DR. LISA VANWAGNER (Orcid ID : 0000-0002-6264-2573)

Article type : Original

[HEP4-18-0312]

Insulin resistance exacerbates genetic predisposition to NAFLD in individuals without diabetes

Llilda Barata[1] (barata@wustl.edu), Mary F. Feitosa[1] (mfeitosa@wustl.edu), Lawrence F. Bielak[2] (lfbielak@umich.edu), Brian Halligan[3] (halligan@med.umich.edu), Abigail S. Baldridge[4] (abigail.baldridge@northwestern.edu), Xiuqing Guo[5] (xguo@labiomed.org), Laura M. Yerges-Armstrong[6] (laura.m.yerges-armstrong@gsk.com), Albert V. Smith[7] (albertvs@umich.edu), Jie Yao[5] (jyao@labiomed.org), Nicholette D. Palmer[8] (nallred@wakehealth.edu), Lisa B. VanWagner[4,9] (lvw@northwestern.edu), J. Jeffrey Carr[10] (j.jeffrey.carr@vumc.org), Yii-Der I. Chen[5] (ichen@labiomed.org), Matthew Allison[11] (mallison@ucsd.edu), Matthew J. Budoff[12] (mbudoff@labiomed.org), Samuel K. Handelman[3] (samuelkh@med.umich.edu), Sharon L.R. Kardia[2] (skardia@umich.edu), Thomas H. Mosley Jr.[13] (tmosley@umc.edu), Kathleen Ryan[6] (KRyan@som.umaryland.edu), Tamara B. Harris[14] (harris99@nia.nih.gov), Lenore J. Launer[14] (LaunerL@nia.nih.gov), Vilmundur Gudnason[15,16] (v.gudnason@hjarta.is), Jerome I. Rotter[5] (jrotter@labiomed.org), Myriam Fornage[17](Myriam.Fornage@uth.tmc.edu), Laura J. Rasmussen-Torvik[4] (ljrtorvik@northwestern.edu), Ingrid Borecki[1] (iborecki28@gmail.com), Jeffrey R. O'Connell[6] (joconnel@som.umaryland.edu), Patricia A. Peyser[2] (ppeyser@umich.edu), Elizabeth K. Speliotes*[3] (espeliot@med.umich.edu), Michael A. Province*[1] (mprovince@wustl.edu)

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/HEP4.1353

[1] Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine; St. Louis, MO, USA

[2] Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA

[3] Division of Gastroenterology, Department of Internal Medicine, Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan USA

[4] Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL USA

[5] The Institute for Translational Genomics and Population Sciences, LABioMed and the Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, CA, USA

[6] Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland USA

[7] Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI,USA

[8] Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC, USA

[9] Division of Gastroenterology and Hepatology, Northwestern University Feinberg School of Medicine, Chicago, IL USA

[10] Department of Radiology, Vanderbilt University School of Medicine, Nashville, TN, USA

[11] Department of Family Medicine and Public Health, University of California, San Diego, CA,USA

[12] Division of Cardiology, Los Angeles Biomedical Research Institute, Torrance, California.

[13] Department of Medicine, Division of Geriatrics, University of Mississippi Medical Center, Jackson, MS, USA

[14] Laboratory of Epidemiology and Population Sciences, National Institute of Aging, Bethesda, MD, USA

[15] Icelandic Heart Association, Kopavogur, Iceland

[16] Faculty of Medicine, University of Iceland, Reykjavik, Iceland

[17] University of Texas Health Science Center, Houston, Texas, USA

** Michael Province and Elizabeth K. Speliotes equally supervised this work.*

Keywords: gene-environment interactions, PNPLA3, hepatic steatosis, metabolic factors, GCKR

Corresponding Authors:

Llilda Barata PhD, MPH

Department of Genetics
Division of Statistical Genomics
Washington University School of Medicine
660 South Euclid Ave
St. Louis, MO, 63108 USA
barata@wustl.edu
(314) 747-4133

Elizabeth K. Speliotes, MD PhD MPH

Division of Gastroenterology
Department of Internal Medicine
University of Michigan
1150 West Medical Center Drive
Ann Arbor, MI, 48109 USA
espeliot@med.umich.edu
(734) 936-4785

Michael A. Province, PhD

Department of Genetics
Division of Statistical Genomics
Washington University School of Medicine
660 South Euclid Ave
St. Louis, MO, 63108 USA
mprovince@wustl.edu
(314) 362-3616

Abbreviations: NAFLD= Nonalcoholic fatty liver disease; TG=triglycerides; LDL=low-density

lipoprotein cholesterol; HDL=high-density lipoprotein cholesterol; BMI=body mass index; $WHR_{adj}BMI$= waist-to-hip ratio adjusted for body mass index; *PNPLA3*=Patatin-like phospholipase domain-containing protein 3 gene; *GCKR* =Glucokinase regulatory protein gene; *NCAN* =Neurocan gene; *TM6SF2* =Transmembrane 6 Superfamily Member 2 gene; *LYPLAL1*=Lysophospholipase-like 1 gene; EA=European ancestry; SNP=single nucleotide polymorphism; AA=African ancestry; LA=liver attenuation; HOMA-IR=homeostatic model of insulin resistance; AGES=Age, Gene/Environment Susceptibility-Reykjavik; Amish=Old Order Amish; CARDIA=Coronary Artery Risk Development in Young Adults; FamHS=Family Heart Study; FHS=Framingham Heart Study; GENOA=Genetic Epidemiology Network of Arteriopathy; MESA=Multi-Ethnic Study of Atherosclerosis; HU=Hounsfield units; IVN=inverse normal transformation; LA_{ivn}=Inverse normal-transformed residuals of LA; SD=standard deviation

Financial Support:

This work was performed under the auspices of the Genetics of Obesity-Related Liver Disease Consortium. Funding for this study was made possible by The Age, Gene/Environment Susceptibility-Reykjavik study (AGES) funded by the National Institute of Health (NIH) contracts N01-AG-1-2100 and 271201200022C, the National Institute of Aging Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN:00-063. The researchers are indebted to the participants for their willingness to participate in the study. The Amish study gratefully acknowledges our Amish liaisons, research volunteers, field workers and Amish Research Clinic staff and the extraordinary cooperation and support of the

Amish community without which these studies would not have been possible. The Amish studies are supported by grants and contracts from NIH, including U01 HL072515, U01 HL84756, U01 HL137181 and P30 DK72488. Funding for the Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (HHSN268201300025C & HHSN268201300026C), Northwestern University (HHSN268201300027C), University of Minnesota (HHSN268201300028C), Kaiser Foundation Research Institute (HHSN268201300029C), Johns Hopkins University School of Medicine (HHSN268200900041C) and Vanderbilt University Medical Center (R01 HL 098445). CARDIA is also partially supported by the Intramural Research Program of NIA and an intra-agency agreement between NIA and NHLBI (AG0005). The National Human Genome Research Institute (NHGRI) supported genotyping of CARDIA participants through grants U01-HG-004729, U01-HG-004446, and U01-HG-004424. This manuscript has been reviewed by CARDIA for scientific content. Exome chip genotyping and data analyses were funded in part by grants U01-HG004729, R01-HL093029 and R01-HL084099 from NIH to Dr. Myriam Fornage. Lisa B. VanWagner is supported by NIH grant K23 HL136891. The Family Heart Study (FamHS) was supported by grant R01-DK-089256 from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and grant R01HL117078 from NHLBI. Elizabeth K. Speliotes and Brian Halligan are supported by NIH grants R01 DK106621, R01 DK107904 and The University of Michigan Department of Internal Medicine. Data from the Framingham Heart Study (FHS) came from the database of Genotypes and Phenotypes (dbGaP). Lawrence F. Bielak and Patricia A. Peyser are supported, in part, by NIH grants R01 DK106621 and R01 DK107904. Support for the Genetic Epidemiology Network of Arteriopathy (GENOA) study was provided by NHLBI (HL054457, HL054464, HL054481, HL087660, and HL085571) of the National Institutes of Health. The Multi-Ethnic Study of Atherosclerosis (MESA) and the MESA SHARe project are conducted and supported by NHLBI in collaboration with MESA investigators. This research was supported by R01 HL071739, and MESA was supported by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169 from the National Heart, Lung, and Blood Institute, and by grants UL1-TR-000040, UL1-TR-001079, and UL1-RR-025005 from National Center for Research Resources. The provision of exome chip genotyping data was supported in part by the NHLBI contract N02-HL-64278, National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center

(DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. Funding support for MESA's NAFLD dataset was provided by grant HL071739-05A2. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of NHLBI, NIA, NIDDK or NIH.

