



Published in final edited form as:

*Gastroenterology*. 2016 May ; 150(5): 1219–1230.e6. doi:10.1053/j.gastro.2016.01.032.

## The *MBOAT7-TMC4* Variant rs641738 Increases Risk of Nonalcoholic Fatty Liver Disease in Individuals of European Descent

Rosellina Margherita Mancina<sup>1,\*</sup>, Paola Dongiovanni<sup>2,\*</sup>, Salvatore Petta<sup>3</sup>, Piero Pingitore<sup>1</sup>, Marica Meroni<sup>4</sup>, Raffaella Rametta<sup>4</sup>, Jan Borén<sup>1</sup>, Tiziana Montalcini<sup>5</sup>, Arturo Pujja<sup>5</sup>, Olov Wiklund<sup>1,6</sup>, George Hindy<sup>7</sup>, Rocco Spagnuolo<sup>8</sup>, Benedetta Maria Motta<sup>1</sup>, Rosaria Maria Pipitone<sup>4</sup>, Antonio Craxi<sup>3</sup>, Silvia Fargion<sup>2,4</sup>, Valerio Nobili<sup>9</sup>, Pirjo Käkälä<sup>10</sup>, Vesa Kärjä<sup>11</sup>, Ville Männistö<sup>12</sup>, Jussi Pihlajamäki<sup>13,14</sup>, Dermot F. Reilly<sup>15</sup>, Jose Castro-Perez<sup>16,17</sup>, Julia Kozlitina<sup>18,#</sup>, Luca Valenti<sup>2,4,#</sup>, and Stefano Romeo<sup>1,5,6,#</sup>

<sup>1</sup>Department of Molecular and Clinical Medicine, University of Gothenburg, Sweden <sup>2</sup>Internal Medicine, Fondazione IRCCS Ca' Granda Ospedale Policlinico Milano, Milan, Italy <sup>3</sup>Department of Gastroenterology, Università di Palermo, Palermo, Italy <sup>4</sup>Department of Pathophysiology and Transplantation Università degli Studi di Milano, Milan, Italy <sup>5</sup>Clinical Nutrition Unit, Department of Medical and Surgical Sciences, University Magna Graecia, Catanzaro, Italy <sup>6</sup>Cardiology Department, Sahlgrenska University Hospital, Gothenburg, Sweden <sup>7</sup>Diabetes and Cardiovascular Disease-Genetic Epidemiology, Lund, Sweden <sup>8</sup>Division of Gastroenterology, Fondazione Tommaso Campanella, University Magna Graecia of Catanzaro, Italy <sup>9</sup>Hepatometabolic Unit, Bambin Gesù Hospital, Rome, Italy <sup>10</sup>Department of Surgery, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland <sup>11</sup>Department of Pathology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland <sup>12</sup> Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland <sup>13</sup>Clinical Nutrition and Obesity Center, Kuopio University Hospital, Kuopio, Finland <sup>14</sup>Department of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland <sup>15</sup>Merck Research Laboratories, Genetics and Pharmacogenomics, Boston, Massachusetts, USA <sup>16</sup>Merck Research

# Corresponding authors **Stefano Romeo, MD, PhD**, Department of Molecular and Clinical Medicine, The Sahlgrenska Academy, University of Gothenburg, Wallenberg Laboratory, Bruna Stråket 16, SE-413 45 Göteborg, Sweden, Tel: +46 (0)313426735, stefano.romeo@wlab.gu.se, **Luca Valenti, MD**, Department of Pathophysiology and Transplantation, University of Milan, Fondazione IRCCS Ca' Granda Ospedale Policlinico, Pad. Granelli, Via Sforza 35, 20122, Milan, Italy, Tel: +39-2-50320278, luca.valenti@unimi.it, **Julia Kozlitina, PhD**, McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, Texas 75390, USA, Tel: +1 214 648 1446, Julia.Kozlitina@UTSouthwestern.edu.

\* Authors share co-first authorship

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Author contributions:** The authors' responsibilities were as follows: SR, LV and JK, designed research, wrote, reviewed and edited the manuscript; RMM, PD and JK performed statistical analysis and wrote, reviewed and edited the manuscript; PP and BMM performed and interpreted the in vitro experiments, reviewed and edited the manuscript; MM, SP, RR, JB, TM, AP, OW, GH, RS, AC, SF and JP contributed to discussion and reviewed the manuscript, JCP and DFR acquired and interpreted the lipidomic data, and SR, JK and LV had primary responsibility for final content. All authors read and approved the final manuscript.

**Disclosures:** The authors declare no conflict of interest.

Laboratories, Diabetes Department, Kenilworth, New Jersey, USA <sup>17</sup>Waters Corporation, Milford, Massachusetts, USA <sup>18</sup>McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

## Abstract

**Background & Aims**—Nonalcoholic fatty liver disease (NAFLD) is a leading cause of liver damage and is characterized by steatosis. Genetic factors increase risk for progressive NAFLD. A genome-wide association study showed that the rs641738 C>T variant in the locus that contains the membrane bound O-acyltransferase domain-containing 7 gene (*MBOAT7*, also called *LPIAT1*) and transmembrane channel-like 4 gene (*TMC4*) increased the risk for cirrhosis in alcohol abusers. We investigated whether the *MBOAT7/TMC4* is a susceptibility locus for the development and progression of NAFLD.

**Methods**—We genotyped rs641738 in DNA collected from 3854 participants from the Dallas Heart Study (a multi-ethnic population-based probability sample of Dallas County residents) and 1149 European individuals from the Liver Biopsy Cross-sectional Cohort. Clinical and anthropometric data were collected, and biochemical and lipidomics were measured in plasma samples from participants. A total of 2736 participants from the Dallas Heart Study underwent also proton magnetic resonance spectroscopy to measure hepatic triglyceride content. In the Liver Biopsy Cross-sectional Cohort, a total of 1149 individuals underwent liver biopsy to diagnose liver disease and disease severity.

