

Metabolomic architecture of obesity implicates metabolomic lactone sulfate in cardiometabolic disease



Swapan K. Das¹, Hannah C. Ainsworth², Latchezar Dimitrov⁷, Hayrettin Okut³, Mary E. Comeau², Neeraj Sharma¹, Maggie C.Y. Ng⁴, Jill M. Norris⁵, Yii-der I. Chen⁶, Lynne E. Wagenknecht⁷, Donald W. Bowden⁸, Fang-Chi Hsu², Kent D. Taylor⁶, Carl D. Langefeld², Nicholette D. Palmer^{8,*}

ABSTRACT

Objective: Identify and characterize circulating metabolite profiles associated with adiposity to inform precision medicine.

Methods: Untargeted plasma metabolomic profiles in the Insulin Resistance Atherosclerosis Family Study (IRASFS) Mexican American cohort ($n = 1108$) were analyzed for association with anthropometric (body mass index, BMI; waist circumference, WC; waist-to-hip ratio, WHR) and computed tomography measures (visceral adipose tissue, VAT; subcutaneous adipose tissue, SAT; visceral-to-subcutaneous ratio, VSR) of adiposity. Genetic data, inclusive of genome-wide array-based genotyping, whole exome sequencing (WES) and whole genome sequencing (WGS), were evaluated to identify the genetic contributors. Phenotypic and genetic association signals were replicated across ancestries. Transcriptomic data were analyzed to explore the relationship between genetic and metabolomic data.

Results: A partially characterized metabolite, tentatively named metabolomic lactone sulfate (X-12063), was consistently associated with BMI, WC, WHR, VAT, and SAT in IRASFS Mexican Americans ($P_{MA} < 2.02 \times 10^{-27}$). Trait associations were replicated in IRASFS African Americans ($P_{AA} < 1.12 \times 10^{-07}$). Expanded analyses revealed associations with multiple phenotypic measures of cardiometabolic health, e.g. insulin sensitivity (S_i), triglycerides (TG), diastolic blood pressure (DBP) and plasminogen activator inhibitor-1 (PAI-1) in both ancestries. Metabolomic lactone sulfate levels were heritable ($h^2 > 0.47$), and a significant genetic signal at the *ZSCAN25/CYP3A5* locus ($P_{MA} = 9.00 \times 10^{-41}$, $P_{AA} = 2.31 \times 10^{-10}$) was observed, highlighting a putative functional variant (rs776746, *CYP3A5**3). Transcriptomic analysis in the African American Genetics of Metabolism and Expression (AAGMEx) cohort supported the association of *CYP3A5* with metabolomic lactone sulfate levels ($P_{FDR} = 6.64 \times 10^{-07}$).

Conclusions: Variant rs776746 is associated with a decrease in the transcript levels of *CYP3A5*, which in turn is associated with increased metabolomic lactone sulfate levels and poor cardiometabolic health.

© 2021 The Author(s). Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords Obesity; Metabolomics; Cardiometabolic disease; Minority populations

1. INTRODUCTION

Despite public health efforts to curb the growing obesity epidemic, the prevalence of obesity continues to increase and has surpassed 40% of the U.S. population, disproportionately impacting ethnic minorities [1]. This excess accumulation of adiposity is associated with an increased risk for the development of numerous metabolic diseases including type 2 diabetes (T2D), cardiovascular disease (CVD) and cancer, all of which contribute significantly to morbidity and mortality [2]. However, how best to measure adiposity remains a controversial topic, with routine measures such as body mass index (BMI) showing variability by age and sex and failing to distinguish between fat and lean mass. More

accurate imaging techniques, such as computed tomography (CT), are expensive but capture region-specific deposition, e.g. metabolically active visceral adipose tissue [3].

The main driver of obesity is energy intake that exceeds energy expenditure, a state easily entrenched given the prevailing conditions of the modern environment [4]. One approach to capture an individual's metabolic state, indicative of cellular processes and environmental exposures to promote disease, is to examine their metabolomic profile. High-throughput profiling of blood metabolites has identified associations between elevated branched chain amino acids (BCAA) and increased obesity and development of T2D, predominantly in European-derived populations [5–9]. In contrast, there

¹Department of Internal Medicine, Wake Forest School of Medicine, Winston–Salem, North Carolina, USA ²Department of Biostatistics and Data Science, Wake Forest School of Medicine, Winston–Salem, North Carolina, USA ³Office of Research, University of Kansas Medical Center, Wichita, Kansas, USA ⁴Division of Genetic Medicine, Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA ⁵Department of Epidemiology, Colorado School of Public Health, Aurora, CO, USA ⁶The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA, USA ⁷Department of Epidemiology and Prevention, Wake Forest School of Medicine, Winston–Salem, North Carolina, USA ⁸Department of Biochemistry, Wake Forest School of Medicine, Winston–Salem, North Carolina, USA

*Corresponding author. Department of Biochemistry, 1 Medical Center Blvd, Winston–Salem, NC, 27040, USA. E-mail: nallred@wakehealth.edu (N.D. Palmer).

Received September 9, 2021 • Accepted September 17, 2021 • Available online 24 September 2021

<https://doi.org/10.1016/j.molmet.2021.101342>

Abbreviations

AA	African American
AAGMEx	African American Genetics of Metabolism and Expression
BCAA	branched chain amino acids
BMI	body mass index
CT	computed tomography
CVD	cardiovascular disease
CYP3A5	Cytochrome P450 Family 3 Subfamily A Member 5
DBP	diastolic blood pressure
IRASFS	Insulin Resistance Atherosclerosis Family Study
MA	Mexican American

PAI-1	plasminogen activator inhibitor-1
SAT	subcutaneous adipose
S _i	insulin sensitivity
T2D	type 2 diabetes
TG	triglycerides
VAT	visceral adipose
VSR	visceral-to-subcutaneous ratio
WC	waist circumference
WES	whole exome sequencing
WFSM	Wake Forest School of Medicine
WGS	whole genome sequencing
WHR	waist-to-hip ratio

is a significant knowledge gap as to whether these results fully or partially translate to other ancestries.

The goal of this study was to address this gap through metabolomic analysis in individuals of Mexican American ethnicity, one of the fastest-growing ethnic minority populations in the U.S. and one disproportionately impacted by the obesity epidemic [10]. Both anthropometric and CT-derived measures of adiposity were included to capture differing aspects of the disease state, e.g. whole-body and depot-specific deposition. An expanded focus on previously uncharacterized metabolites further enhances the likelihood of identifying novel metabolites of significant effect. Genetic and transcriptomic data can provide insight into genetic regulation, contributing to metabolite characterization. These results will provide new insights into the etiology of disease and identify prospective biomarkers indicative of metabolic health status associated with increased adipose deposition.

2. MATERIALS AND METHODS

2.1. Study populations

The Insulin Resistance Atherosclerosis Family Study (IRASFS) was designed to investigate the genetic and epidemiologic basis of glucose homeostasis and abdominal adiposity. The study design, recruitment methods and phenotype assessment have been described previously [11]. Specific to this study, Mexican American families were recruited from two clinical centers including San Antonio, TX, and San Luis Valley, CO, and African Americans were recruited from one clinical center in Los Angeles, CA.

The African American Genetics of Metabolism and Expression (AAGMEx) study was designed to identify the genetic regulatory mechanisms associated with insulin resistance. The study design, recruitment methods and phenotype assessment have been described previously [12]. Specific to this report, healthy, self-identified African Americans, aged 18–60 years and with a BMI between 18 and 42 kg/m², were recruited to the Clinical Research Unit at Wake Forest School of Medicine (WFSM).

All study protocols were approved by the Institutional Review Board of each participating clinical and analysis site and all participants provided written informed consent.