ABSTRACT

The accumulation of excess fat in the liver (hepatic steatosis), in the absence of heavy alcohol consumption, causes nonalcoholic fatty liver disease (NAFLD), which has become a global epidemic. Identifying metabolic risk factors that interact with the genetic risk of NAFLD is important for reducing disease burden. We tested whether serum glucose, insulin, insulin resistance, triglycerides, low density lipoprotein cholesterol, high density lipoprotein cholesterol, body mass index (BMI), and waist-to-hip ratio adjusted for BMI interact with genetic variants in or near the patatin-like phospholipase domain containing 3 gene *(PNPLA3)*, the glucokinase regulatory protein gene *(GCKR)*, the neurocan gene (*NCAN/TM6SF2),* and the lysophospholipase-like 1 gene *(LYPLAL1)* to exacerbate hepatic steatosis, estimated by liver attenuation (LA). We performed association analyses in ten population-based cohorts separately and then meta-analyzed results in up to 14,751 individuals (11,870 of European

ancestry and 2,881 of African ancestry). We found that *PNPLA3*-rs738409 significantly interacted with insulin, insulin resistance, BMI, glucose, and TG to increase hepatic steatosis in nondiabetic individuals carrying the G-allele. Additionally, *GCKR*-rs780094 significantly interacted with insulin, insulin resistance and TG. Conditional analyses, using the two largest European ancestry cohorts in the study, showed that insulin levels accounted for most of the interaction of *PNPLA3*-rs738409 with BMI, glucose, and TG in nondiabetic individuals. Insulin, *PNPLA*-rs738409, and their interaction accounted for at least 8% of the variance in hepatic steatosis in these two cohorts.

Conclusion: Our results suggest that insulin resistance, either directly or via the resultant elevated insulin levels, more than other metabolic traits, amplifies the *PNPLA3* rs738409-G genetic risk for hepatic steatosis. These results suggest that improving insulin resistance in nondiabetic individuals carrying *PNPLA3*-rs738409-G may preferentially decrease hepatic steatosis.

Nonalcoholic fatty liver disease (NAFLD) is a result of the excess accumulation of lipids in hepatocytes (hepatic steatosis) in the absence of heavy alcohol consumption(1). Hepatic steatosis is also associated with the risk of developing dyslipidemia or dysglycemia(2), as well as cardiovascular disease, which is the number one cause of death in individuals with NAFLD(3, 4). Hepatic steatosis may progress to advanced liver disease in the form of nonalcoholic steatohepatitis, fibrosis (cirrhosis), and cancer (hepatocellular carcinoma)(5-7). In the U.S., the prevalence of hepatic steatosis in the adult population is between 10% to 30%; worldwide it is 25% to 45%(8). While the pathogenesis of NAFLD is not entirely understood, both genetic factors and metabolic traits increase the risk of hepatic steatosis.

Heritability of hepatic steatosis ranges from 22 to 38% across all ancestries suggesting that specific genotypes may predispose individuals to NAFLD(1). Previously, the Genetics of Obesity-Related Liver Disease Consortium conducted a genome-wide association study in 7,176 individuals of European ancestry (EA) with replication in histology-based samples (9). This study identified that rs738409 (*PNPLA3*), a missense single nucleotide polymorphism (SNP) first associated with hepatic fat content a decade ago (10), the missense variant rs2228603 (*NCAN/TM6SF2*) and intronic variants rs12137855 (*LYPLAL1*) and rs780094 (*GCKR*) were significantly associated with hepatic steatosis(9). We and others have replicated the association of these common variants with hepatic steatosis in other populations and ethnicities (11-13), and the associations are consistent between those of EA and African

ancestry (AA) (direction of effect is similar)(11). Further, the G allele for rs738409 was associated with susceptibility to nonalcoholic steatohepatitis (OR 2.64, 95% CI: 1.85-3.75, p≤ 1.0E-04), nonalcoholic steatohepatitis severity (OR 1.85, 95% CI: 1.05-3.26, p≤ 3.5E-02) and fibrosis (OR 1.95, 95% CI: 1.17-3.26, p≤ 1.3E-02) in EA individuals(14).

Traits that predispose to metabolic syndrome, i.e. higher body mass index (BMI) (15), dyslipidemia, hyperglycemia, and insulin resistance are associated with hepatic steatosis (2, 3, 16). Eighty to ninety percent of obese (BMI ≥ 30 kg/m^2) adults have hepatic steatosis(17), while 20-80% of individuals with hepatic steatosis also have higher levels of triglyceride (TG) and low-density lipoprotein cholesterol (LDL), but lower levels of high-density lipoprotein cholesterol (HDL)(18). Diabetes is also commonly associated with hepatic steatosis(19). How these modifiable metabolic traits interact with genetic variation to influence risk for hepatic steatosis is not known.

In this cross-sectional study, we tested whether several metabolic traits interact with the four genetic variants previously associated with hepatic steatosis(9) to affect liver attenuation (LA), a computed tomographic quantitative measure that is inversely related to histologically measured liver fat (20). The metabolic traits tested were: insulin resistance (as homeostatic model of insulin resistance (HOMA-IR)), fasting insulin, fasting glucose, BMI, centralized fat deposition measured by waist-to-hip ratio adjusted for BMI ($WHR_{adj}BMI$), fasting TG, fasting HDL and fasting LDL. We first carried out interaction analyses between each of these traits and each of the genetic variants in ten separate population-based cohorts from seven different studies. Then we meta-analyzed results across cohorts in up to 14,751 individuals (EA, n=11,870 and AA, n=2,881). We then carried out conditional analyses in the two largest EA cohorts in the study to determine the driving metabolic factor.

POPULATION AND METHODS

Ethics Statement

The Institutional Review Boards or equivalent committees of all participating studies approved this study. The principal investigator of each institution obtained written consent from participants.

Study Description

The study was comprised of up to 14,751 individuals (EA, n=11,870 and AA, n=2,881); 56% of participants were female. The sample derived from seven population-based studies participating in the Genetics of Obesity-Related Liver Disease Consortium: Age, Gene/Environment Susceptibility-Reykjavik (AGES), Old Order Amish (Amish), Coronary Artery Risk Development in Young Adults (CARDIA), Family Heart Study (FamHS), Framingham Heart Study (FHS), Genetic Epidemiology Network of Arteriopathy (GENOA), and Multi-Ethnic Study of Atherosclerosis (MESA). In total, ten cohorts were included in the analysis, as three studies contributed two ethnic groups (AA, EA). Each ethnic group was analyzed separately. CARDIA, MESA, and AGES have unrelated individuals while FHS, Amish, GENOA, and FamHS are family-based. Detailed information about the characteristics and design of each study is provided in **Supplementary Table 1**.

Outcome variable and metabolic traits

The outcome variable was LA (liver attenuation), measured non-invasively with computed tomography in Hounsfield units (HU) (21). LA is inversely proportional to liver fat, i.e. lower LA values indicate a higher fat content in the liver (more hepatic steatosis)(2). The procedures followed by each cohort to measure LA are described in **Supplementary Table 2**. Individuals with active malignancies, focal lesions, or other incidental findings on computed tomography were excluded from the studies.