**Results**—The genotype rs641738 at the *MBOAT7/TMC4* locus associated with increased hepatic fat content in the 2 cohorts, and with more severe liver damage and increased risk of fibrosis compared to subjects without the variant. *MBOAT7*, but not *TMC4*, was found to be highly expressed in the liver. The *MBOAT7* rs641738 T allele was associated with lower protein expression in the liver and changes in plasma phosphatidylinositol species consistent with decreased *MBOAT7* function.

**Conclusions**—We provide evidence for an association between the *MBOAT7* rs641738 variant and the development and severity of NAFLD in individuals of European descent. This association seems to be mediated by changes in the hepatic phosphatidylinositol acyl-chain remodeling.

## Keywords

PNPLA3; TM6SF2; NASH; Arachidonic Acid

---

Nonalcoholic fatty liver disease (NAFLD) is a leading cause of liver damage, especially in Western countries<sup>1</sup>. Accumulation of triglycerides exceeding 5% of liver weight, namely steatosis, in the absence of excess alcohol intake is the hallmark of NAFLD<sup>2</sup>. Steatosis may progress to liver damage and inflammation, namely nonalcoholic steatohepatitis (NASH)<sup>3, 4</sup> and ultimately to cirrhosis and hepatocellular carcinoma<sup>5</sup>. Importantly, the fibrosis progression rate is higher in individuals with NASH than with simple steatosis<sup>6</sup>.

Predisposition to progressive NAFLD is strongly influenced by genetic heritability<sup>7, 8</sup>. Two common genetic variants increasing liver fat content have been identified by genome wide

association studies<sup>9-11</sup>, namely the I148M variant of the *Patatin-like phospholipase domain-containing-3* gene (*PNPLA3*) and the E167K variant of *Transmembrane 6 superfamily member 2* gene (*TM6SF2*). Both *PNPLA3* and *TM6SF2* variants were found to increase the susceptibility to NAFLD<sup>12-14</sup>. *PNPLA3* is a lipase<sup>15-17</sup> involved in hepatocellular lipid remodeling<sup>17-19</sup> and retinol metabolism<sup>20, 21</sup> while *TM6SF2* is involved in hepatic very-low density lipoproteins (VLDL) secretion<sup>9, 10, 22-23</sup>.

Recently, a genome wide association study identified the rs641738 C>T single nucleotide polymorphism (SNP) in the Membrane bound O-acyltransferase domain containing 7/ Transmembrane Channel-Like 4 (*MBOAT7/TMC4*) locus that has been shown to predispose to cirrhosis development in alcohol abusers<sup>24</sup>. However, the impact of this variant on hepatic fat accumulation and metabolism, and on liver damage related to NAFLD, is unknown.

In this work we tested the hypothesis that the *MBOAT7/TMC4* is a susceptibility locus for NAFLD. We found that the locus associates with increased hepatic fat content in two cohorts and with the entire spectrum of histological liver damage related to NAFLD. This association seems to be mediated by changes in hepatic phosphatidylinositol acyl-chain remodeling.

## PATIENTS AND METHODS

### Study cohorts

**Dallas Heart Study (DHS)**—The Dallas heart study (DHS) is a multiethnic population-based probability sample of Dallas County residents. The study design and recruitment procedures have been previously described in detail<sup>25</sup>. The original cohort was enrolled between 2000 and 2002, and all participants as well as their spouses or significant others were invited for a repeat evaluation in 2007-2009 (DHS-2). The study was approved by the Institutional Review Board of University of Texas Southwestern Medical Center and all individuals provided written informed consent. During each visit, all participants completed a detailed survey including questions about demographics, socioeconomic status, medical history, and current medication use, and underwent a health examination that involved measurement of blood pressure, anthropometry, blood and urine sample collection, and imaging studies. Ethnicity was self-reported. Hepatic triglyceride content was measured with proton magnetic resonance spectroscopy (<sup>1</sup>H-MRS) as previously described<sup>26, 27</sup>. The present investigation includes a total of 3854 individuals (51% African-Americans, 30% European-Americans, 17% Hispanics, and 2% other ethnicities), who provided blood samples for genetic and lipidomic analysis and underwent imaging studies during the initial or the follow-up visit. All analyses were based on cross-sectional data. Lipidomic analyses were restricted to 3447 DHS-1 participants. Hepatic triglyceride measurements were available for a total of 2736 participants (including 2220 of the DHS-1 and 516 of the DHS-2 participants). Presence of liver steatosis was defined as hepatic triglycerides content >5.5%. Given the low prevalence of heavy drinking (>30 g/day), we did not exclude subjects based on alcohol intake.

**Liver Biopsy cross-sectional Cohort**—The Liver Biopsy cross-sectional Cohort has been previously described<sup>12</sup>. Briefly, a total of 1201 individuals of European descent were