2.2. Clinical phenotyping in IRASFS

Phenotypic assessment has been described previously [11]. Briefly, cardiometabolic phenotyping was conducted using a standardized protocol and included a frequently sampled intravenous glucose tolerance test (FSIGT), anthropometric measures, adipose deposition by computed tomography (CT) scanning [11], resting blood pressure, fasting blood draw and spot urine collection. Interviews included questions on medical history, physical activity, alcohol intake, tobacco

use and demographic characteristics. Laboratory measures included cholesterol levels and a selection of biomarkers.

2.3. Metabolite profiling in IRASFS

We performed metabolite profiling of fasting plasma samples collected in ethylenediaminetetraacetic acid (EDTA) tubes and stored at –80 °C since baseline collection from 1999 to 2002. Detection and quantification of metabolites was completed by Metabolon, Inc. (Morrisville, North Carolina) using untargeted liquid chromatography–mass spectrometry (LCMS; DiscoveryHD4 panel). Data were block corrected for run day, normalized by batch, and missing data by metabolite were imputed to the lowest measured value. The correlation structure among all measured metabolites prior to normalization was examined using Spearman's rank-order correlation (r_s). Prior to analysis, the metabolites were individually Box–Cox power transformed [13] to best approximate distributional assumptions (conditional normality, homogeneity of variance) on a per-metabolite basis to minimize false positives.

2.4. Genotyping in IRASFS

A genome-wide association study (GWAS) was conducted on 1034 IRASFS Mexican American samples using the Illumina OmniExpress and 1S arrays (Illumina Inc.; San Diego, CA, USA) as described previously [14]. Data were imputed via IMPUTE2 [15] using two reference panels: 1) 1000 Genomes phase 3 and 2) WGS data from 623 IRASFS Mexican American samples. Whole exome sequencing was performed on 1205 IRASFS Mexican American samples as described previously [16]. Whole genome sequencing was performed on 623 IRASFS Mexican American samples using the Illumina HiSeq X Ten instrument by Macrogen, targeting a mean depth of 30X (paired-end, 150 bp reads). Genotyping in IRASFS African American samples was performed using the Illumina Multi-Ethnic Genotyping Array (MEGA; Illumina, San Diego, CA), with imputation to 1000 Genomes phase 3 and the African Genome Variation Project as described previously [17].

2.5. Multiomics in AAGMEx

Phenotypic and genomic assessment has been described previously [12]. Specific to this report, abdominal subcutaneous adipose near the umbilicus and vastus lateralis skeletal muscle biopsies were obtained under fasting conditions with local anesthesia. Genome-wide expression data were generated using the HumanHT-12 v4 Expression BeadChip (Illumina, San Diego, CA). Metabolite profiling of fasting plasma samples stored at –80 °C since collection was performed by Metabolon, Inc. (Morrisville, North Carolina) using untargeted liquid chromatography–mass spectrometry (DiscoveryHD4 panel). Prior to analysis, data were block corrected for run day and normalized by batch; missing data by metabolite were imputed to the lowest measured value and metabolite

values were log transformed. Genome-wide array-based genotyping was completed using Infinium HumanOmni5Exome-4 v1.1 DNA Analysis BeadChips (Illumina, San Diego, CA).

2.6. Statistical analysis

For IRASFS, clinical phenotypes and metabolites were transformed to best approximate the distributional assumptions of conditional normality (conditional on the covariates) and homogeneity of the variance. Clinical phenotype correlations were estimated using a Pearson's correlation (r_P) after regressing out the effects of the covariates (age, sex and recruitment center (Mexican American samples only)). Variance component models as implemented in SOLAR [18] were used to test for associations accounting for familial relationships using a random effect model. Tests of association between individual metabolites (outcome) and clinical phenotypes or genetic variants (predictors) were computed using the Wald test. All models were minimally adjusted for age, sex and recruitment center (for Mexican American samples only). The analysis of cardiometabolic traits was additionally adjusted for BMI. A conservative Bonferroni correction was used to account for multiple testing. In addition, the association of metabolites with future risk of diabetes or cardiovascular disease (CVD) was assessed using the composite phenotypes of metabolic syndrome (MetS) [19] and the Framingham Heart Study (FHS) CVD 10-year Risk Score [20]. Covariates included age, sex and recruitment center (MetS) and recruitment center only (FHS CVD risk score) owing to the inclusion of a subset of the standard covariates in the composite phenotypes. Owing to the younger age of the IRASFS cohort, the FHS CVD 10-year Risk Score was calculated both with inclusion of all participants with phenotypes winsorized to the algorithm minimum or with the exclusion of participants who did not meet the algorithm criteria, i.e. at least 30 years of age, 200 mmHg > SBP > 90 mmHg, HDL < 100 mg/dL and TC > 100 mg/dL. Heritability (h^2) was estimated by the proportion of total variation in phenotype due to genetic effects in SOLAR [18]. GWAS analysis was performed using RVTESTS [21] accounting for age, sex, recruitment center (Mexican American samples only), ancestry estimates (ADMIXTURE version 1.21 (<http://www.genetics.ucla.edu/software/admixture>)) and relatedness (kinship matrix). Statistical significance for the analysis of genetic variants was set at $P < 5.00 \times 10^{-08}$. For significant variants, the proportion of variance explained was calculated using SOLAR [18].

In AAGMEX, clinical phenotypes and metabolites were transformed to approximate conditional normality and homogeneity of the variance. Linear regression models were fit to test for the association between individual metabolites (outcome) and clinical phenotypes or genetic variants (predictors). Models were adjusted for age, sex and admixture estimates. For gene expression studies, a Benjamini-Hochberg false-discovery rate (FDR) adjusted P -value was computed by trait and a P_{FDR} value < 0.01 (1%) was considered significant. In addition, multivariate and mediation analyses [22] were performed to delineate the relative contributions and unravel the complex relationships among genotype, transcript, and plasma metabolite levels.

3. RESULTS

3.1. Subject characteristics

Recruitment for IRASFS was population-based and not based on disease status, e.g. diabetes. A maximum of 1205 Mexican Americans and 569 African Americans were included in these analyses. Char-

acteristics of the study participants are shown in **ST1**. Overall, participants were more frequently female (59%) and were overweight (BMI, 28.92 and 29.99 kg/m², respectively). Adiposity phenotype correlations are presented in **ST2**.

3.2. Metabolite associations with adiposity in IRASFS Mexican Americans

A total of 882 metabolites were transformed and analyzed in up to 1108 Mexican American samples. Of these, 609 metabolites belonged to the categories of amino acids, carbohydrates, cofactors and vitamins, energy, lipids, nucleotides and peptide metabolism. The remaining 273 metabolites were uncharacterized, i.e. "unknowns". Among anthropometric traits in Mexican American samples, 275, 299 and 185 metabolites were associated with BMI, WC and WHR, respectively (**SF1**, **ST4**). For all anthropometric traits, lipid metabolism was the predominant superpathway, highlighting phospholipid ($n = 32$), lysolipid ($n = 16$) and sphingolipid ($n = 12$) metabolism. Association with branched chain amino acid (BCAA) metabolism, i.e. leucine, isoleucine and valine, was also a significant feature ($n = 13$). A total of 163 metabolites were shared between the three traits, with 263 metabolites shared between BMI and WC, the two most correlated anthropometric traits ($r_P = 0.92$, **ST1**). Among these traits, 12, 21 and 7 metabolites, respectively, were unique. Among all three phenotypes, the most significant metabolite was a partially characterized metabolite tentatively named metabolonic lactone sulfate (X-12063; $P = 4.51 \times 10^{-56}$, 4.05×10^{-59} and 2.02×10^{-27} , respectively; **Table 1**) that was positively associated.