Metabolic traits of interest were harmonized across cohorts following standard clinical definitions. Overall adiposity was characterized by BMI (kg/m^2), and abdominal adiposity by waist-to-hip ratio adjusted for BMI ($WHR_{adj}BMI$, cm). Since waist-to-hip ratio is correlated with both BMI and visceral fat, we chose to use $WHR_{adj}BMI$ to have a measure that is independent of overall fatness (i.e. BMI), but does reflect visceral adiposity, and is easily measured in the clinic. Fasting insulin (mU/L) and fasting glucose (mmol/L) were measured from plasma or serum using standard laboratory techniques detailed in **Supplementary Table 2**. When fasting glucose was measured from whole blood, it was converted to plasma glucose using a correction factor of 1.13 (22). HOMA-IR was assessed using fasting glucose (mmol/L) x fasting insulin (mU/L) divided by 22.5 (23). Each cohort assayed fasting TG (mg/dL) and fasting HDL (mg/dL) using methods described in **Supplementary Table 2**. If fasting LDL (mg/dL) was assayed, it was used. Otherwise, LDL was calculated using the Friedewald formula, LDL_F =(Total cholesterol(mg/dL) - HDL(mg/dL) -TG(mg/dL)/5.0), only if TG < 400 mg/dL (24).

Alcohol consumption, history of diabetes, and use of lipid lowering medications were acquired by questionnaire. Total alcohol consumption, defined in drinks per week, was calculated from daily intake of beer, wine, and spirits. One drink was defined as a serving of 14 grams of ethanol, the same as a 12 oz. bottle or can of beer, 5 oz. glass of wine, or 1.5 oz. shot of 80-proof spirits such as gin, vodka, or whiskey(25). Heavy drinking was defined as ≥ 8 drinks per week for women and ≥ 15 drinks per week for men (26). Diabetes (Type 1, Type 2) was defined as having fasting plasma glucose levels ≥ 7 mmol/L (126 mg/dL), or self-reporting the use of insulin or oral antidiabetic medications, or having a physician diagnosis of diabetes. The use of statins was assessed from medication questionnaires.

Genotyping and Imputation

Four common variants were included in the analyses: rs738409 - a missense variant in the patatin-like phospholipase domain containing 3 gene (*PNPLA3*); rs780094, an intronic variant within the glucokinase regulatory protein gene (*GCKR*) that is in high linkage disequilibrium (r^2=0.93) with rs1260326, a likely functional missense variant in this gene; rs2228603, a missense variant in the neurocan gene (*NCAN*) that is in high linkage disequilibrium (r^2=0.798) with rs585422926, a likely functional missense variant in the transmembrane 6 Superfamily Member 2 gene (*TM6SF2)*; and rs12137855, an intronic variant in the lysophospholipase-like 1 gene (*LYPLAL1*). These variants were either directly genotyped (allele counts were coded 0, 1, or 2), or dosages were imputed from HapMap II or 1000G. Genotype calling algorithms and imputation methods are detailed in **Supplementary Table 3**.

STATISTICAL ANALYSIS

Cohort-specific analyses

Cohorts performed analyses separately in each ancestry group (EA, AA). LA and metabolic traits, used as continuous variables in all analyses, were adjusted for sex, age, principal component estimates of ancestry, and study-specific covariates using linear regression as detailed in **Supplementary Table 2**. LA was also adjusted for alcohol consumption, a continuous variable (drinks/week), and for scan penetrance using phantom or spleen density. Residuals from adjusted LA and metabolic traits were transformed using inverse normal

transformation (IVN) to reduce the influence of outliers and to standardize the phenotypes across cohorts. Inverse normal-transformed residuals of LA, (LA_{ivn}), and each metabolic trait (MT_{ivn}) were used to fit the interaction models.

Each cohort tested for statistical interactions between each variant and each metabolic trait using multivariable linear regression or mixed linear modeling. LA_{ivn} was the dependent variable. The independent variables were each SNP and MT_{ivn}, plus the interaction: $LA_{ivn} = \alpha + \beta_1 (SNP) + \beta_2 (MT_{ivn}) + \beta_3 (SNP \times MT_{ivn}) + \epsilon$. An additive model of inheritance was assumed. Studies with family data (FHS, GENOA, Amish, and FamHS) used linear mixed models to account for family relatedness among participants and computed robust standard errors. Participants with diabetes (Type1 and Type 2) were excluded from the insulin, glucose and HOMA-IR models, and those taking statins were excluded from the LDL model. As a secondary analysis, BMI was included as a covariate in the models to investigate whether the effect of the interaction between each SNP and each metabolic trait on LA_{ivn} occurred independent of overall adiposity. Associations were carried out using MMAP(27), R(28), and SAS (29) software.

Meta-analyses

We conducted fixed-effects meta-analyses by ancestry and overall on the parameter estimates (β-coefficients and standard errors) for the main effects and interaction effects. We utilized the inverse variance weighting method implemented in METAL (30). Using Cochran's Q test (31), we tested for heterogeneity of effects across all analyses. Within ancestries, focusing on interactions, we found evidence of heterogeneity only for the interaction between TG and *GCKR* in the EA cohorts. We did not find any heterogeneity for the interaction in the meta-analyses between the two ancestry groups (EA vs AA); thus, we report the combined ancestry meta-analyses. To determine the level of statistical significance while accounting for multiple testing, we applied a Bonferroni correction that consisted of grouping correlated traits into three metabolic domains: insulin-glucose, adiposity, and lipids. The critical p-value α=0.05 was divided by 12 (4 variants x 3 metabolic domains) to obtain a corrected p-value. Meta-analyses results and heterogeneity tests were considered significant if the two-tailed p-value was ≤ 4.17E-03. As a secondary analysis, to investigate whether the statistically significant interactions were consistent between genders, we fit the interaction models in men and women separately, and meta-analyzed results within gender.

Conditional Analyses in FamHS and FHS

To determine whether the interaction of BMI, glucose or TG with *PNPLA3*-rs738409 was independent of insulin, we analyzed each trait's interaction effect before and after including insulin in the model. The analyses were performed with EA individuals in FamHS and replicated in FHS. We chose these two cohorts because they are the two largest cohorts in the study; together they represent more than 1/3 of our total sample. Individuals with diabetes and/or missing information for the metabolic traits of interest were excluded resulting in a sample of 2,280 individuals in FamHS and 2,581 in FHS. After adjusting LA for phantom in both cohorts, and for field centers in FamHS, LA residuals were transformed using inverse normal transformation to approximate normality. LA transformed residuals (LA_{ivn}) were used as the dependent variable. Using linear mixed models, we first regressed LA_{ivn} on either BMI, glucose, or TG, and their interaction with *PNPLA3*-rs738409 **(Supplementary Text)**. We then added insulin to the models and its interaction with *PNPLA3*-rs738409 and the metabolic trait (either BMI, glucose, or TG). Insulin and TG were log-transformed due to the presence of influential outliers. Models were adjusted for age, sex, and alcohol consumption (drinks/week), and for genotype batch effects in FamHS. Results from conditional analyses in each cohort were then meta-analyzed.

The conditional models included principal components to adjust for population stratification. Because the principal components were not associated with LA_{ivn} in either cohort, and their inclusion in the conditional models did not change the inferences, we present the models without them. We also performed conditional analyses after excluding individuals from FamHS (n=231), and FHS (n=371) who reported heavy alcohol use (≥ 8 drinks per week for women, and ≥ 15 drinks per week for men **(Supplementary Tables 10-12)** (26). Since the inferences were unchanged, to increase power, we included all individuals, and adjusted for alcohol as a covariate. Additionally, we conducted the conditional analyses with log-transformed HOMA-IR instead of log-transformed insulin **(Supplementary Tables 13-15)**. Insulin and HOMA-IR provided similar inferences. Because glucose explains significantly less of the variation in LA_{inv}, we focused on insulin over HOMA-IR since there was no added benefit of measuring glucose on variance explained by HOMA-IR than with just measuring insulin.