consecutively enrolled in four European centers: 1149 out of 1201 individuals were included in the present study. Forty-eight percent were from Milan (the Metabolic Liver Diseases outpatient service and from the Fondazione IRCCS Ca' Granda Ospedale Policlinico Milano, Milan, Italy)<sup>28</sup>, 23% were from the Gastrointestinal & Liver Unit of the Palermo University Hospital, Palermo, Italy<sup>29</sup>, 12% were children and adolescences from the Bambino Gesù Children's Hospital, Rome, Italy<sup>30</sup>, and 16% were from the Northern Savo Hospital District, Kuopio, Finland<sup>31</sup>. The inclusion criteria were liver biopsy for suspected NASH or severe obesity, availability of DNA samples and clinical data. Individuals with increased alcohol intake (men, >30 g/day; women, >20 g/day), viral and autoimmune hepatitis or other causes of liver disease were excluded. All subjects gave written informed consent. In the present analyses the cohort is composed of 1149 individuals as compared to 1201 in our previous publication<sup>12</sup> for two reasons: a) lack of DNA samples b) lack of informed consent from the patient to proceed with this specific genetic analysis. Diagnosis of NASH was based on the presence of steatosis with lobular necro-inflammation and ballooning or fibrosis. Disease activity was assessed according to the NAFLD Activity Score (NAS); fibrosis was staged according to the recommendations of the NAFLD clinical research network<sup>3</sup>. The scoring of liver biopsies was performed by independent pathologists unaware of patients' status and genotype. This information was provided in the previous publications using this cohort<sup>12, 32</sup>. The concordance between pathologists within this cohort was very good for fibrosis and good for steatosis with a coefficient inter-observer agreements for fibrosis, steatosis grade, lobular inflammation and ballooning of 0.89, 0.76, 0.60 and 0.55.

### Genotyping

All DHS participants included in the current study were previously genotyped for rs58542926 (*TM6SF2* E167K), rs738409 (*PNPLA3* I148M), and rs641738 (*MBOAT7/TMC4*) using the Illumina Infinium HumanExome BeadChip, as described<sup>9, 10</sup>. The genotype frequencies for rs641738 were in Hardy-Weinberg equilibrium in European-American and Hispanic participants of the DHS ( $p=0.95$  and  $p=0.79$ , respectively, using the exact test for HWE<sup>33</sup>). However, a small deviation from HWE was found in African-Americans included in our analytical sample ( $p=0.04$ ). Since no deviation from HWE was observed in the entire sample of African-Americans genotyped using the Exome chip ( $p=0.1$ ), this deviation is likely due to random sampling rather than genotyping error.

The Liver Biopsy cross-sectional Cohort has been genotyped for the rs58542926 (*TM6SF2* E167K) and rs738409 (*PNPLA3* I148M) as previously described<sup>12</sup>. The rs641738 *MBOAT7/TMC4* genotyping has been performed in duplicate by TaqMan 5'-nuclease assays (Life Technologies, Carlsbad, CA). In this population, partially selected due to referral for suspected steatohepatitis, there was a borderline deviation from Hardy Weinberg equilibrium. The duplicate genotype concordance rate was 100%.

### Gene Expression in human tissues and human primary hepatic cells

Real-Time quantitative PCR (qPCR) was used to assess *MBOAT7* and *TMC4* mRNA expression in a selection of 14 human tissues using TissueScan Human Normal cDNA Array (Origene, Rockville, MD; USA). The *MBOAT7* and *TMC4* mRNA were assessed in human

primary hepatocytes (ScienceCell, Carlsbad, CA, USA), human primary hepatic stellate cells and human primary hepatic sinusoidal endothelial cells (ScienceCell). The mRNA expression levels were determined using TaqMan probes and master mix (Life Technologies) according to the manufacturer's protocol.

### Generation of stable cell line overexpressing *MBOAT7*

Human wild type *MBOAT7* cDNA was synthesized and cloned into a pcDNA3.1 vector (Thermo Fisher Scientific, Rockford, IL, USA) with a V5 epitope tag (coding for a 14-amino-acid sequence: GKPIPPLLGLDST) at the C-terminus and before the stop codon by GeneArt Gene Synthesis (Thermo Fisher Scientific). Human liver carcinoma (HepG2) cells were transfected with *MBOAT7* pcDNA 3.1 using TurboFect (Thermo Scientific, Rockford, IL, USA) reagent according to the manufacturer's protocol. Forty-eight hours after transient transfection, cells were treated with 500 µg/ml of geneticin-G418 (Life Technologies), to select resistant cells. After 15 days, three single colonies per each genotype were selected and expanded.

### Membrane fractionation

Endoplasmic reticulum, crude mitochondrial, purified mitochondrial, mitochondria-associated membrane fractions were separated as previously described<sup>34</sup>.

Lipid droplets were separated as previously described<sup>35</sup>. One ml from the bottom of the tube was collected as a negative fraction control and indicated as infranatant.

Next, post-nuclear supernatant, cytosol, endoplasmic reticulum, crude mitochondria, pure mitochondria, mitochondria-associated membrane, lipid droplets and infranatant were analyzed by immunoblotting.

### Gene expression analysis in liver biopsies

**Patients**—*MBOAT7* and *TMC4* hepatic gene expression was measured in a subset of 98 obese patients belonging to the Milan cohort, after percutaneous liver biopsy performed during bariatric surgery for staging of liver damage severity, and for whom liver samples were available. These included 24 patients with normal liver histology, 28 with simple steatosis, and 46 with NASH. Informed written consent was obtained from each subject, and the study protocol was approved by the Ethical Committees of the Fondazione IRCCS Ca' Granda of Milano.

**RNA isolation**—RNA was extracted from liver biopsies using Trizol reagent (Life Technologies), according to the manufacturer's instruction. One µg of total RNA was retro-transcribed by using the VILO random hexamers synthesis system (Life Technologies). Gene expression was evaluated by Real-Time qPCR, and performed by the 7500 Fast thermocycler (Life Technologies), using the Taqman Universal PCR Master Mix (Life Technologies) and Taqman probes for human *MBOAT7* and *TMC4*. All reactions were delivered in triplicate, and data were normalized to beta-actin gene expression. Association analysis between rs641738 variant (additive model) and gene expression was conducted by the PLINK v1.07 genetic analysis software.

**Immunoblotting**—Proteins were separated by electrophoresis and transferred to nitrocellulose membranes. Immunoblot analysis was performed according to standard procedures. Protein lysates were extracted from liver biopsies of 21 obese patients, using RIPA buffer. Seven samples for each rs641738 genotype (CC/CT/TT) were collected and pooled according to genotype. All reactions were performed in duplicate.