Among the CT measures, 256, 256 and 1 metabolites were associated with VAT, SAT and VSR, respectively (**SF1**, **ST4**). Despite similar numbers, the results for VAT and SAT were not identical, i.e. 204 overlapping and 52 unique metabolites each, consistent with trait correlations ($r_P = 0.65$, **ST2**). Nearly half of the metabolites associated with imaging measures of adiposity represented lipid metabolism ($n = 125$) with results again highlighting phospholipid ($n = 38$), lysolipid ($n = 19$) and sphingolipid ($n = 16$) metabolism. Metabolonic lactone sulfate was the second most significant result for VAT ($P = 1.02 \times 10^{-43}$), followed closely behind sphingomyelin (d18:0/18:0, d19:0/17:0) ($P = 2.21 \times 10^{-44}$), and the most significant result for SAT ($P = 1.31 \times 10^{-47}$). These results were consistent with the positive association observed with the anthropometric traits (**Table 1**). Metabolonic lactone sulfate was not significantly associated with VSR ($P = 0.026$). Metabolite correlations with metabolonic lactone sulfate are presented in **ST3**.

3.3. Transancestry replication in IRASFS African Americans

Owing to the highly significant association of metabolonic lactone sulfate with multiple adiposity measures in IRASFS Mexican Americans, transancestry replication was explored in up to 569 IRASFS African American samples. Metabolonic lactone sulfate levels were significantly associated with BMI ($P = 1.84 \times 10^{-24}$), WC ($P = 7.91 \times 10^{-24}$) and WHR ($P = 1.12 \times 10^{-07}$) with a consistent direction of effect, i.e. increasing metabolonic lactone sulfate levels were associated with increasing adiposity, in the transancestry analysis (**Table 2**, **ST4**). Significance levels were comparatively reduced, likely owing to the reduced sample size. Among the CT measures, metabolonic lactone sulfate levels were significantly associated with VAT ($P = 3.38 \times 10^{-14}$) and SAT ($P = 1.00 \times 10^{-20}$) with a consistent direction of effect (**Table 2**, **ST4**). Results with VSR were non-significant ($P = 0.20$).

Table 1 — Association of metabolonic lactone sulfate with cardiometabolic phenotypes in the IRASFS Mexican Americans.

	Mexican Americans				
	<i>N</i>	Beta (SE)	<i>P</i>	Beta (SE)	<i>P</i> _{adjBMI}
Glucose Homeostasis					
Insulin Sensitivity (<i>S_I</i>)	921	−0.20 (0.016)	1.09E-35	−0.11 (0.017)	3.03E-11
Glucose Effectiveness (<i>S_G</i>)	921	−26.78 (3.51)	1.37E-14	−11.13 (3.24)	7.31E-04
Acute Insulin Response (<i>AIR_G</i>)	921	0.00021 (0.000048)	1.80E-05	0.000031 (0.000012)	0.49
Disposition Index (<i>DI</i>)	921	−0.00023 (0.000011)	5.09E-19	−0.00013 (0.000024)	1.17E-07
Fasting Glucose (<i>FG</i>)	980	0.032 (0.00028)	4.27E-21	0.016 (0.00031)	7.77E-07
Fasting Insulin (<i>FI</i>)	978	0.030 (0.0027)	2.96E-26	0.0093 (0.0018)	0.0022
Homeostatic Model Assessment of Insulin Resistance (<i>HOMA_{IR}</i>)	939	0.32 (0.029)	6.30E-28	0.12 (0.032)	1.07E-04
Homeostatic Model Assessment of Beta Cell Function (<i>HOMA_B</i>)	939	0.0056 (0.00041)	7.45E-17	0.0014 (0.00019)	0.038
Metabolic Clearance Rate of Insulin (<i>MCR_I</i>)	928	−0.16 (0.012)	2.86E-41	−0.084 (0.014)	2.01E-09
Adiposity					
Body Mass Index (<i>BMI</i>)	1108	0.075 (0.0045)	4.51E-56		
Waist Circumference (<i>WC</i>)	1105	0.032 (0.00056)	4.05E-59		
Waist-to-Hip Ratio (<i>WHR</i>)	1108	3.42 (0.070)	2.02E-27		
Visceral Adipose Tissue (<i>VAT</i>)	1058	0.0079 (0.00054)	1.02E-43		
Subcutaneous Adipose Tissue (<i>SAT</i>)	1058	0.0029 (0.00019)	1.31E-47		
Visceral-Subcutaneous Ratio (<i>VSR</i>)	1018	0.46 (0.21)	0.026		
Alanine Aminotransferase (<i>ALT</i>)	1029	0.020 (0.0029)	2.72E-10	0.012 (0.0029)	1.90E-05
Aspartate Aminotransferase (<i>AST</i>)	1079	0.012 (0.0030)	9.70E-05	0.0067 (0.0027)	0.013
Gamma-glutamyltransferase (<i>GGT</i>)	1047	0.00038 (0.00056)	0.55	−0.00039 (0.00056)	0.49
Lipids					
Triglycerides (<i>TRIGS</i>)	1103	0.0030 (0.00016)	1.20E-17	0.0020 (0.00032)	8.11E-10
Total Cholesterol (<i>TC</i>)	1106	0.0032 (0.000035)	1.13E-04	0.0026 (0.00075)	6.23E-04
Low Density Lipoprotein Cholesterol (<i>LDL</i>)	1093	0.0027 (0.0010)	0.0078	0.0016 (0.00091)	0.084
High Density Lipoprotein Cholesterol (<i>HDL</i>)	1105	−0.016 (0.0021)	1.60E-11	−0.0058 (0.0022)	0.0090
Blood Pressure					
Pulse Pressure (<i>PP</i>)	1108	0.0013 (0.0021)	0.598511	−0.0024 (0.0022)	0.29
Mean Arterial Pressure (<i>MAP</i>)	1104	0.024 (0.00090)	2.87E-16	0.011 (0.00047)	1.21E-04
Systolic Blood Pressure (<i>SBP</i>)	1108	0.011 (0.00062)	2.68E-08	0.0033 (0.00036)	0.069
Diastolic Blood Pressure (<i>DBP</i>)	1108	0.027 (0.0027)	1.94E-17	0.013 (0.0030)	7.91E-06
Biomarkers					
Fibrinogen (<i>FIBRIN</i>)	1106	0.002261 (0.000012)	7.61E-06	−0.00044 (0.000012)	0.37
Plasminogen Activator Inhibitor (<i>PAI-1</i>)	1107	0.008951 (0.000748)	4.57E-31	0.0047 (0.00078)	2.20E-09
Adiponectin (<i>ADP</i>)	1099	−0.037745 (0.004406)	4.59E-17	−0.019 (0.0043)	1.10E-05
C Reactive Protein (<i>CRP</i>)	1023	0.05165 (0.007386)	1.32E-12	0.015 (0.0071)	0.041
Interleukin 6 (<i>IL6</i>)	848	0.005737 (0.011034)	0.60	−0.014 (0.010)	0.15
Retinol Binding Protein 4 (<i>RBP4</i>)	927	0.016958 (0.004519)	1.93E-04	0.020 (0.0040)	9.27E-07

Abbreviations: *N*, number of participants; SE, standard error; *P*_{adjBMI}, *p*-value with inclusion of BMI as a covariate. *P*-values in bold remained significant after multiple comparisons correction, i.e. 0.05/32 phenotypes = 0.0016.