Illustration in FamHS of the interaction between insulin and PNPLA3-rs738409 in individuals without diabetes

To assess the interaction effect of insulin with *PNPLA3*-rs738409 on hepatic steatosis prevalence in FamHS, we plotted the percentage of individuals with LA ≤ 60 HU per *PNPLA3*-rs738409 genotype by the lowest and highest quartile of insulin. Individuals with diabetes and/or missing information for insulin were excluded and ancestries were combined to obtain a sample of n=2,725. LA and insulin were not adjusted or transformed. The LA cut point of ≤ 60 HU, which corresponds to a liver/spleen ratio of 1.1, has previously been shown to identify individuals with moderate to severe macrovesicular steatosis (≥ 30% of the liver parenchyma with fat) at histology with a high diagnostic accuracy (32). In the literature, ≥ 30% liver fat suggests moderate to severe hepatic steatosis (33).

RESULTS

Demographics and clinical characteristics across the study cohorts are presented in **Table 1**. The mean age ± standard deviation (SD) across cohorts ranged from 49.47±3.86 to 76.38±5.46 years old. All cohorts included more women than men. The mean±SD of LA across cohorts ranged from 55.05±12.28 HU to 65.40±9.83 HU. Mean±SD of fasting insulin levels in non-diabetics ranged from 8.30±5.73 to 13.02±10.22 mU/L and fasting blood glucose levels ranged from 4.90±0.58 to 5.49±0.50 mmol/L. The lowest mean±SD for HOMA-IR in non-diabetics was 1.99±1.27 and the highest was 3.14±2.69. The mean±SD of BMI ranged from 27.00±4.49 to 32.71±7.37 kg/m^2. Several cohorts reported mean fasting TG >100 mg/dL. Mean±SD for fasting LDL cholesterol in non-statin users was borderline high in Amish (141.31±8.66 mg/dL) and AGES (146.84±5.73 mg/dL). Across cohorts, the range of fasting HDL was within the recommended limit of ≥ 40 mg/dL. Heavy drinking varied among studies with GENOA having the lowest percentage (0%) and CARDIA the highest (37%).

PNPLA3-rs738409 and GCKR-rs780094 interact with several metabolic traits

We found significant interactions for *PNPLA3*-rs738409 and *GCKR-rs780094* with several metabolic traits in combined ancestries after adjusting for multiple comparisons (**Table 2, Supplementary Table 4).** *PNPLA3*-rs738409 interacted with insulin (p= 4.79E-14), HOMA-IR (p= 4.68E-15), glucose (p= 1.26E-03), BMI (p= 8.13E-08) and TG (p=2.95E-03). As each of these metabolic traits increased, a decrease in LA_{ivn} (i.e. higher fat content in the liver) became more pronounced in presence of the G allele at *PNPLA3*-rs738409 as compared to the presence of the C allele. Additionally, *GCKR*-rs780094 interacted with insulin (p= 4.57E-04), HOMA-IR (p= 1.32E-03), and TG (p= 4.17E-03). As levels of insulin, HOMA-IR, and TG increased, a decrease in LA_{ivn} (i.e. higher fat content in the liver) became more pronounced in

the presence of the T allele at *GCKR*-rs780094, compared to the C allele. All interactions remained significant after adjusting for BMI **(Supplementary Table 5)** suggesting that overall adiposity did not alter these effects. We did not find evidence of significant interactions between any of the four genetic variants and WHR_{adj}BMI, LDL, or HDL. Although the interaction between WHR_{adj}BMI and *PNPLA3* did not reach the Bonferroni significance level, it was borderline significant. This suggests that a larger sample size may be needed to detect an interaction. Alternatively, the lack of statistical significance could be because WHR_{adj}BMI does not represent overall fatness to the extent that BMI or other anthropometric measurements do.

We also carried out meta-analyses in men and women separately to investigate possible gender differences focusing only on the statistically significant interactions with *PNPLA3*-rs738409 and *GCKR*-rs780094 **(Supplementary Table 6)**. Women made up 56% of our study sample. The interaction effects of insulin and HOMA-IR with *PNPLA3*-rs738409 did not differ between men and women, and both reached statistical significance (women= p=3.24E-11, men=7.24E-05; and women: p=1.62E-11, men: p=2.88E-05, respectively). For glucose, the interaction effect was slightly less in men than in women (beta smaller), and did not reach significance in men. These results suggest that gender did not alter the interactions between *PNPLA3*-rs738409 and insulin/HOMAIR and the interaction effect of glucose was still present only in women in the present study. Further, the interaction effects of BMI with *PNPLA3*-rs738409 were similar between men and women, and reached significance in both (p=1.20E-03 and p=3.39E-05, respectively). The interaction effect of TG with PNPLA3- rs738409 did not reach statistical significance in either gender. Moreover, the interaction effects of both insulin and HOMA-IR with *GCKR*-rs780094 reached significance only in women (p=1.02E-03 and p= 6.46E-04, respectively). Similarly, the interaction of TG with *GCKR*-rs780094 was significant only in women (p=8.71E-04). Stratifying by gender substantially reduced our sample size, and as a result power.

Conditional analyses suggest that insulin may mediate the interaction effect of BMI, TG and glucose on LA_{ivn} in individuals without diabetes

We observed that the interaction of insulin with *PNPLA3*-rs738409 had a greater effect on LA_{ivn} (hepatic steatosis defined by liver attenuation) than that of BMI, TG, or glucose. To determine if the interaction of BMI, TG, or glucose with *PNPLA3*-rs738409 was independent of insulin, we carried out conditional analyses in FamHS and FHS, and meta-analyzed results. We found that the interaction of BMI (p=7.57E-02), TG (p=3.49E-01), or glucose (p=9.09E-01) with *PNPLA3*-

rs738409 was no longer statistically significant after including insulin as a main effect and interactor with *PNPLA3*-rs738409 and the respective metabolic trait in the models (**Supplementary Tables 7-9)**. In contrast, the interaction of insulin with *PNPLA3*-rs738409 remained significant after controlling for BMI, TG, or glucose ($p_{insulin\text{-}BMI}$= 4.04E-04; $p_{insulin\text{-}TG}$= 3.24E-06; $p_{insulin\text{-}glucose}$= 8.40E-08), although the effect sizes and p-values were attenuated. These results suggest that insulin may account for most of the interaction effect of BMI, glucose, and TG with *PNPLA3*-rs738409 on LA_{ivn}. Previously, we reported that *PNPLA3*-rs738409 explained 2.4% of the variance in hepatic steatosis, estimated by LA, in EA individuals (11). In the present study, *PNPLA3*-rs738409, insulin and their interaction together explain as much as 8% of the variance in hepatic steatosis in the two largest EA cohorts excluding individuals with diagnosed diabetes. This suggests that insulin levels/insulin resistance may be a key contributor to NAFLD. Excluding heavy drinkers from the conditional analyses did not change our inferences regarding *PNPLA3*-rs738409 **(Supplementary Table 10-12)**. We were not powered to carry out these analyses for *GCKR*-rs780094.