The following antibodies were used: mouse anti-V5 (Invitrogen-Life Technologies), rabbit anti-calnexin (Sigma-Aldrich, Saint Louis, Missouri, USA), mouse anti-COXIV (Cell Signalling, Beverly, MA, USA), mouse anti-FACL4 (Abcam, Cambridge, UK), rabbit anti-MBOAT7 (LENG4, lot number: B1315, Santa Cruz Biotechnology, Santa Cruz, CA), goat polyclonal anti- $\beta$ -actin (Santa Cruz Biotechnology).

**Lipidomic analysis**—Lipids were measured in fasting plasma samples from 3447 DHS-1 participants using high resolution mass spectrometry which consisted of Ultra Performance Liquid Chromatography coupled with a hybrid quadrupole orthogonal time of flight mass spectrometer (UPLC-QTOF, Waters Corp., Milford, MA) The data was collected in positive and negative ion electrospray using accurate mass mode so that elemental compositions were obtained <5ppm. Quantification of lipids was achieved by reference to appropriate heavy labeled deuterated internal standards for each lipid class detected. In total, 14 distinct species of phosphatidylinositol (PI) and 9 free fatty acids (FFA) were measured and analyzed in the present study. Lipid side chain composition is denoted by  $a.b$ , where  $a$  is the number of carbons in the side chain and  $b$  the number of double bonds. The precise composition of PI species (position of the double bonds and the distribution of carbon atoms in the two side-chains) was not determined with the lipidomics analysis described above.

**Statistical analysis**—For descriptive statistics, continuous variables are shown as mean and standard deviation or median and interquartile range. Categorical variables are presented as number and proportion. For the DHS, demographic, anthropometric, and clinical characteristics of participants were compared using linear regression (for continuous) and logistic regression (for categorical characteristics). All models were adjusted for age, gender, ethnicity, statin use (triglyceride and LDL cholesterol), BMI and IFG/T2DM (triglyceride and hepatic triglycerides content), *PNPLA3* I148M and *TM6SF2* E167K genotype (hepatic triglycerides content). The odds of hepatic steatosis per *MBOAT7/TMC4* rs641738 T alleles were estimated by logistic regression models and adjusted for ethnicity, age, gender, BMI, IFG/T2DM, *PNPLA3* I148M and *TM6SF2* E167K genotypes. All genetic analyses were performed assuming an additive model (by coding the genotypes as 0, 1, and 2 for wild-type homozygotes, heterozygotes, and alternate allele homozygotes, respectively). Interactions among genes and between genes and environmental factors were tested by adding the corresponding gene  $\times$  gene and gene  $\times$  environment interaction terms to the multivariable-adjusted regression models.

For the Liver Biopsy cross-sectional Cohort, characteristics of participants were compared across *MBOAT7/TMC4* genotypes using linear regression (for continuous characteristics) and logistic regression (for categorical characteristics). All models were adjusted for age, gender, body mass index (BMI), impaired fasting glucose (IFG)/ type 2 diabetes mellitus (T2DM), the number of *PNPLA3* 148M alleles, for the presence of the *TM6SF2* E167K

variant, and study cohort when appropriate. The association between the *MBOAT7/TMC4* rs641738 variant and the components of the NAFLD activity score (severity of steatosis, necro-inflammation, and ballooning) have been tested by multivariate ordinal regression analysis adjusted for age, gender, BMI, presence of IFG/T2DM, the number of *PNPLA3* 148M alleles, presence of the *TM6SF2* E167K variant, and indication of liver biopsy (severe obesity vs. nonalcoholic fatty liver with increased liver enzymes). Logistic regression models adjusted for age, gender, BMI, IFG/T2DM, the number of *PNPLA3* 148M alleles, presence of the *TM6SF2* E167K variant (wild type vs. heterozygotes+homozygotes, due to the presence of few homozygotes) and the number of *MBOAT7/TMC4* rs641738 T alleles, and indication for liver biopsy (severe obesity vs. increased liver enzymes in NAFLD) were fit to test the association with binary traits (steatosis, NASH, and fibrosis).

Population attributable risk of genetic variants on histological traits has been estimated as previously described<sup>36</sup>.

Statistical analyses were carried out using the IBM Statistical Package for Social Sciences (IBM SPSS, version 19.0, Inc. Chicago, IL, USA) or R statistical analysis software version 3.2.0 (<http://www.R-project.org/>). P-values <0.05 were considered statistically significant.

## RESULTS

### The *MBOAT7/TMC4* rs641738 variant increases hepatic triglyceride content

To test whether the minor (T) allele of rs641738 was associated with hepatic steatosis we first examined the differences in hepatic triglyceride content by rs641738 genotype in participants of the Dallas Heart Study with available liver spectroscopy (n=2736). Carriers of the minor T allele had higher hepatic triglyceride content than homozygotes for the common allele [median (interquartile range) =3.5% (2-6) in CC, 3.6% (2-7) in CT, and 4.1% (2-8) in TT; p=0.005, Figure 1]. Each copy of the T allele increased the odds of hepatic steatosis by 20% (p=0.006) after correction for demographic and anthropometric (age, gender, BMI), clinical (IFG or T2DM and statin use) and genetic (*PNPLA3* 148M and *TM6SF2* E167K) factors (Table 1). No differences in clinical, anthropometrical and biochemical indices, including aminotransferases, were found across the *MBOAT7* rs641738 genotypes in the Dallas Heart Study (Supplementary table 1). Importantly, rs641738 was not associated with average alcohol intake or the prevalence of heavy drinking (>30 g/day) (Supplementary table 1), and the association of rs641738 with hepatic steatosis was unchanged after an additional adjustment for alcohol intake (p=0.005) or after removing heavy drinkers from analysis (p=0.004). When we stratified the analysis by ethnic groups, the minor T allele was associated with higher hepatic triglyceride content only in African Americans (p=0.019, Supplementary table 2) although a comparable (but not statistically significant) increase in hepatic triglyceride content was observed in European-Americans. Conversely, the association with hepatic steatosis remained significant only in European Americans (Table 1: OR: 1.37; 95% CI: 1.09-1.72; p=0.007). The estimated odds ratio in African Americans was directionally consistent (OR = 1.17), but not statistically significant.