3.4. Association of metabolonic lactone sulfate with cardiometabolic disease

Metabolonic lactone sulfate was the most significantly associated metabolite with adiposity-related phenotypes in IRASFS Mexican Americans, which was replicated in IRASFS African Americans. In addition to adiposity-related phenotypes, metabolonic lactone sulfate was also associated with multiple features of cardiometabolic disease. Among IRASFS Mexican Americans (Table 1), an inverse association was observed with insulin sensitivity (*S_I*; $P = 1.09 \times 10^{-35}$, $P_{\text{adjBMI}} = 3.03 \times 10^{-11}$) and metabolic clearance rate of insulin (*MCR_I*; $P = 2.86 \times 10^{-41}$, $P_{\text{adjBMI}} = 2.01 \times 10^{-09}$); both are dynamic measures of glucose homeostasis. Positive associations were also observed with total triglyceride (*TRIGS*; $P = 1.20 \times 10^{-17}$, $P_{\text{adjBMI}} = 8.11 \times 10^{-10}$), diastolic blood pressure (*DBP*; $P = 1.94 \times 10^{-17}$, $P_{\text{adjBMI}} = 7.91 \times 10^{-06}$) and plasminogen activator inhibitor (*PAI-1*; $P = 4.57 \times 10^{-31}$, $P_{\text{adjBMI}} = 2.20 \times 10^{-09}$). Results from the sex-stratified analysis were broadly consistent, although more significant associations were observed in females with a slightly larger sample size (ST5). The same associations with consistent direction of effect were observed in IRASFS African Americans (Table 2). Unique to the African American analyses, metabolonic lactone sulfate was significantly associated with adiponectin (*ADP*;

$P = 7.8 \times 10^{-17}$, $P_{\text{adjBMI}} = 9.1 \times 10^{-10}$) compared with a more attenuated association observed in Mexican Americans ($P = 4.6 \times 10^{-17}$, $P_{\text{adjBMI}} = 1.1 \times 10^{-05}$). Consistent with these data, metabolonic lactone sulfate was significantly associated with the presence of MetS ($\beta = 0.36 \pm 0.036$, $P = 2.14 \times 10^{-24}$), indicative of future development of diabetes and CVD [23]. In contrast, it was nominally associated with the FHS CVD Risk Score ($\beta = 0.0024 \pm 0.012$, $P = 0.040$ with winsorization and $\beta = 0.0024 \pm 0.014$, $P = 0.081$ with the exclusion of phenotypic algorithm outliers), indicative of future development of CVD [20].

3.5. Genetic architecture of metabolonic lactone sulfate

Taking advantage of the multigenerational pedigrees in IRASFS, the heritability of metabolonic lactone sulfate was estimated to be 0.47 ($P = 1.46 \times 10^{-21}$) in IRASFS Mexican Americans and 0.51 ($P = 1.57 \times 10^{-15}$) in IRASFS African Americans in a model accounting for age, sex, recruitment center (for Mexican American samples only) and BMI. Given the evidence of a strong genetic contribution, a genome-wide association study (GWAS) was performed to identify the genetic contributors to metabolonic lactone sulfate. A GWAS in IRASFS Mexican Americans ($\lambda = 1.05$) identified 1194 genome-wide significant variants ($P < 5.00 \times 10^{-08}$) associated with

Table 2 — Association of metabolonic lactone sulfate with cardiometabolic phenotypes in the IRASFS African Americans.

	African Americans				
	<i>N</i>	Beta (SE)	<i>P</i>	Beta (SE)	<i>P</i> _{adjBMI}
Glucose Homeostasis					
Insulin Sensitivity (<i>S_I</i>)	493	−0.31 (0.034)	4.63E-19	−0.20 (0.036)	3.41E-08
Glucose Effectiveness (<i>S_G</i>)	493	−30.18 (5.74)	1.90E-07	−19.68 (5.41)	2.86E-04
Acute Insulin Response (<i>AI_{R₀}</i>)	492	0.00014 (0.000042)	0.0091	0.000057 (0.00)	0.24
Disposition Index (<i>DI</i>)	492	−0.00015 (0.00)	1.70E-05	−0.000097 (0.000019)	0.0038
Fasting Glucose (<i>FG</i>)	507	0.023 (0.00064)	2.43E-08	0.012 (0.0021)	0.0053
Fasting Insulin (<i>FI</i>)	507	0.024 (0.0034)	1.11E-11	0.015 (0.0034)	1.10E-05
Homeostatic Model Assessment of Insulin Resistance (<i>HOMA_{IR}</i>)	482	0.32 (0.039)	1.36E-14	0.19 (0.043)	8.45E-06
Homeostatic Model Assessment of Beta Cell Function (<i>HOMA_B</i>)	482	0.0055 (0.000064)	2.20E-09	0.0033 (0.00084)	2.43E-04
Metabolic Clearance Rate of Insulin (<i>MCR_I</i>)	491	−0.18 (0.019)	7.48E-19	−0.10 (0.022)	2.31E-06
Adiposity					
Body Mass Index (<i>BMI</i>)	565	0.062 (0.0058)	1.84E-24		
Waist Circumference (<i>WC</i>)	564	0.024 (0.00074)	7.91E-24		
Waist-to-Hip Ratio (<i>WHR</i>)	569	2.59 (0.086)	1.12E-07		
Visceral Adipose Tissue (<i>VAT</i>)	519	0.0068 (0.00085)	3.38E-14		
Subcutaneous Adipose Tissue (<i>SAT</i>)	519	0.0022 (0.000020)	1.00E-20		
Visceral-Subcutaneous Ratio (<i>VSR</i>)	501	−0.41 (0.32)	0.20		
Alanine Aminotransferase (<i>ALT</i>)	562	0.0011 (0.0079)	0.88	0.0010 (0.0072)	0.89
Aspartate Aminotransferase (<i>AST</i>)	565	−0.0082 (0.006122)	0.17	−0.0089 (0.0054)	0.10
Gamma-glutamyltransferase (<i>GGT</i>)	556	0.00079 (0.0011)	0.48	−0.00025 (0.0010)	0.80
Lipids					
Triglycerides (<i>TRIGS</i>)	569	0.0022 (0.00059)	2.65E-04	0.0015 (0.00054)	0.0048
Total Cholesterol (<i>TC</i>)	569	0.0032 (0.0012)	0.0060	0.0021 (0.0011)	0.049
Low Density Lipoprotein Cholesterol (<i>LDL</i>)	566	0.0042 (0.0012)	9.96E-04	0.0021 (0.0012)	0.075
High Density Lipoprotein Cholesterol (<i>HDL</i>)	569	−0.016 (0.0034)	6.00E-06	−0.0062 (0.0031)	0.060
Blood Pressure					
Pulse Pressure (<i>PP</i>)	569	−0.0030 (0.0026)	0.34	−0.0061 (0.0029)	0.034
Mean Arterial Pressure (<i>MAP</i>)	566	0.017 (0.0039)	1.70E-05	0.0095 (0.00020)	0.0086
Systolic Blood Pressure (<i>SBP</i>)	569	0.0053 (0.00037)	0.034	0.00083 (0.0024)	0.72
Diastolic Blood Pressure (<i>DBP</i>)	569	0.018 (0.0040)	4.19E-06	0.012 (0.0037)	8.66E-04
Biomarkers					
Fibrinogen (<i>FIBRI</i>)	569	0.0025 (0.000020)	2.21E-04	0.00024 (0.000029)	0.72
Plasminogen Activator Inhibitor (<i>PAI-1</i>)	569	0.011 (0.0014)	2.92E-14	0.0060 (0.0015)	3.00E-05
Adiponectin (<i>ADP</i>)	569	−0.067 (0.0077)	7.79E-17	−0.048 (0.0078)	9.14E-10
C Reactive Protein (<i>CRP</i>)	558	0.053 (0.0083)	1.60E-10	0.030 (0.0082)	2.61E-04
Interleukin 6 (<i>IL6</i>)	387	0.0096 (0.010)	0.34	−0.0025 (0.0099)	0.80
Retinol Binding Protein 4 (<i>RBP4</i>)	518	0.018618 (0.005275)	9.73E-04	0.019 (0.0026)	2.45E-04

Abbreviations: *N*, number of participants; SE, standard error; *P*_{adjBMI}, *p*-value with inclusion of BMI as a covariate. *P*-values in **bold** remained significant after multiple comparisons correction, i.e. 0.05/32 phenotypes = 0.0016.