Interaction effect of insulin with PNPLA3 on hepatic steatosis prevalence in FamHS

We also assessed the interaction effect of insulin with *PNPLA3-rs738409* on hepatic steatosis prevalence in individuals without diabetes **(Figure 1)**. In the lowest quartile of insulin levels (≤ 5.20 mU/L), the percentage of individuals with ≥ 30% liver fat (i.e. moderate to severe hepatic steatosis) was 23.42%, 35.81%, and 39.47% for CC, CG and GG individuals, respectively. In the highest quartile of insulin levels (≥ 13.06 mU/L), the percentage of individuals with ≥ 30% liver fat was 54.44%, 76.32% and 95.29% for CC, CG and GG individuals, respectively. The data show that as insulin levels increase the percentage of individuals with moderate to severe hepatic steatosis increases. However, among those with the GG genotype, this effect is magnified. The difference in the percentage of individuals with moderate to severe hepatic steatosis increases by 55 percentage points between the lowest and highest insulin quartiles among those with GG genotype, and increases by 41 percentage points among heterozygotes, while that difference increases only by 31 percentage points among those with the CC genotype. These data suggest that insulin has a strong effect on exacerbating the accumulation of liver fat in individuals without diabetes who have 1 or 2 G- alleles at *PNPLA3*-rs738409.

DISCUSSION

In a sample of 14,751 EA and AA individuals, we found interactions between *PNPLA3*-rs738409 and insulin, HOMA-IR, BMI, glucose, and TG on LA_{inv} (hepatic steatosis) after adjusting for

differences in age, sex, and alcohol consumption. We also found interactions between *GCKR*-rs780094 and insulin, HOMA-IR, and TG on LA_{inv}. Conditional analyses in more than 5,000 EA individuals suggest that insulin, more than glucose, BMI, or TG drive the interaction with *PNPLA3*-rs738409 to affect LA_{inv} in non-diabetics. We did not see significant interactions between *PNPLA3*-rs738409 and BMI, TG or glucose once insulin was accounted for, whereas the reverse was not true. That is, there was still evidence for an interaction between *PNPLA3*-rs738409 and insulin even after accounting for the other metabolic traits. These results persist after accounting for alcohol intake, gender and overall adiposity. We estimated in FamHS and FHS that as much as 8% of the variance in hepatic steatosis is explained by *PNLPA3*-rs738409, insulin and their interaction in non-diabetic EA individuals. In our previous study, *PNPLA3*-rs738409 alone explained only 2.4% of hepatic steatosis variance in EA individuals (11).

Our findings suggest that non-diabetic individuals with *PNPLA3*-rs738409-G and high insulin levels may have a particularly high risk for hepatic steatosis. The *PNPLA3* gene encodes adiponutrin, an enzyme found on the membrane of lipid droplets within hepatocytes (34). Its function may be to break down TG stored in the droplets, helping regulate hepatic TG content (34, 35). The missense polymorphism rs738409 (C > G) in *PNPLA3* substitutes the amino acid isoleucine for methionine at residue 148 (I148M), changing the configuration of adiponutrin's catalytic site, and rendering the enzyme inactive (10, 36). The accumulation of the inactive enzyme on lipid droplets is associated with TG buildup in hepatocytes (36). Humans and mice carrying one or two copies of the I148M mutation (rs738409 CG or GG genotype) accumulate excess TG in lipid droplets, and show more pronounced hepatic steatosis and NAFLD than those without the mutation(35, 36).

It is possible that having high insulin levels in addition to the *PNPLA3-rs738409* G allele may result in a strong synergistic effect that exacerbates the accumulation of fat in the liver of non-diabetic individuals, predisposing them to NAFLD. Insulin resistance stimulates the hydrolysis of TG in adipose tissue releasing fatty acids in the bloodstream, which are taken up by the liver in an unregulated manner promoting the accumulation of TG in hepatocytes (37). Higher insulin levels also activate fatty acid synthesis in the liver further driving the formation and storage of TG (34). In addition, insulin resistance elevates plasma glucose, which is sequestered by the liver, phosphorylated, and metabolized to make glycerol and acetyl-CoA, the building blocks for the synthesis of TG (34,38). In this context, it is possible that increased lipid synthesis and fatty acid delivery to the liver may combine with the inability of hepatocytes to dispose of TG from

lipid droplets, due to the presence of *PNPLA3*-rs738408-G, and lead to increased hepatic steatosis. High insulin levels and *PNPLA3*-rs738409-G may also be involved in molecular feedback loops that increase hepatic steatosis. Insulin resistance and increased insulin levels augment the activity of transcription factors such as SREBP-1c (39). These transcription factors may promote TG synthesis in the liver and up-regulate the expression of *PNPLA3* I148M by binding to its promoter in a positive feedback loop (39). In this way, insulin and *PNPLA3* I148M may synergize to promote hepatic steatosis. This conjecture is also consistent with the enhanced risk of steatosis and liver damage as evident by elevated liver enzymes and liver fat content seen with liver directed long-acting insulin analogues in type 2 diabetics carrying the *PNPLA-3* variant (40).

When taken together, results show evidence that insulin and *PNPLA3*-rs738409 interact to have an important role in hepatic steatosis, and as a result NAFLD. Consequently, lowering the risk of hepatic steatosis and its liver complications in individuals with *PNPLA3*-rs738409-G may be achieved by reducing insulin resistance and concomitant high levels of insulin. One way to accomplish this could be through lifestyle changes that include increased exercise, weight loss, and better nutrition (41). For example, decreasing exposure to carbohydrate rich diets, which adversely increase insulin levels, may mitigate risk (42, 43). Also, treatments that target insulin resistance may be of greater benefit for preventing or treating hepatic steatosis than drugs that simply lower glucose. For example, insulin sensitizing medications such as pioglitazone may be an option; it has already been shown to improve NAFLD, although at the expense of weight gain (44). More studies are warranted to better understand the effect of the relationship between insulin levels and *PNPLA3*-rs738409-G on hepatic steatosis in different populations.

We also observed significant interactions of *PNPLA3*-rs738409 with BMI, glucose, and TG. Our results support the findings of Stender et al. who reported that high BMI augmented the effect of *PNPLA3*-rs738409-G on hepatic steatosis conferring susceptibility to NAFLD (45). Graff et al. also showed an interaction effect between *PNPLA3*-rs738409 and visceral fat content, a measure of metabolic dysfunction (46). However, we found that the effect of BMI in exacerbating hepatic steatosis in the presence of *PNPLA3*-rs738409-G is attenuated by controlling for insulin levels in the model. We made the same observation for glucose and TG suggesting that insulin/insulin resistance in the presence of *PNPLA3*-rs738409-G may confer most of the risk for hepatic steatosis on its own or through other metabolic intermediates.

Studies have reported an association between LDL and hepatic steatosis (47, 48). However, our study did not find an interaction between any of the genetic variants considered and LDL. This suggests that for individuals carrying *PNPLA3*-rs738409-G, reducing insulin levels or insulin resistance may have a greater effect on reducing the risk of hepatic steatosis than reducing LDL.

In addition to *PNPLA3*, we found that *GCKR* interacts with insulin resistance to increase susceptibility to hepatic steatosis. *GCKR* encodes the glucokinase regulatory protein, which has an important role in glucose metabolism(49). The glucokinase regulatory protein binds to the glucose metabolizing enzyme, glucokinase, to inhibit its role in the uptake and storage of dietary glucose via stimulating de novo lipogenesis(49). The variant rs780094/rs12060326 in the glucokinase regulatory protein reduces its ability to inhibit glucokinase (49). This results in an increased activity of glucokinase in the liver, which promotes de novo lipogenesis. When this mutation is combined with insulin resistance, it may amplify de novo lipogenesis to promote hepatic steatosis. We did not replicate the interaction between *TM6SF2* and BMI reported by Stender et al. (45); however, our results show a similar trend. The interaction was borderline non-significant in the combined ancestry meta-analyses (B_{int}= -0.05, p=5.89E-02). Some differences between Stender et al. and this study may explain why we did not detect a statistically significant interaction. First, Stender et al. used proton magnetic resonance spectrometry to measure steatosis, which is a more sensitive measure than computed tomography. Second, they used the genotyped missense variant, rs58542926; we used the proxy, imputed variant, rs2228603. The two variants are in high linkage disequilibrium (D'=0.926, r^2=0.798). Third, Stender et al. combined the heterozygotes (EK), and homozygotes (KK), and compared them to those without the risk allele (EE). These three differences may have increased their power to see the weak effect they reported.