We did not see an interaction effect among *PNPLA3*, *TM6SF2* and *MBOAT7* variants on mean hepatic triglyceride content or the risk of hepatic steatosis (p>0.05). Instead the three

genetic variants appeared to act in an additive fashion, with a step-wise increase in mean hepatic fat content per each additional risk allele (Supplementary figure 1).

### **The MBOAT7/TMC4 rs641738 variant increases the risk of the entire NAFLD spectrum**

To confirm the finding and test whether the *MBOAT7/TMC4* locus was also associated with disease severity, we genotyped the rs641738 variant in a large cohort of individuals (n=1149) with suspected nonalcoholic steatohepatitis, namely the Liver Biopsy Cohort (Supplementary table 3). Success rate of genotyping was >99%. In this population, partially selected due to referral for suspected steatohepatitis, there was a borderline deviation from Hardy-Weinberg equilibrium (p=0.017, Supplementary table 4).

The association of *MBOAT7/TMC4* genotype with the spectrum of liver damage related to NAFLD is shown in table 2. In multivariate logistic regression analysis each T allele conferred an increased risk of steatosis (OR: 1.42; 95% CI: 1.07-1.91; p=0.015), NASH (OR: 1.18; 95% CI: 1.00-1.40; p=0.050) and clinically significant fibrosis stage F2-4 (OR: 1.30; 95% CI: 1.06-1.70; p=0.012) after adjustment for demographic and anthropometric (age, gender, BMI), clinical (indication for liver biopsy, IFG/T2DM) and genetic (*PNPLA3* I148M and *TM6SF2* E167K) factors.

The impact of *MBOAT7/TMC4* genotype on the single components of liver damage associated with NAFLD in multivariate ordinal regression analysis, adjusted for covariates as specified above, is shown in figure 2. The minor (T) allele was associated with a higher degree of liver steatosis (p=0.045), more severe necro-inflammation (p=0.008) and more advanced fibrosis (p=0.027). No differences in clinical, anthropometric or biochemical parameters including serum aminotransferases and circulating lipids were detected across *MBOAT7/TMC4* genotypes (Supplementary table 3).

We next examined the proportion of attributable risk conferred by the *MBOAT7*, *PNPLA3* and *TM6SF2* genetic variants. As expected, the *PNPLA3* I148M variant had the largest impact on the whole spectrum of liver damage, while the *MBOAT7/TMC4* rs641738 variant had generally a larger impact than the E167K *TM6SF2* variant (Supplementary table 5). Finally, there was no interaction effects among *PNPLA3*, *TM6SF2* and *MBOAT7* variants on histological traits in the Liver Biopsy Cohort (all p-values >0.5).

### **MBOAT7 is highly expressed in the liver and is attached to endoplasmic reticulum, mitochondria-associated membranes and lipid droplets**

As a first step to elucidate the role of *MBOAT7* and *TMC4* in liver disease we examined the mRNA expression of these two genes in several human tissues (Figure 3 A). *MBOAT7* is expressed in many human tissues including the liver where it is expressed with a >10 fold higher expression than *TMC4*. We next examined the *MBOAT7* and *TMC4* mRNA expression in three different human hepatic cell types, namely hepatocyte, hepatic stellate cells and hepatic sinusoidal endothelial cells. Overall, the mRNA expression levels were high in all three human cell types (ct in hepatocytes = 25; stellate cells = 26; sinusoidal cells = 23) with the highest expression in sinusoidal endothelial cells. *TMC4* was expressed at a very low level in all three hepatic cell types (ct > 32 in all three cell types see Figure 3 B). The intracellular localization of the endogenous MBOAT7 was therefore evaluated in human



hepatoma cell line (HepG2). We observed MBOAT7 co-localizing with purified endoplasmic reticulum (ER) membranes, lipid droplets (LD) and with mitochondria-associated membranes (MAM), but not with pure mitochondria membranes (Figure 4 A, B, C). To confirm our data we examined the localization of a recombinant V5 tagged MBOAT7 in membrane fractions of HepG2 stably overexpressing this V5 recombinant protein (see supplementary figure 2) and we obtained identical results. We next examined MBOAT7 protein levels in human liver samples, in HepG2 and in a NAFLD-hepatocellular carcinoma cell line (HCC1<sup>37</sup>) as shown in supplementary figure 3. We observed a similar synthesis of MBOAT7 in human liver and in tumor cell lines.

### The rs641738 associates with lower MBOAT7 protein levels in the liver

To investigate whether the mechanism whereby rs641738 variant predisposes to NAFLD is by changing the gene expression, we examined hepatic mRNA levels of *MBOAT7* and *TMC4* in a subset of patients for whom liver samples were available (n=98). This subgroup is representative of the whole cohort, as the rs641738 T allele was associated with the presence and the severity of steatosis (p=0.01 and p=0.04, respectively, Figure 4 D). We found that *MBOAT7* mRNA levels were about 20-fold higher than those of *TMC4* (p<0.001). Importantly, the rs641738 T allele was associated with lower *MBOAT7* mRNA levels (estimate  $-0.08 \pm 0.03$ ; p=0.03; p=0.049 after adjustment for age, sex, presence of T2DM and the severity of liver disease). Conversely, there was no significant association between rs641738 and *TMC4* mRNA levels. Finally, we examined MBOAT7 protein levels in liver biopsies of 21 obese individuals (7 for each genotype) and found that the number of T alleles was associated with lower protein expression (p<0.005, see Figure 4 E). These data suggest that the rs641738 T allele may predispose to NAFLD by down-regulating the *MBOAT7* expression and synthesis.