metabolonic lactone sulfate (**SF2 and ST6**). The three most significant variants (rs6465750, rs10242455 and rs776746; $P = 9.00 \times 10^{-41}$) were common variants (reference allele frequency (RAF) = 14%) in complete linkage disequilibrium ($r^2 = 1.0$). For these variants, the minor allele was associated with a significant decrease in metabolonic lactone sulfate levels ($\beta = -0.87 \pm 0.06$; **SF3**). This locus, as defined by one of the three associated variants, explained 20% of the trait variance for metabolonic lactone sulfate. All three variants are located intronically in *ZSCAN25* and *CYP3A5*. The associated interval spanned chr7: 98.7–99.4 Mb and was flanked by recombination hotspots (Figure 1A). Conditional analysis with the inclusion of any one of the three variants as a covariate diminished the association in this region to $P > 7.88 \times 10^{-05}$ (Figure 1B and ST7). These three genetic variants were not significantly associated among the cardiometabolic traits assessed in IRASFS Mexican Americans (**ST8**). Transancestry replication of this region was observed in IRASFS African Americans (Figure 1C, **ST9**). Among the three most significant variants observed in IRASFS Mexican Americans evaluated *a priori* in IRASFS African Americans, there was a significant change in the allele frequency, with rs6465750 being the most significant ($P = 2.31 \times 10^{-31}$) but with an RAF of 67%. An agnostic scan of the associated interval (chr7: 98.7–99.4 Mb) in IRASFS African Americans identified one additional variant

(rs2687135, $P = 2.22 \times 10^{-10}$). In IRASFS Mexican Americans, this variant was highly significant ($P = 4.35 \times 10^{-34}$) but not among the top observed associations. Analysis of variants identified from WES (**ST10**) and WGS (**ST11**) in the associated interval in IRASFS Mexican Americans did not identify additional variants more strongly associated with metabolonic lactone sulfate.

3.6. Transcriptomic architecture of metabolonic lactone sulfate

With significant transancestry replication observed in IRASFS African Americans, the genomic characterization of metabolonic lactone sulfate was extended to the African American Genetics of Metabolism and Expression (AAGMEx) cohort (**ST12**). Consistent with the observations in IRASFS, log-transformed metabolonic lactone sulfate in AAGMEx was significantly associated with BMI ($\beta = 0.21$, $P = 3.90 \times 10^{-07}$) and *S_I* ($\beta = -0.28$, $P = 2.60 \times 10^{-05}$) with a consistent direction of effect. The *a priori* evaluation of the three most significant variants associated with metabolonic lactone sulfate from the IRASFS Mexican Americans yielded significant positive associations under an additive genetic model: rs6465750 ($P = 3.92 \times 10^{-08}$), rs10242455 ($P = 1.78 \times 10^{-08}$) and rs776746 ($P = 1.85 \times 10^{-08}$). An agnostic scan of the associated interval (chr7: 98.7–99.4 Mb) in AAGMEx did not identify any variants with stronger associations (**ST13**).

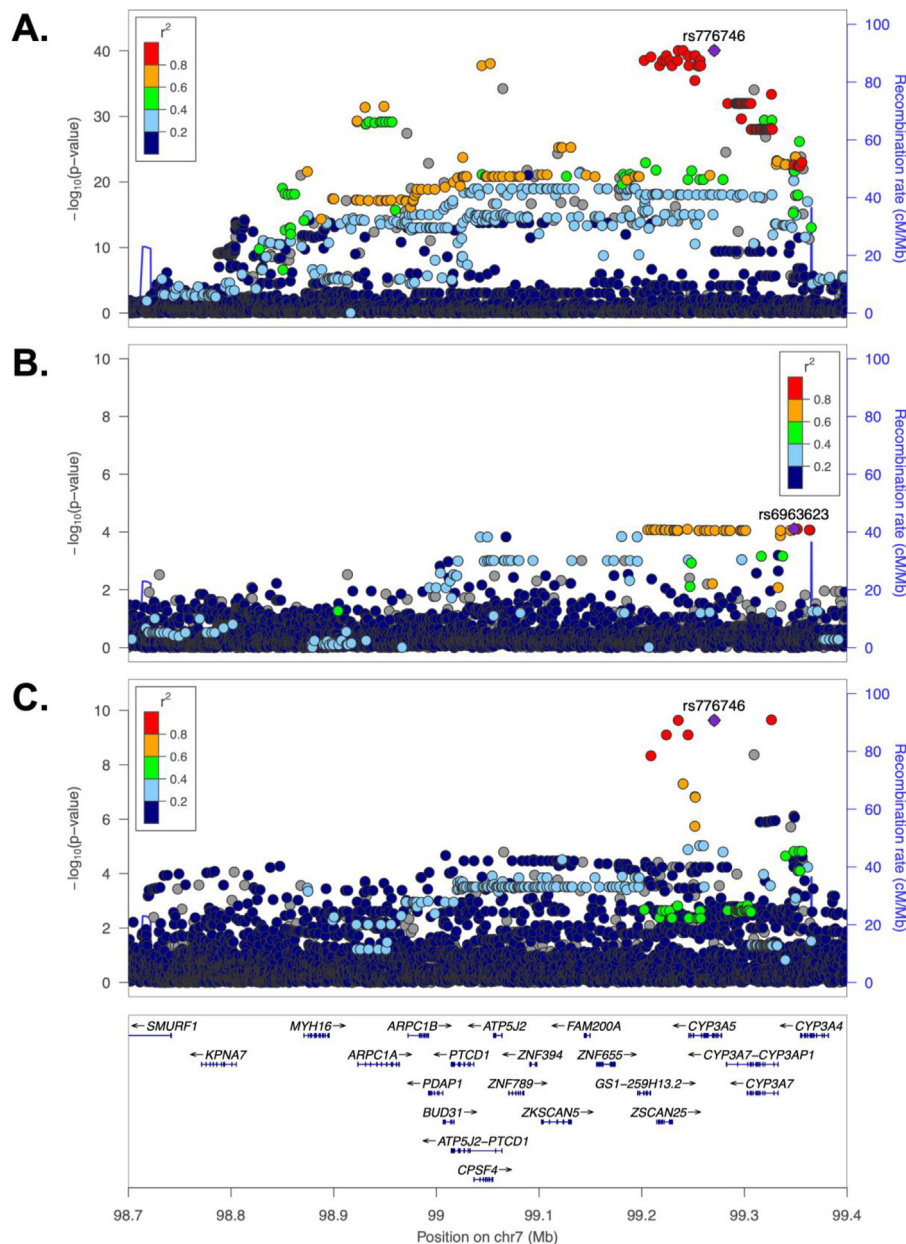


Figure 1: Regional association plot for metabolonic lactone sulfate in IRASFS. A. Mexican Americans, B. Mexican Americans, conditional on rs776746 and C. African Americans. Variants are plotted with their p -values (as $-\log_{10}$ values) as a function of genomic position (hg19). The index variant is represented by a purple diamond. Color of additional variants indicates correlation with the index SNP (red, $r^2 \geq 0.80$; orange, $0.60 \leq r^2 < 0.80$; green, $0.40 \leq r^2 < 0.60$; light blue, $0.20 \leq r^2 < 0.40$; dark blue, $r^2 < 0.20$; gray, no r^2 value available) based on pairwise r^2 values from 1000 Genomes (AMR for Mexican Americans and AFR for African Americans). Estimated recombination rates (taken from HapMap) are plotted to reflect the local linkage disequilibrium (LD) structure. Gene annotations were taken from the UCSC genome browser.

In addition to genetic and metabolic data, AAGMEx also has adipose and muscle transcriptomic expression profiling. Metabolonic lactone sulfate plasma levels were positively and negatively associated with 694 and 579 adipose transcripts ($P_{\text{FDR}} < 0.01$), respectively (ST14). Subsetting the results to chromosome 7 identified 54 associated transcripts, with two located within the genetic association interval. The most significant transcript was *CYP3A5*, which was negatively associated with metabolonic lactone sulfate plasma levels ($P_{\text{FDR}} = 1.40 \times 10^{-10}$). No muscle transcripts were associated with metabolonic lactone sulfate plasma levels ($P_{\text{FDR}} > 0.01$).