Our study has several limitations. It is a cross-sectional design that cannot prove temporal causality of insulin exposure on increasing hepatic steatosis. Because we used population-based cohorts that lacked biopsy information, we do not know whether we included individuals with advanced stages of NAFLD such as nonalcoholic steatohepatitis, fibrosis, or cirrhosis. We also could not differentiate peripheral insulin resistance from hepatic insulin resistance with our data. Moreover, even though in euglycemic individuals HOMA-IR was highly correlated to a single value of insulin (r^2=0.98), we do not have direct measures of dynamic glucose regulation. Therefore, functional studies are needed to gain more insight into the biological processes

driving our observations. Finally, our study did not include the genetic variant *MBOAT7* (rs641738), which has been associated with hepatic fat accumulation (50). In our prior association analyses (11), we did not see an association between *MBOAT7* and LA (Beta= -0.03, p=0.15). Because our inclusion criteria for variants was that they needed to be associated with LA, and we could not substantiate the association of *MBOAT7* in our sample, we excluded it.

In conclusion, to our knowledge, this is the largest study examining the interaction between multiple metabolic traits and four genetic variants on hepatic steatosis in multiple cohorts representing two different ancestry groups. Our findings suggest that insulin levels/insulin resistance more than other correlated metabolic traits including glucose, TG, and BMI interact with genetic variants in *PNPLA3* to promote hepatic steatosis. Through conditional analyses, we show that insulin levels explain the interactions observed between *PNPLA3*-rs738409 and BMI, as well as the interactions between *PNPLA3*-rs738409 and glucose and TG, in almost 5,000 nondiabetic, EA individuals. Our work suggests that improving insulin resistance and reducing insulin levels in pre-diabetic individuals carrying fatty liver promoting alleles at *PNPLA3*-rs738409 may offer preferential benefit and mitigate their risk of developing NAFLD. Although *PNPLA3* genotype information is not currently used to make clinical decisions, it may be helpful in the future not only to risk stratify individuals, but also to tailor their treatment. Our work contributes to the understanding of the pathophysiology of NAFLD, and informs further interventional research to better diagnose and/or treat individuals with increased risk of NAFLD.

REFERENCES

1. Kahali B, Halligan B, Speliotes EK. Insights from Genome-Wide Association Analyses of Nonalcoholic Fatty Liver Disease. Semin Liver Dis 2015;35:375-391.

2. Speliotes EK, Massaro JM, Hoffmann U, Vasan RS, Meigs JB, Sahani DV, et al. Fatty liver is associated with dyslipidemia and dysglycemia independent of visceral fat: the Framingham Heart Study. Hepatology 2010;51:1979-1987.

3. Gaggini M, Morelli M, Buzzigoli E, DeFronzo RA, Bugianesi E, Gastaldelli A. Non-alcoholic fatty liver disease (NAFLD) and its connection with insulin resistance, dyslipidemia, atherosclerosis and coronary heart disease. Nutrients 2013;5:1544-1560.

4. Hassan K, Bhalla V, El Regal ME, HH AK. Nonalcoholic fatty liver disease: a comprehensive review of a growing epidemic. World J Gastroenterol 2014;20:12082-12101.

5. McPherson S, Hardy T, Henderson E, Burt AD, Day CP, Anstee QM. Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: implications for prognosis and clinical management. J Hepatol 2015;62:1148-1155.

6. Pais R, Charlotte F, Fedchuk L, Bedossa P, Lebray P, Poynard T, et al. A systematic review of follow-up biopsies reveals disease progression in patients with non-alcoholic fatty liver. J Hepatol 2013;59:550-556.

7. Wong VW, Wong GL, Choi PC, Chan AW, Li MK, Chan HY et al. Disease progression of non-alcoholic fatty liver disease: a prospective study with paired liver biopsies at 3 years. Gut 2010;59:969-974.

8. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 2016;64:73-84.

9. **Speliotes EK, Yerges-Armstrong LM, Wu J, Hernaez R, Kim LJ**, Palmer CD, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS Genet 2011;7:e1001324.

10. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40:1461-1465.

11. Palmer ND, Musani SK, Yerges-Armstrong LM, Feitosa MF, Bielak LF, Hernaez R, et al. Characterization of European ancestry nonalcoholic fatty liver disease-associated variants in individuals of African and Hispanic descent. Hepatology 2013;58:966-975.

12. Hernaez R, McLean J, Lazo M, Brancati FL, Hirschhorn JN, Borecki IB, Harris TB, et al. Association between variants in or near PNPLA3, GCKR, and PPP1R3B with ultrasound-defined steatosis based on data from the third National Health and Nutrition Examination Survey. Clin Gastroenterol Hepatol 2013;11:1183-1190 e1182.

13. Lin YC, Chang PF, Chang MH, Ni YH. Genetic variants in GCKR and PNPLA3 confer susceptibility to nonalcoholic fatty liver disease in obese individuals. Am J Clin Nutr 2014;99:869-874.
14. Liu YL, Patman GL, Leathart JB, Piguet AC, Burt AD, Dufour JF, et al. Carriage of the PNPLA3 rs738409 C >G polymorphism confers an increased risk of non-alcoholic fatty liver disease associated hepatocellular carcinoma. J Hepatol 2014;61:75-81.
15. **Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH**, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197-206.
16. Chatrath H, Vuppalanchi R, Chalasani N. Dyslipidemia in patients with nonalcoholic fatty liver disease. Semin Liver Dis 2012;32:22-29.
17. Bellentani S, Scaglioni F, Marino M, Bedogni G. Epidemiology of non-alcoholic fatty liver disease. Dig Dis 2010;28:155-161.
18. Zhang QQ, Lu LG. Nonalcoholic Fatty Liver Disease: Dyslipidemia, Risk for Cardiovascular Complications, and Treatment Strategy. J Clin Transl Hepatol 2015;3:78-84.
19. Petaja EM, Yki-Jarvinen H. Definitions of Normal Liver Fat and the Association of Insulin Sensitivity with Acquired and Genetic NAFLD-A Systematic Review. Int J Mol Sci 2016;17.
20. Limanond P, Raman SS, Lassman C, Sayre J, Ghobrial RM, Busuttil RW, et al. Macrovesicular hepatic steatosis in living related liver donors: correlation between CT and histologic findings. Radiology 2004;230:276-280.
21. Graffy PM, Pickhardt PJ. Quantification of hepatic and visceral fat by CT and MR imaging: relevance to the obesity epidemic, metabolic syndrome and NAFLD. Br J Radiol 2016;89:20151024.
22. D'Orazio P, Burnett RW, Fogh-Andersen N, Jacobs E, Kuwa K, Kulpmann WR, et al. Approved IFCC recommendation on reporting results for blood glucose (abbreviated). Clin Chem 2005;51:1573-1576.
23. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985;28:412-419.
24. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18:499-502.
25. National Institute of Alcohol Abuse and Alcoholism . What is a Standard Drink? Available at https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/what-standard-drink. Accessed 2017