### The rs641738 associates with differences in phosphatidyl-inositol species in plasma

MBOAT7, also known as lysophosphatidylinositol acyltransferase 1 (LPIAT1), is an enzyme involved in the phospholipid acyl-chain remodeling, that catalyzes a desaturation of the second acyl-chain of phospholipids and specifically transfers polyunsaturated fatty acids such as arachidonoyl-CoA to lyso-phosphatidylinositol (LPI) and other lyso-phospholipids<sup>38</sup>. To test the hypothesis that the rs641738 association is due to changes in MBOAT7 activity we examined the differences in the concentration of phosphatidylinositol species and free fatty acids in plasma samples of the Dallas Heart Study participants (Table 3). We found significant differences in absolute levels of most of the 14 species of PI examined. In line with the reduction of enzyme synthesis associated with the risk variant, the T allele was associated with a PI composition pattern suggestive of lower levels of arachidonoyl-PI/total PI (e.g., PI 36:4) and with higher levels of oleyl-PI/total PI linoleoyl-PI/total PI ratios (e.g., PI 36:2, PI 34:1, PI 34:2) (Table 3). We observed even stronger signal when we looked at the pairwise ratios between the concentrations of individual PI species, particularly ratios of PI containing polyunsaturated fatty acid chains to those most likely composed of saturated and monounsaturated fatty acid chains (e.g., PI 36:4/PI 34:2). These differences were highly significant across all three major ethnic groups in the DHS (Figure 3 C). We found no differences across genotypes in circulating levels of free fatty acids

(Supplementary table 6) or any other class of lipids measured, including triglycerides, cholesteryl esters, phospholipids, ceramides, and sphingomyelins (data not shown).

## DISCUSSION

In this work we show for the first time an association of rs641738 C>T genetic variant in the *MBOAT7/TMC4* locus with increased risk of the entire spectrum of NAFLD.

The *MBOAT7/TMC4* was recently identified as a locus involved in alcoholic cirrhosis<sup>33</sup>. Since alcoholic and nonalcoholic cirrhosis share the major common genetic risk determinants, namely the *PNPLA3* I148M and *TM6SF2* E167K variants<sup>9, 10</sup>, we tested the hypothesis that *MBOAT7/TMC4* was associated with an increased risk of NAFLD and liver damage development. We observed a higher hepatic fat content in two cohorts: measured by MRS in a multiethnic general population from the US, the Dallas Heart Study; and by histological evaluation in at-risk individuals of European descent, the Liver Biopsy Cohort. In the Liver Biopsy Cohort, the rs641738 T risk allele was also associated with higher severity of necro-inflammation and fibrosis. The observed association of rs641738 with NASH was borderline significant. This may be due to the fact that the variant was associated with necro-inflammation, but not with hepatocellular ballooning (Figure 2). Most importantly the *MBOAT7* variant was independently associated with development of fibrosis (Table 2 and Figure 2), which represents the major determinant of the prognosis of patients with NAFLD<sup>39, 40</sup>. These results show a substantial overlap between genes involved in the pathogenesis of alcoholic and nonalcoholic liver disease, indicating that these conditions have a common etiology involving altered regulation of hepatic lipid metabolism.

Interestingly, as previously reported for *PNPLA3* and *TM6SF2* variants, the rs641738 polymorphism was not associated with indices of insulin resistance suggesting that moderate differences in hepatic fat content or inflammation do not cause insulin resistance, and that NASH is rather a consequence than a cause of insulin resistance.

Notably, our previous genome-wide association studies that identified *PNPLA3* and *TM6SF2* variants failed to detect the association of *MBOAT7/TMC4* with liver fat content. This is due to the modest effect size of rs641738 on hepatic TG content and the conservative significance thresholds used in genome-wide studies in order to guard against false positive signals. Our results indicate that there are additional loci affecting liver fat that remain to be discovered. Concerning the mechanism underpinning the association, as a first step to elucidate the role of *MBOAT7/TMC4* locus in liver disease we examined the expression of the two genes in a panel of human tissues. While *MBOAT7* was expressed in a wide range of human tissues, it had a ten-fold higher expression in the liver as compared to *TMC4*. At a cellular level *MBOAT7* but not *TMC4* was highly expressed in human hepatocytes, sinusoidal endothelial and stellate cells. In hepatoblastoma cells stably overexpressing *MBOAT7* the recombinant protein was localized specifically in membrane fractions rich in phospholipids<sup>34</sup> i.e.: the endoplasmic reticulum and the mitochondria-associated membranes. These data are consistent with previous published data<sup>41</sup>.