A mediation analysis was performed to examine the relationship among the rs776746 genotype, *CYP3A5* transcript levels and metabolonic lactone sulfate plasma levels in the AAGMEx cohort. Results of the multivariate modeling and corresponding mediation analysis (Table 3) indicated that there was an association between the rs776746 genotype and metabolonic lactone sulfate plasma levels even after adjusting for *CYP3A5* transcript levels, age, sex and admixture estimates (1.29×10^{-05} ; Table 3A). In addition, there was an association between *CYP3A5* transcript levels and metabolonic lactone sulfate plasma levels after adjusting for age, sex, admixture

Table 3 — Mediation analysis to examine the relationship between rs776746 genotype, *CYP3A5* transcript levels and metabolonic lactone sulfate plasma levels in AAGME African Americans. A. Multivariate analysis. B. Causal mediation analysis.

A.			
Dependent Variable	Variable	Estimate ± SE	P-value
log (metabolonic lactone sulfate)	Intercept	3.82 ± 0.61	1.42E-09
	rs776746	0.15 ± 0.034	1.29E-05
	<i>CYP3A5</i> transcript	−0.69 ± 0.12	5.97E-08
	Age	0.0016 ± 0.0019	0.40
	Female	−0.0014 ± 0.045	0.98
	Admixture estimate	0.18 ± 0.28	0.52
B.			
Test	P-value		
Causal inference test - omnibus	0.021		
Association between metabolonic lactone sulfate plasma levels and the rs776746 genotype	1.85E-08		
Association between metabolonic lactone sulfate plasma levels and <i>CYP3A5</i> transcript levels given the rs776746 genotype	5.97E-08		
Association between <i>CYP3A5</i> transcript levels and the rs776746 genotype given metabolonic lactone sulfate plasma levels	0.021		
Equivalence test of the rs776746 genotype and metabolonic lactone sulfate given <i>CYP3A5</i> transcript levels	3.59E-05		

estimates and the rs776746 genotype ($P = 5.97 \times 10^{-08}$; Table 3A). Thus, although the rs776746 genotype and *CYP3A5* transcript levels are correlated, they both have independent contributions to metabolonic lactone sulfate plasma levels. The mediation analysis underscores these patterns and highlights their independent and combined contributions to metabolite levels (Table 3B). Furthermore, the relationship between the rs776746 genotype and the *CYP3A5* transcript level remains significant even after adjusting for the metabolite; thus, the association between rs776746 genotype and *CYP3A5* transcript level are not due to their respective associations with metabolonic lactone sulfate plasma levels ($P = 0.021$).

4. DISCUSSION

Using untargeted metabolomic profiling in plasma, we identified a previously unknown metabolite tentatively named metabolonic lactone sulfate (X-12063) to be significantly and consistently associated with markers of cardiometabolic health. Increasing levels of this metabolite were associated with obesity (increased waist circumference, BMI and SAT), insulin resistance (reduced insulin sensitivity and increased fasting glucose), dyslipidemia (increased triglycerides) and hypertension (increased diastolic blood pressure and mean arterial pressure). Levels of metabolonic lactone sulfate are genetically regulated by genetic variation at the *CYP3A5* locus. Variant rs776746 encodes a cryptic splice site resulting in protein truncation. With predicted decreasing levels of functional *CYP3A5*, levels of metabolonic lactone sulfate increase, associated with worsening cardiometabolic health. Obesity is a complex multifactorial disease that is a strong risk factor for insulin resistance, cardiovascular disease (CVD) and certain types of cancer [3]. Metabolonic lactone sulfate was the most significant metabolite identified for its association with routinely assessed anthropometric measures of adiposity in a population-based cohort of Mexican Americans from IRASFS. These associations capture the total and central adipose deposition, despite varying degrees of correlation among traits. Consistent with these findings, this metabolite was also the most associated metabolite with SAT, which is significantly

correlated with BMI and WC. With a demonstrated role in central adiposity, it was also associated with VAT, which is predictive of future development of metabolic disease [24]. Notably, these findings were generalized and extended to the IRASFS African American cohort, where significant associations of metabolonic lactone sulfate with adiposity were observed with consistent direction of effect.

Consistent with the role of obesity in the promotion of metabolic disease, metabolonic lactone sulfate was significantly associated with multiple cardiometabolic measures in IRASFS Mexican Americans that persisted after adjustment for BMI in both Mexican Americans and African Americans. Specific to a role in type 2 diabetes (T2D), it was associated with measures indicative of insulin resistance. Insulin sensitivity, a dynamic measure of glucose homeostasis and direct assessment of insulin action, was negatively associated with metabolonic lactone sulfate, indicating increased insulin resistance. HOMA_{IR}, a basal measure of glucose homeostasis, was also associated, however, likely driven through the contribution of fasting glucose, which persisted after adjustment for BMI. Fasting insulin was not associated after adjustment for BMI, despite a decrease in the metabolic clearance rate of insulin (MCRI). This decrease in MCRI may be related to the measurement assay as the FSIGT mainly reflects hepatic insulin clearance [25]. Among lipid phenotypes, metabolonic lactone sulfate was positively associated with triglycerides. In the general population, triglyceride levels have risen in concert with the prevalence of obesity [26] and represent a significant risk factor for CVD [27]. In addition, metabolonic lactone sulfate was positively associated with DBP; high blood pressure represents an additional risk factor for CVD [28]. When examined as composite phenotypes, associations observed with MetS and the FHS CVD Risk Score support an association of metabolonic lactone sulfate with future development of diabetes and, to a lesser extent, CVD [20,23].

Metabolonic lactone sulfate is a partially characterized metabolite (X-12063) identified by Metabolon on their LC–MS global metabolomics platform, i.e. routinely observed with a mass spectrometric signature. However, the chemical structure has not been confirmed. Metabolonic lactone sulfate has been previously associated with insulin resistance but was overshadowed by the characterized metabolite α -hydroxybutyrate, an early marker for dysglycemia [29,30] and progression to T2D [31]. Subsequently, metabolonic lactone sulfate was found to be inversely associated with selenoprotein P levels, consistent with the pattern observed for branched chain amino acids [32]. More recent studies have evolved to incorporate a systems biology approach to metabolite identification, specifically focusing on genetic contributions. These results have implicated genotype–metabolite associations for metabolonic lactone sulfate at *CYP3A5* (chromosome 7) and *SLC01B1* (chromosome 12) [33–35]. Functional annotation using Gaussian graphical modeling (GGM) implicates dehydroisoandrosteronesulfate (DHEA-S) as an association partner. When GGM was combined with GWAS, metabolonic lactone sulfate had a consensus prediction for a role in steroid metabolism. Notably, these results were derived from a European population. In comparison, correlation estimates between metabolonic lactone sulfate and DHEA-S from the ethnic minorities presented herein were low. However, the most highly correlated annotated metabolite was 5 α -androstan-3 α , 17 β -diol mono-sulfate from the steroid pathway supporting this pathway designation. Consistent with a previous study [35], metabolonic lactone sulfate has a high heritability, suggesting a significant genetic component. In IRASFS, highly significant associations spanning the *ZSCAN25-CYP3A5* locus were observed with individual genetic variation explaining 20% of the metabolite trait variance. Among the three correlated variants identified and replicated from GWAS, rs776746 has

a putative functional role. Located in intron 3, rs776746-G is denoted as CYP3A5*3 using the pharmacogenomics star allele nomenclature and creates a cryptic splice site. This resultant aberrantly spliced mRNA includes a premature termination codon resulting in a non-functional protein [36]. Allele frequencies for CYP3A5*3 vary widely from 0.96 in European Americans (CEU) to 0.77 in Mexican Americans (MEX) to 0.31 in African Americans (ASW). GWASs published to date have reported a role for this variant in kidney disease and differential response to immunotherapies. This is the first report to indirectly implicate rs776746 in cardiometabolic disease through metabolomic lactone sulfate as this variant was not directly associated with measures of cardiometabolic health.

