics. *Mol Syst Biol*. 2012;8:43.
39. **Ganna A, Salihovic S, Sundström J et al. Large-scale Metabolomic Profiling Identifies Novel Biomarkers for Incident Coronary Heart Disease. PLoS Genetics 2014;10:e1004801.**
40. **Okudaira S, Yukiura H, Aoki J.** Biological roles of lysophosphatidic acid signaling through its production by autotaxin. *Biochimie*. 2010;92:698–706.
41. **Ferry G, Tellier E, Try A et al.** Autotaxin is released from adipocytes, catalyzes lysophosphatidic acid synthesis, and activates preadipocyte proliferation. Up-regulated expression with adipocyte differentiation and obesity. *J Biol Chem*. 2003;278:18162–9.
42. **Rancoule C, Dusaulcy R, Treguer K et al.** Depot-specific regulation

of autotaxin with obesity in human adipose tissue. *J Physiol Biochem*. 2012;68:635–44.
43. **Dusaulcy R, Rancoule C, Gres S et al.** Adipose-specific disruption of autotaxin enhances nutritional fattening and reduces plasma lysophosphatidic acid. *J Lipid Res*. 2011;52:1247–55.
44. **Moolenaar WH, Perrakis A.** Insights into autotaxin: how to produce and present a lipid mediator. *Nat Rev Mol Cell Biol*. 2011;12(10):674–9.
45. **Nakanaga K, Hama K, Aoki J.** Autotaxin–an LPA producing enzyme with diverse functions. *J Biochem*. 2010;148(1):13–24.
46. **Rancoule C, Attane C, Gres S et al.** Lysophosphatidic acid impairs glucose homeostasis and inhibits insulin secretion in high-fat diet obese mice. *Diabetologia*. 2013;56:1394–402.
47. **Merrill A.** Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. *Chemical reviews*. 2011;111:6387–6422.
48. **Barenholz Y.** Sphingomyelin and cholesterol: from membrane biophysics and rafts to potential medical applications. *Subcell Biochem*. 2004;37:167–215.
49. **De Smet E, Mensink RP, Plat J.** Effects of plant sterols and stanols on intestinal cholesterol metabolism: suggested mechanisms from past to present. *Mol Nutr Food Res*. 2012;56:1058–72.
50. **Miettinen TA, Tilvis RS, Kesaniemi YA.** Serum plant sterols and cholesterol precursors reflect cholesterol absorption and synthesis in volunteers of a randomly selected male population. *American Journal of Epidemiology*. 1990;131:20–31.
51. **Escurriol V, Cofán M, Serra M et al.** Serum sterol responses to increasing plant sterol intake from natural foods in the Mediterranean diet. *European Journal of Nutrition*. 2009;48:373–382.
52. **Gylling H, Hallikainen M, Kolehmainen M et al.** Cholesterol synthesis prevails over absorption in metabolic syndrome. *Translational Research*. 2007;149:310–316.
53. **Andersen LW, Mackenhauer J, Roberts JC, Berg KM, Cocchi MN, Donnino MW.** Etiology and therapeutic approach to elevated lactate levels. *Mayo Clin Proc*. 2013;88:1127–40.
54. **Cox KC, NC, Carney EE, Howell MD, Donnino MW. Prevalence and significance of lactic acidosis in diabetic ketoacidosis. Emergency Medicine, Beth Israel Deaconess Medical Center; Boston: Boston, MA, USA: 2008.**
55. **Lovejoy J, Newby FD, Gebhart SS, DiGirolamo M.** Insulin resistance in obesity is associated with elevated basal lactate levels and diminished lactate appearance following intravenous glucose and insulin. *Metabolism*. 1992;41:22–7.
56. **Crawford SO, Hoogeveen RC, Brancati FL et al.** Association of blood lactate with type 2 diabetes: the Atherosclerosis Risk in Communities Carotid MRI Study. *International journal of epidemiology*. 2010;39:1647–1655.
57. **Juraschek SP, Shantha GP, Chu AY et al.** Lactate and risk of incident diabetes in a case-cohort of the atherosclerosis risk in communities (ARIC) study. *PLoS One*. 2013;8:30.
58. **Goldberg GR, Greene CL.** Update on inborn errors of metabolism: primary lactic acidemia. *Journal of pediatric health care : official publication of National Association of Pediatric Nurse Associates, Practitioners*. 1992 Jul-Aug;6(4):176–181.
59. **Manach C, Scalbert A, Morand C, Rémésy C, Jiménez L.** Polyphenols: food sources and bioavailability. *The American journal of clinical nutrition*. 2004;79:727–747.
60. **Clifford MN, Stoupi S, Kuhnert N.** Profiling and characterization by LC-MSn of the galloylquinic acids of green tea, tara tannin, and tannic acid. *J Agric Food Chem*. 2007;55(8):2797–807.
61. **Rechner AR, Wagner E, Van buren L, Van de put F, Wiseman S, Rice-evans CA. Black tea represents a major source of dietary phenolics among regular tea drinkers. Free Radic Res. 2002;36(10):1127–35.**
62. **Pero RW, Lund H.** In Vivo Treatment of Humans with Quinic Acid Enhances DNA Repair and Reduces the Influence of lifestyle Factors on Risk to Disease. *International Journal of Biotechnology and Biochemistry*. 2009;5:293–305.
63. **Jurgonski A, Juskiewicz J, Zdunczyk Z, Krol B.** Caffeoylquinic acid-rich extract from chicory seeds improves glycemia, atherogenic index, and antioxidant status in rats. *Nutrition*. 2012;28:300–6.
64. **Cho A-S, Jeon S-M, Kim M-J et al.** Chlorogenic acid exhibits anti-obesity property and improves lipid metabolism in high-fat diet-induced-obese mice. *Food and Chemical Toxicology*. 2010;48:937–943.
65. **Cherniack EP.** Polyphenols: Planting the seeds of treatment for the metabolic syndrome. *Nutrition*. 2011;27:617–623.
66. **Wahl S, Vogt S, Stückler F, et al.** Multi-omic signature of body weight change: results from a population-based cohort study. *BMC Med*. 2015;13:48.