As rs641738, to our knowledge, is not in linkage with non-synonymous functional variants of *MBOAT7*, we next examined the impact of this variant on mRNA and protein levels of *MBOAT7* stratified by rs641738 genotype. We observed a reduction in both mRNA expression and protein synthesis associated with the risk T allele, suggesting that the mechanism underlying the association is related to lower enzymatic activity due to lower protein levels. We do not know whether rs641738 is causing the decreased expression of *MBOAT7* or whether the observed association is due to linkage disequilibrium between rs641738 and variants in the 3'UTR of *MBOAT7* (e.g., rs8736,  $r^2 > 0.95$  in the CEU population from the 1000 genomes project<sup>42</sup>), which may influence mRNA stability/translation. *MBOAT7*, also known as lysophosphatidylinositol acyltransferase 1 (LPIAT1), is an enzyme involved in the phospholipid acyl-chain remodeling, known as the Lands' cycle<sup>43</sup>. *MBOAT7* specifically transfers polyunsaturated fatty acids such as arachidonoyl-CoA to lysophosphatidylinositol (LPI) and other lyso-phospholipids<sup>38</sup>. Furthermore, a recent genome-wide association study showed that variants in the *MBOAT7* locus were associated with arachidonic acid-containing metabolites<sup>44</sup>. As a further step to demonstrate that the rs641738 association is due to changes in the *MBOAT7* activity, we examined the differences in the concentration of phosphatidylinositol in plasma samples of Dallas Heart Study participants. In line with the predicted protein function, we observed significant differences in the levels of multiple PI species, but not the other lipid classes. Although we could not establish the precise fatty acid chain composition of the individual PI species, the observed association pattern of the T allele was broadly suggestive of lower levels of arachidonoyl-PI/total PI, docosapentaenoyl-PI/total PI and with higher levels oleyl-PI/total PI, linoleoyl-PI/total PI and docosahexanoyl-PI/total PI ratios. Even bigger differences were observed for the ratios between the concentrations of pairs of PI species. In cases where a pair of PI species is likely related to the substrate and product of a given enzymatic reaction, such ratios can serve as a surrogate for enzymatic activity<sup>45</sup>. The observed associations are therefore strongly indicative of reduced enzymatic activity. These data are consistent with the changes found in liver PI levels of a *MBOAT7* KO mouse<sup>46</sup>. Taken together, these data suggest that rs641738, or the variant causing decreased *MBOAT7* expression, predisposes to NAFLD/NASH by affecting the acyl remodeling of phosphatidylinositol in the liver. Interestingly, it has recently been shown that genetic deletion of liver *Mboat5*, which has similar enzymatic activity, is associated with fatty liver development due to altered lipid kinetics within hepatocytes<sup>47</sup>. Furthermore, higher levels of circulating bioactive proinflammatory metabolites of arachidonic acid were reported in patients with NASH<sup>48</sup>. Several limitations of this study should be noted. The analyses of hepatic fat content and hepatic steatosis in the DHS stratified by ethnic group were underpowered, given the modest effect size of *MBOAT7/TMC4* rs641738. Biopsies were performed in two different settings: either for suspected NASH or at the time of bariatric surgery and this may affect the true estimate of the effect of the genotypes. The impact of the *MBOAT7* genetic variant on all histological traits in NAFLD was smaller than that of acquired risk factors, such as hyperglycemia. Our analyses were not designed to predict NAFLD progression or to see whether the addition of genetic variants to clinical parameters may improve their discrimination accuracy. Finally, an independent replication cohort would have further strengthened the results concerning the association of the *MBOAT7* variant with progressive liver disease.

In conclusion, we showed an association between variation in the *MBOAT7* locus, and the development and severity of NAFLD in individuals of European descent. This association is mediated by lower hepatic protein expression of MBOAT7 resulting in changes in the hepatic phosphatidylinositol acyl-chain remodeling. Further studies are needed to understand whether rs641738 is the causal variant or is in linkage disequilibrium with a causal variant, and to elucidate the mechanism linking altered hepatic PI remodelling to NAFLD development and progression.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

**Grant support:** This work was supported by the Swedish Research Council [Vetenskapsrådet (VR), 254439006], the Swedish Heart Lung Foundation [244439007], the Swedish federal government funding under the Agreement on Medical Training and Medical Research (ALF) [76290], the Novonordisk Foundation Grant for Excellence in Endocrinology [244439012], the Swedish Diabetes Foundation [DIA 2014-052] (S.R.), the Wilhelm and Martina Lundgren Science Fund (R.M.M, P.P, B.M.M and S.R.), the Nilsson-Ehle funds from the Fysiografiska Sällskapet in Lund (R.M.M.) and from Ricerca Corrente Fondazione Ca' Granda IRCCS Policlinico of Milan, Associazione Malattie Metaboliche del Fegato ONLUS, and the Fondazione Policlinico – INGM Molecular Medicine grant 2014-2016, My First AIRC Grant project code 16888 (L.V.). The Dallas Heart Study was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under award number UL1TR001105 (J.K.).

## Abbreviations

<b>ALT</b>	alanine transaminase
<b>AST</b>	aspartate transaminase
<b>BMI</b>	body mass index
<b>CI</b>	confidence interval
<b>ct</b>	threshold cycle
<b>DHS</b>	Dallas Heart Study
<b>HDL</b>	high density lipoproteins
<b>ER</b>	endoplasmic reticulum
<b>HWE</b>	Hardy-Weinberg equilibrium
<b>IFG</b>	impaired fasting glucose
<b>LD</b>	lipid droplets
<b>LDL</b>	low density lipoproteins
<b>LPIAT1</b>	lysophosphatidylinositol acyltransferase 1
<b>MAM</b>	mitochondria-associated membrane
<b>MBOAT7</b>	Membrane bound O-acyltransferase domain containing 7
<b>NAFLD</b>	nonalcoholic fatty liver disease