Consistent with the genetic findings, transcriptomic analysis in adipose tissue highlighted a significant inverse association between CYP3A5 transcript levels and metabolomic lactone sulfate, i.e. decreased levels of the transcript are associated with increased levels of this metabolite. Cytochrome P450 proteins are mono-oxygenases that catalyze reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipid species [37]. The CYP3A protein family is composed of four members, with CYP3A4 and CYP3A5 levels being highly abundant in the liver and intestine and accounting for the majority of the enzymatic activity in this subfamily [36]. Among the subfamily members, CYP3A5 is also expressed in other tissues including the kidneys, lungs and pancreas, suggesting the regulation of endogenous paracrine or endocrine factors. CYP3A5 is highly polymorphic with 25 allelic variants (numbered *1–*9). Individuals with two copies of CYP3A5*3 (rs776746) are termed non-expressors and do not metabolize CYP3A substrates as rapidly as CYP3A5*1 [38].

Strengths of the current study include untargeted metabolomic profiling of a population-based cohort with multiple traditional and sophisticated measures of adiposity. Inclusion of two ethnic minority cohorts further facilitated generalizability of the findings. Expansion of the phenotypic assessments to address the overall role of adiposity in cardiometabolic health was possible through comprehensive phenotyping, which included measures of glucose homeostasis, lipids and blood pressure. In addition to comprehensive genetic data, transcriptomic analysis of adipose tissue further validated the association of genetic variation with metabolite levels. However, this study is not without limitations. The current metabolite datasets represent untargeted metabolomic profiling using LCMS, which captures only a proportion of the metabolome. This coverage could be enhanced with complementary technologies, e.g. nuclear magnetic resonance (NMR) or gas chromatography mass spectrometry, and the application of an untargeted approach, although, the latter requires significant investment in metabolite identification. In addition, the results described herein were based on cross-sectional data analysis. We posit that levels of metabolomic lactone sulfate will continue to increase over time, consistent with a worsening cardiometabolic profile. Longitudinal studies would not only be able to assess the time course for the increase but also determine whether metabolite levels can be applied prospectively to target therapeutic interventions.

5. CONCLUSIONS

In summary, using untargeted metabolomic profiling in a Mexican American population-based cohort, we have identified a metabolite, tentatively named metabolomic lactone sulfate, that is positively associated with multiple measures of adiposity. Beyond adiposity, this metabolite is associated with multiple phenotypes indicating

worsening cardiometabolic health, i.e. insulin resistance, dyslipidemia and hypertension. These findings were replicated in an African American cohort. Consistent with a strong genetic heritability, metabolite levels were significantly associated with genetic variation at the CYP3A5 locus; variant rs776746 (CYP3A5*3) creates a cryptic splice site resulting in protein truncation. Transcriptional profiling of adipose tissue confirmed that this genetic variant was associated with lower CYP3A5 transcript levels and that lower transcript levels were associated with increased levels of metabolomic lactone sulfate. One interpretation of these results is that the absence of the CYP3A5 protein could result in an unregulated excess of metabolomic lactone sulfate, contributing to worsening cardiometabolic health. Alternatively, this observation could be the result of cardiometabolic disturbances. Further investigation is warranted to determine whether the levels of metabolomic lactone sulfate can prospectively predict worsening metabolic health, thereby facilitating the identification of individuals for earlier targeted interventions.

FUNDING

Grant support for IRASFS was from the National Heart, Lung and Blood Institute (NHLBI; HL060944, HL061019 and HL060919). Data analysis was through the GUARDIAN Study, with support from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; DK085175 and DK118062). The provision of genome-wide array-based data in IRASFS Mexican Americans was supported by NIDDK grants (DK085175 and DK063491) and the National Center for Advancing Translational Science grant (UL1TR001881, Clinical and Translational Science Institute). Acquisition of WES and WGS data was supported by the National Human Genome Research Institute (NHGRI; HG007112). Genome-wide array-based data and metabolomics data in IRASFS African Americans was supported by NIDDK grants (DK105556 and DK066358). Clinical and genomic studies in the AAGMEx cohort were supported by NIDDK grants R01 DK090111 and R01 DK118243 with metabolite data generation supported by NIDDK grants U01 DK105556 and R01 DK066358.

ACKNOWLEDGEMENTS

The authors thank the investigators, staff and participants of the studies for their valuable contributions.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.molmet.2021.101342>.

CONFLICT OF INTEREST

None declared.

REFERENCES

- [1] Hales, C.M., Carroll, M.D., Fryar, C.D., Ogden, C.L., 2020. Prevalence of obesity and severe obesity among adults: United States, 2017–2018. Hyattsville, MD.
- [2] Masters, R.K., Powers, D.A., Link, B.G., 2013. Obesity and US mortality risk over the adult life course. American Journal of Epidemiology 177:431–442. <https://doi.org/10.1093/aje/kws325>.