26. USDA. Dietary Guidelines for Americans 2015-2020 (Appendix 9: Alcohol). Available at https://health.gov/dietaryguidelines/2015/guidelines/appendix-9/ (Accessed 2017).
27. O'Connell J. MMAP:Mixed Model Analysis for Pedigrees and Populations. Available: https//mmap.github.io (Accessed 4 April, 2017). In; 2013.
28. R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 URL http://www.R-project.org/
29. SAS Institute Inc; SAS Software, Version 9.4, Cary, NC: SAS Institute Inc. 2011.
30. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190-2191.
31. Cochran WG. The combination of estimates from different experiments. Biometrics 1954;10:101-129.
32. Park SH, Kim PN, Kim KW, Lee SW, Yoon SE, Park SW et al. Macrovesicular hepatic steatosis in living liver donors: use of CT for quantitative and qualitative assessment. Radiology 2006;239:105-112.
33. Zeb I, Li D, Nasir K, Katz R, Larijani VN, Budoff MJ. Computed tomography scans in the evaluation of fatty liver disease in a population based study: the multi-ethnic study of atherosclerosis. Acad Radiol 2012;19:811-818.
34. Chamoun Z, Vacca F, Parton RG, Gruenberg J. PNPLA3/adiponutrin functions in lipid droplet formation. Biol Cell 2013;105:219-233.
35. Li JZ, Huang Y, Karaman R, Ivanova PT, Brown HA, Roddy T, et al. Chronic overexpression of PNPLA3I148M in mouse liver causes hepatic steatosis. J Clin Invest 2012;122:4130-4144.
36. Smagris E, BasuRay S, Li J, Huang Y, Lai KM, Gromada J, et al. Pnpla3I148M knockin mice accumulate PNPLA3 on lipid droplets and develop hepatic steatosis. Hepatology 2015;61:108-118.
37. Johnson AM, Olefsky JM. The origins and drivers of insulin resistance. Cell 2013;152:673-684.
38. Saponaro C, Gaggini M, Gastaldelli A. Nonalcoholic fatty liver disease and type 2 diabetes: common pathophysiologic mechanisms. Curr Diab Rep 2015;15:607.
39. Dubuquoy C, Robichon C, Lasnier F, Langlois C, Dugail I, Foufelle F, et al. Distinct regulation of adiponutrin/PNPLA3 gene expression by the transcription factors ChREBP and SREBP1c in mouse and human hepatocytes. J Hepatol 2011;55:145-153.

40. Pillai S, Duvvuru S, Bhatnagar P, Foster W, Farmen M, Shankar S, Harris C, et al. The PNPLA3 I148M variant is associated with transaminase elevations in type 2 diabetes patients treated with basal insulin peglispro. Pharmacogenomics J 2018;18:487-493.

41. Maglio C. The PNPLA3 I148M variant and chronic liver disease: When a genetic mutation meets nutrients. Food Research International 2014;63:293-243.

42. Davis JN, Le KA, Walker RW, Vikman S, Spruijt-Metz D, Weigensberg MJ, et al. Increased hepatic fat in overweight Hispanic youth influenced by interaction between genetic variation in PNPLA3 and high dietary carbohydrate and sugar consumption. Am J Clin Nutr 2010;92:1522-1527.

43. Stojkovic IA, Ericson U, Rukh G, Riddestrale M, Romeo S, Orho-Melander M. The PNPLA3 Ile148Met interacts with overweight and dietary intakes on fasting triglyceride levels. Genes Nutr 2014;9:388.

44. Sanyal AJ, Chalasani N, Kowdley KV, McCullough A, Diehl AM, Bass NM, Neuschwander-Tetri BA, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. N Engl J Med 2010;362:1675-1685.

45. Stender S, Kozlitina J, Nordestgaard BG, Tybjaerg-Hansen A, Hobbs HH, Cohen JC. Adiposity amplifies the genetic risk of fatty liver disease conferred by multiple loci. Nat Genet 2017.

46. Graff M, North KE, Franceschini N, Reiner AP, Feitosa M, Carr JJ, et al. PNPLA3 gene-by-visceral adipose tissue volume interaction and the pathogenesis of fatty liver disease: the NHLBI family heart study. Int J Obes (Lond) 2013;37:432-438.

47. **Sun DQ, Liu WY, Wu SJ**, Zhu GQ, Braddock M, Zhang DC, et al. Increased levels of low-density lipoprotein cholesterol within the normal range as a risk factor for nonalcoholic fatty liver disease. Oncotarget 2016;7:5728-5737.

48. Papandreou D, Karabouta Z, Rousso I. Are dietary cholesterol intake and serum cholesterol levels related to nonalcoholic Fatty liver disease in obese children? Cholesterol 2012;2012:572820.

49. Raimondo A, Rees MG, Gloyn AL. Glucokinase regulatory protein: complexity at the crossroads of triglyceride and glucose metabolism. Curr Opin Lipidol 2015;26:88-95.

50. **Mancina RM**, **Dongiovanni P**, Petta S, Pingitore P, Meroni M, Rametta R, Boren J, et al. The MBOAT7-TMC4 Variant rs641738 Increases Risk of Nonalcoholic Fatty Liver Disease in Individuals of European Descent. Gastroenterology 2016;150:1219-1230 e1216.

765

766 Author names in bold designate shared co-first authorship.

767

768 **ACKNOWLEDGMENT**

769 We thank the reviewers for their helpful suggestions, which have improved this paper.

770 **Conflict of Interest**

771 Drr Laura Yerges-Armstrong is a current employee stockholder for GlaxoSmithKline, however,
772 the current work was conducted while at University of Maryland School of Medicine. Dr. Ingrid
773 Borecki owns stock in Regeneron Pharmaceuticals. Dr. Jeffrey R. O'Connell was a consultant
774 for Regeneron Pharmaceuticals for a period of time during this study.

775

776

777

778

779

780

781 **Figure 1.** Shown is the percentage of non-diabetic individuals in FamHS with ≥ 30% fat in the
782 liver (moderate to severe hepatic steatosis) per *PNPLA3*-rs738409 genotype in the lowest and
783 highest quartile of insulin levels. As the level of insulin increases, the percentage of individuals
784 with ≥ 30% fat in the liver increases more markedly with increasing copies of the G risk allele
785 (non-parallel lines show interaction). Among those with the GG genotype, the difference (Δ) in
786 the percentage of individuals with moderate to severe liver fat increases by 55 percentage
787 points between the lowest and highest insulin quartiles. In contrast, this difference is lower
788 among those with the CG genotype (41%), and CC genotype (31%).

Table 1. Demographic and Characteristics of Study Participants in each Cohort by Ancestry