- [3] Hruby, A., Hu, F.B., 2015. The epidemiology of obesity: a big picture. *Pharmacoeconomics* 33:673–689. <https://doi.org/10.1007/s40273-014-0243-x>.
- [4] Hill, J.O., Wyatt, H.R., Peters, J.C., 2012. Energy balance and obesity. *Circulation* 126:126–132. <https://doi.org/10.1161/CIRCULATIONAHA.111.087213>.
- [5] Newgard, C.B., An, J., Bain, J.R., Muehlbauer, M.J., Stevens, R.D., Lien, L.F., et al., 2009. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metabolism* 9:311–326. <https://doi.org/10.1016/j.cmet.2009.02.002>.
- [6] Cheng, S., Rhee, E.P., Larson, M.G., Lewis, G.D., McCabe, E.L., Shen, D., et al., 2012. Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation* 125:2222–2231. <https://doi.org/10.1161/CIRCULATIONAHA.111.067827>.
- [7] Gaudet, M.M., Falk, R.T., Stevens, R.D., Gunter, M.J., Bain, J.R., Pfeiffer, R.M., et al., 2012. Analysis of serum metabolite profiles in women with endometrial cancer and controls in a population-based case-control study. *Journal of Clinical Endocrinology & Metabolism* 97:3216–3223. <https://doi.org/10.1210/jc.2012-1490>.
- [8] Wang, T.J., Larson, M.G., Vasan, R.S., Cheng, S., Rhee, E.P., McCabe, E., et al., 2011. Metabolite profiles and the risk of developing diabetes. *Nature Medicine* 17:448–453. <https://doi.org/10.1038/nm.2307>.
- [9] Floegel, A., Stefan, N., Yu, Z., Mühlenbruch, K., Drogan, D., Joost, H.G., et al., 2013. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* 62:639–648. <https://doi.org/10.2337/db12-0495>.
- [10] Hales, C.M., Fryar, C.D., Carroll, M.D., Freedman, D.S., Ogden, C.L., 2018. Trends in obesity and severe obesity prevalence in US youth and adults by sex and age, 2007–2008 to 2015–2016. *Journal of the American Medical Association* vol. 319:1723–1725. <https://doi.org/10.1001/jama.2018.3060>.
- [11] Henkin, L., Bergman, R.N., Bowden, D.W., Ellsworth, D.L., Haffner, S.M., Langefeld, C.D., et al., 2003. Genetic epidemiology of insulin resistance and visceral adiposity. The IRAS Family Study design and methods. *Annals of Epidemiology* 13:211–217. S104727970200412X [pii].
- [12] Sharma, N.K., Sajuthi, S.P., Chou, J.W., Calles-Escandon, J., Demons, J., Rogers, S., et al., 2016. Tissue-specific and genetic regulation of insulin sensitivity-associated transcripts in african Americans. *Journal of Clinical Endocrinology & Metabolism* 101:1455–1468. <https://doi.org/10.1210/jc.2015-3336>.
- [13] Box, G.E.P., Cox, D.R., 1964. *An analysis of transformations*. p. 211–52.
- [14] Palmer, N.D., Goodarzi, M.O., Langefeld, C.D., Wang, N., Guo, X., Taylor, K.D., et al., 2015. Genetic variants associated with quantitative glucose homeostasis traits translate to type 2 diabetes in Mexican Americans: the GUARDIAN (genetics underlying diabetes in hispanics) consortium. *Diabetes* 64:1853–1866. <https://doi.org/10.2337/db14-0732>.
- [15] Howie, B.N., Donnelly, P., Marchini, J., 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genetics* 5:e1000529. <https://doi.org/10.1371/journal.pgen.1000529>.
- [16] Gao, C., Tabb, K.L., Dimitrov, L.M., Taylor, K.D., Wang, N., Guo, X., et al., 2018. Exome sequencing identifies genetic variants associated with circulating lipid levels in Mexican Americans: the insulin resistance Atherosclerosis family study (IRASFS). *Scientific Reports* 8:5603. <https://doi.org/10.1038/s41598-018-23727-2>.
- [17] Guan, M., Keaton, J.M., Dimitrov, L., Hicks, P.J., Xu, J., Palmer, N.D., et al., 2019. Genome-wide association study identifies novel loci for type 2 diabetes-attributed end-stage kidney disease in African Americans. *Human Genomics* 13:21. <https://doi.org/10.1186/s40246-019-0205-7>.
- [18] Almasy, L., Blangero, J., 1998. Multipoint quantitative-trait linkage analysis in general pedigrees. *The American Journal of Human Genetics* 62:1198–1211. <https://doi.org/10.1086/301844>.
- [19] Grundy, S.M., Brewer Jr., H.B., Cleeman, J.I., Smith Jr., S.C., Lenfant, C., American Heart, A., et al., 2004. Definition of metabolic syndrome: report of the national Heart, lung, and blood Institute/American Heart association conference on scientific issues related to definition. *Circulation* 109:433–438. <https://doi.org/10.1161/01.CIR.0000111245.75752.C6>.
- [20] D'Agostino, R.B., Vasan, R.S., Pencina, M.J., Wolf, P.A., Cobain, M., Massaro, J.M., et al., 2008. General cardiovascular risk profile for use in primary care. *Circulation* 117:743–753. <https://doi.org/10.1161/circulationaha.107.699579>.
- [21] Zhan, X., Hu, Y., Li, B., Abecasis, G.R., Liu, D.J., 2016. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics* 32:1423–1426. <https://doi.org/10.1093/bioinformatics/btw079>.
- [22] Millstein, J., Chen, G.K., Breton, C.V., 2016. cit: hypothesis testing software for mediation analysis in genomic applications. *Bioinformatics* 32:2364–2365. <https://doi.org/10.1093/bioinformatics/btw135>.
- [23] Wilson, P.W., D'Agostino, R.B., Parise, H., Sullivan, L., Meigs, J.B., 2005. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation* 112:3066–3072. <https://doi.org/10.1161/CIRCULATIONAHA.105.539528>.
- [24] Pouliot, M.C., Despres, J.P., Nadeau, A., Moorjani, S., Prud'Homme, D., Lupien, P.J., et al., 1992. Visceral obesity in men. Associations with glucose tolerance, plasma insulin, and lipoprotein levels. *Diabetes* 41:826–834. <https://doi.org/10.2337/diab.41.7.826>.
- [25] Goodarzi, M.O., Langefeld, C.D., Xiang, A.H., Chen, Y.D., Guo, X., Hanley, A.J., et al., 2014. Insulin sensitivity and insulin clearance are heritable and have strong genetic correlation in Mexican Americans. *Obesity* 22:1157–1164. <https://doi.org/10.1002/oby.20639>.
- [26] Carroll, M.D., Lacher, D.A., Sorlie, P.D., Cleeman, J.I., Gordon, D.J., Wolz, M., et al., 2005. Trends in serum lipids and lipoproteins of adults, 1960–2002. *Journal of the American Medical Association* vol. 294:1773–1781. <https://doi.org/10.1001/jama.294.14.1773>.
- [27] Austin, M.A., Hokanson, J.E., Edwards, K.L., 1998. Hypertriglyceridemia as a cardiovascular risk factor. *The American Journal of Cardiology* 81:7B–12B. [https://doi.org/10.1016/s0002-9149\(98\)00031-9](https://doi.org/10.1016/s0002-9149(98)00031-9).
- [28] Lewington, S., Clarke, R., Qizilbash, N., Peto, R., Collins, R., Prospective Studies, C., et al., 2002. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 360:1903–1913. [https://doi.org/10.1016/s0140-6736\(02\)11911-8](https://doi.org/10.1016/s0140-6736(02)11911-8).
- [29] Gall, W.E., Beebe, K., Lawton, K.A., Adam, K.P., Mitchell, M.W., Nakhle, P.J., et al., 2010. alpha-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. *PLoS One* 5:e10883. <https://doi.org/10.1371/journal.pone.0010883>.
- [30] Cobb, J., Eckhart, A., Motsinger-Reif, A., Carr, B., Groop, L., Ferrannini, E., et al., 2016. Alpha-hydroxybutyric acid is a selective metabolite biomarker of impaired glucose tolerance. *Diabetes Care* 39:988–995. <https://doi.org/10.2337/dc15-2752>.
- [31] Peddinti, G., Cobb, J., Yengo, L., Froguel, P., Kravic, J., Balkau, B., et al., 2017. Early metabolic markers identify potential targets for the prevention of type 2 diabetes. *Diabetologia* 60:1740–1750. <https://doi.org/10.1007/s00125-017-4325-0>.
- [32] di Giuseppe, R., Koch, M., Nothlings, U., Kastenmuller, G., Artati, A., Adamski, J., et al., 2019. Metabolomics signature associated with circulating serum selenoprotein P levels. *Endocrine* 64:486–495. <https://doi.org/10.1007/s12020-018-1816-9>.
- [33] Krumsiek, J., Suhre, K., Evans, A.M., Mitchell, M.W., Mohnen, R.P., Milburn, M.V., et al., 2012. Mining the unknown: a systems approach to metabolite identification combining genetic and metabolic information. *PLoS Genetics* 8:e1003005. <https://doi.org/10.1371/journal.pgen.1003005>.

- [34] Shin, S.Y., Fauman, E.B., Petersen, A.K., Krumsiek, J., Santos, R., Huang, J., et al., 2014. An atlas of genetic influences on human blood metabolites. *Nature Genetics* 46:543–550. <https://doi.org/10.1038/ng.2982>.
- [35] Long, T., Hicks, M., Yu, H.C., Biggs, W.H., Kirkness, E.F., Menni, C., et al., 2017. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nature Genetics* 49:568–578. <https://doi.org/10.1038/ng.3809>.
- [36] Kuehl, P., Zhang, J., Lin, Y., Lamba, J., Assem, M., Schuetz, J., et al., 2001. Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nature Genetics* 27:383–391. <https://doi.org/10.1038/86882>.
- [37] Zanger, U.M., Schwab, M., 2013. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacology & Therapeutics* 138:103–141. <https://doi.org/10.1016/j.pharmthera.2012.12.007>.
- [38] Lamba, J., Hebert, J.M., Schuetz, E.G., Klein, T.E., Altman, R.B., 2012. PharmGKB summary: very important pharmacogene information for CYP3A5. *Pharmacogenetics and Genomics* 22:555–558. <https://doi.org/10.1097/FPC.0b013e328351d47f>.