	AGES	Amish	CARDIA	FamHS	FHS	MESA	CARDIA	FamHS	GENOA	MESA
Demographic	**European Ancestry (11,870)**						**African Ancestry (2,881)**			
N=14,751	2,865	541	1,282	2,684	2,966	1,532	642	620	560	1,059
Age	76.38 ± 5.46	56.84 ± 12.81	50.74 ± 3.33	57.14 ± 13.28	50.54 ± 10.14	63.05 ± 10.49	49.47 ± 3.86	53.35 ± 10.82	68.86 ± 8.01	63.17 ± 10.00
Men (6,444)	1,139 (40%)	252 (47%)	595 (46%)	1,207 (45%)	1,454 (49%)	746 (49%)	233 (36%)	212 (34%)	141 (25%)	465 (44%)
Women (8,307)	1,726 (60%)	289 (53%)	687 (54%)	1,477 (55%)	1,512 (51%)	786 (51%)	409 (64%)	408 (66%)	419 (75%)	594 (56%)
Characteristics										
Liver Attenuation (HU)[¥]	59.22 ± 8.64	63.05 ± 7.76	55.05 ± 12.28	59.14 ± 11.19	65.40 ± 9.83	59.33 ± 12.43	56.38 ± 10.86	59.52 ± 9.23	60.10 ± 9.39	61.18 ± 9.06
Insulin (mU/L)	9.22 ± 6.39	11.75 ± 6.22	9.26 ± 6.77	9.88 ± 7.16	9.05 ± 7.38	8.90 ± 4.94	11.60 ± 8.29	13.02 ± 10.22	8.30 ± 5.73	9.61 ± 5.53
HOMA-IR[‡]	2.31 ± 1.76	2.69 ± 1.63	2.20 ± 1.77	2.37 ± 1.87	2.32 ± 2.31	1.99 ± 1.27	2.75 ± 2.14	3.14 ± 2.69	2.02 ± 1.46	2.19 ± 1.41
Glucose (mmol/L)	5.49 ± 0.50	4.94 ± 0.52	5.18 ± 0.50	5.25 ± 0.53	5.47 ± 1.12	4.90 ± 0.58	5.17 ± 0.54	5.25 ± 0.57	5.37 ± 0.50	5.03 ± 0.59
BMI (kg/m^2)	27.00 ± 4.49	27.72 ± 4.85	28.50 ± 6.18	28.86 ± 5.69	27.51 ± 5.22	28.06 ± 5.05	31.94 ±7.48	32.71 ± 7.37	32.71 ± 7.27	29.95 ± 5.77
Obese[†]	618 (22%)	155 (29%)	425 (33%)	972 (36%)	769 (26%)	449 (29%)	352 (55%)	377 (61%)	332 (59%)	465 (44%)
WHR (cm)[§]	nval	0.87 ± 0.07	0.85 ± 0.10	0.91 ± 0.10	0.94 ± 0.08	0.93 ± 0.09	0.85 ± 0.08	0.92 ± 0.07	0.89 ± 0.08	0.92 ± 0.08
TG (mg/dL)	106.48 ± 59.06	90.42 ± 57.45	121.64 ± 85.07	144.03± 94.05	126.11 ± 88.07	136.55 ± 99.31	101.55 ± 73.24	111.82 ± 80.09	100.28 ± 62.67	103.82 ± 60.61
LDL (mg/dL)	146.84 ± 35.73	141.31± 38.66	116.27 ± 30.15	112.9 ± 34.22	117.70 ± 31.71	120.24 ± 30.42	112.59 ± 33.83	115.39 ± 36.05	123.85 ± 33.59	118.39 ± 32.87
HDL (mg/dL)	61.75 ± 17.31	57.05 ± 15.37	58.43 ± 18.42	48.82 ± 14.37	54.16 ± 16.77	51.68 ± 15.59	57.59 ± 16.70	53.55 ± 15.41	57.31 ± 16.52	52.39 ± 15.14
Alcohol (drinks/week)	1.09 ± 2.37	nval	5.73 ± 10.07	2.98 ± 7.10	5.39 ± 7.88	5.06 ± 8.40	3.86 ± 10.60	3.24 ± 9.45	0.28 ± 1.18	3.86 ± 8.89
Heavy drinkers*	17 (0.59%)	nval	470 (37%)	152 (6%)	424 (14.3%)	335 (22%)	144 (22%)	69 (11%)	0	139 (13%)

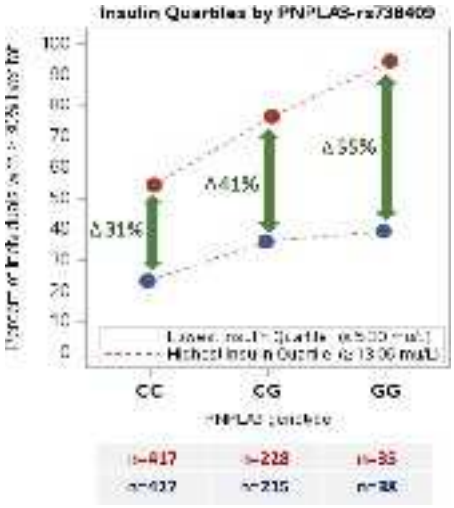
Statistics are presented as mean ± standard deviation (SD), or as n (%). The table includes individuals with liver attenuation and genetic information from each cohort that were included in analyses. LA and metabolic traits were not adjusted for covariates. The sample size for each trait varied from N depending on the data available. Summary statistics for fasting insulin, HOMA-IR and fasting glucose excludes diabetics; fasting LDL excludes statin users. ¥ Raw liver attenuation measured in Hounsfield units. ‡ Calculated as [fasting insulin (mU/L) x fasting glucose (mmol/L)/22.5]; † Defined as BMI ≥ 30 kg/m^2; § not adjusted for BMI; nval= not available in

cohort. *Defined as ≥ 8 drinks per week for women and ≥ 15 drinks per week for men. The Amish do not consume alcohol. Units in the table are HU=Hounsfield units; mU/L=milliunits per liter; mmol/L=millimoles per liter; kg/m^2= kilograms divided by height in meters squared; cm=centimeters; mg/dL=milligram per deciliter.

Table 2. Meta-analyses results for interactions between four SNPs and inverse normal-transformed residuals of metabolic traits on LA_{ivn} in combined ancestries.

	rs738409				rs780094*				rs2228603*				rs12137855			
	Gene	Chr	Alleles (Ref/O)	Ref AF	Gene	Chr	Alleles (Ref/O)	Ref AF	Gene	Chr	Alleles (Ref/O)	Ref AF	Gene	Chr	Alleles (Ref/O)	Ref AF
	PNPLA3	22	G/C	0.24	*GCKR*	2	T/C	0.39	*NCAN/ TM6SF2*	19	T/C	0.13	*LYPLAL1*	8	C/T	0.79
	(SNP x Metabolic Traits)															
Metabolic Traits	**β_{int}**	**SE**	**P-value**	**N**	**β_{int}**	**SE**	**P-value**	**N**	**β_{int}**	**SE**	**P-value**	**N**	**β_{int}**	**SE**	**P-value**	**N**
Insulin	**-0.11**	**0.02**	**4.79E-14**	**12,651**	**-0.04**	**0.01**	**4.57E-04**	**12,651**	-0.06	0.03	4.37E-02	12,651	-0.02	0.02	1.55E-01	12,651
HOMA-IR	**-0.12**	**0.02**	**4.68E-15**	**12,554**	**-0.04**	**0.01**	**1.32E-03**	**12,554**	-0.06	0.03	3.63E-02	12,554	-0.02	0.02	1.38E-01	12,554
Glucose	**-0.05**	**0.02**	**1.26E-03**	**12,742**	-0.01	0.01	4.37E-01	12,742	-0.06	0.03	7.41E-02	12,742	-0.02	0.02	1.91E-01	12,742
BMI	**-0.08**	**0.01**	**8.13E-08**	**14,693**	-0.03	0.01	6.31E-03	14,693	-0.05	0.03	5.89E-02	14,693	-0.02	0.01	8.18E-01	14,693
$WHR_{adj}BMI$	-0.05	0.02	7.59E-03	10,051	-0.04	0.02	1.32E-02	10,051	-0.08	0.03	1.26E-02	10,051	0.01	0.02	7.76E-01	10,051
TG	**-0.05**	**0.02**	**2.95E-03**	**14,551**	**-0.04**	**0.01**	**4.17E-03**	**14,551**	0.00	0.03	9.77E-01	14,551	-0.03	0.02	5.75E-02	14,551
LDL	0.00	0.02	7.94E-01	12,123	0.00	0.01	9.33E-01	12,123	-0.06	0.03	5.50E-02	12,123	0.02	0.02	2.29E-01	12,123
HDL	-0.04	0.02	1.41E-02	14,543	0.03	0.01	3.72E-02	14,543	0.00	0.03	9.55E-01	14,543	-0.01	0.02	3.72E-01	14,543

Chr, Chromosome; Ref/O, Reference/Other allele (reference allele is the effect allele of each SNP); Ref AF, Reference allele frequency; β_{int}, interaction effect size; SE, standard error. P-values that reached significance threshold (P≤ 4.17E-03) are in bold; N is the highest sample size in meta-analyses. *rs780094 is in LD (r^2=0.93) with rs1260326, a functional missense variant in GCKR; *rs2228603 is in LD (r^2=0.79) with rs58542926, a functional missense variant in TM6SF2.



hep4_1353_f1.tif