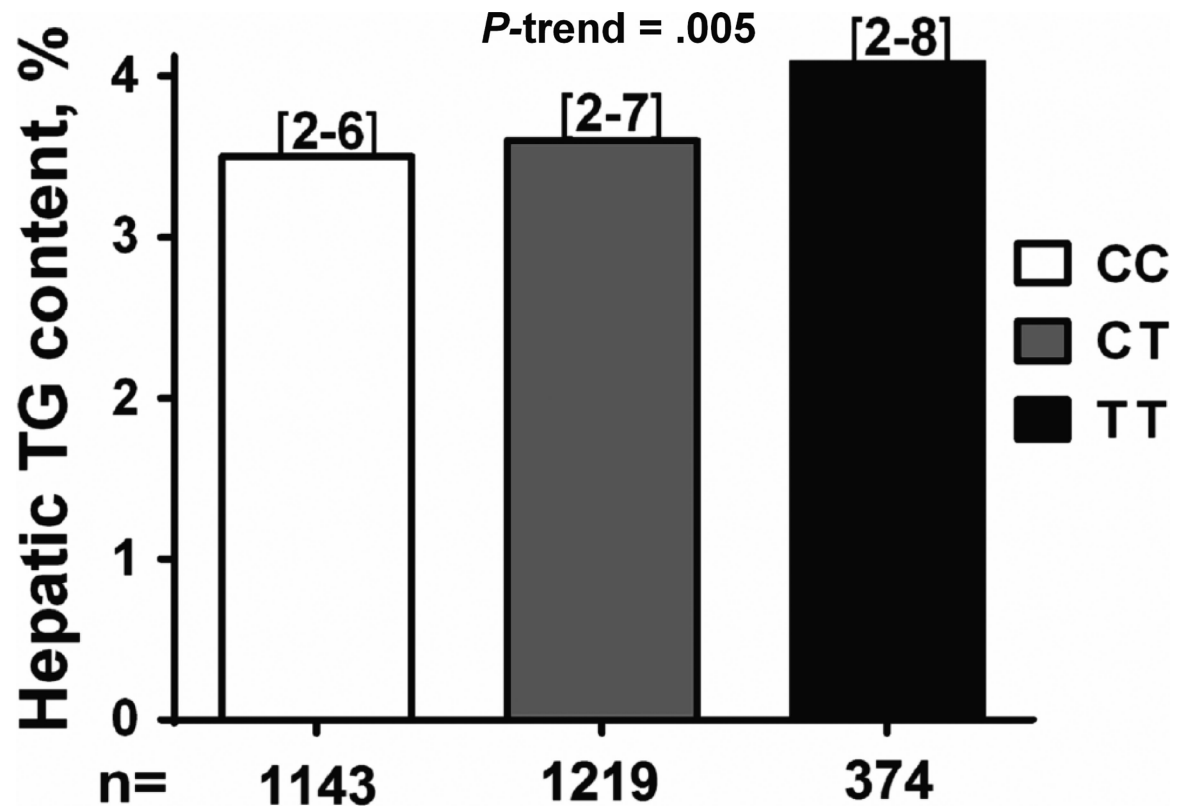
<b>NAS</b>	NAFLD Activity Score
<b>NASH</b>	nonalcoholic steatohepatitis
<b>OR</b>	odds ratio
<b>PNPLA3</b>	Patatin-like phospholipase domain-containing-3
<b>T2DM</b>	type 2 diabetes mellitus
<b>TG</b>	triglycerides
<b>TM6SF2</b>	Transmembrane 6 superfamily member 2
<b>TMC4</b>	Transmembrane Channel-Like 4
<b>PI</b>	phosphatidylinositol

## REFERENCES

- Guerrero R, Vega GL, Grundy SM, et al. Ethnic differences in hepatic steatosis: an insulin resistance paradox? *Hepatology*. 2009; 49:791–801. [PubMed: 19105205]
- Korenblat KM, Fabbrini E, Mohammed BS, et al. Liver, muscle, and adipose tissue insulin action is directly related to intrahepatic triglyceride content in obese subjects. *Gastroenterology*. 2008; 134:1369–75. [PubMed: 18355813]
- Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology*. 2005; 41:1313–21. [PubMed: 15915461]
- Day CP. From fat to inflammation. *Gastroenterology*. 2006; 130:207–10. [PubMed: 16401483]
- Bugianesi E, Leone N, Vanni E, et al. Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma. *Gastroenterology*. 2002; 123:134–40. [PubMed: 12105842]
- Singh S, Allen AM, Wang Z, et al. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. *Clin Gastroenterol Hepatol*. 2015; 13:643–54. e1–9. quiz e39–40. [PubMed: 24768810]
- Dongiovanni P, Anstee QM, Valenti L. Genetic Predisposition in NAFLD and NASH: Impact on Severity of Liver Disease and Response to Treatment. *Curr Pharm Des*. 2013
- Loomba R, Schork N, Chen CH, et al. Heritability of Hepatic Fibrosis and Steatosis Based on a Prospective Twin Study. *Gastroenterology*. 2015
- Kozlitina J, Smagris E, Stender S, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet*. 2014; 46:352–6. [PubMed: 24531328]
- Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet*. 2008; 40:1461–5. [PubMed: 18820647]
- Speliotes EK, Yerges-Armstrong LM, Wu J, 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. [PubMed: 21423719]
- Dongiovanni P, Petta S, Maglio C, et al. Transmembrane 6 superfamily member 2 gene variant disentangles nonalcoholic steatohepatitis from cardiovascular disease. *Hepatology*. 2015; 61:506–14. [PubMed: 25251399]
- Liu YL, Reeves HL, Burt AD, et al. TM6SF2 rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. *Nat Commun*. 2014; 5:4309. [PubMed: 24978903]
- Valenti L, Al-Serri A, Daly AK, et al. Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. *Hepatology*. 2010; 51:1209–17. [PubMed: 20373368]

15. Pingitore P, Pirazzi C, Mancina RM, et al. Recombinant PNPLA3 protein shows triglyceride hydrolase activity and its I148M mutation results in loss of function. *Biochim Biophys Acta*. 2014; 1841:574–80. [PubMed: 24369119]
16. Jenkins CM, Mancuso DJ, Yan W, et al. Identification, cloning, expression, and purification of three novel human calcium-independent phospholipase A2 family members possessing triacylglycerol lipase and acylglycerol transacylase activities. *J Biol Chem*. 2004; 279:48968–75. [PubMed: 15364929]
17. Huang Y, Cohen JC, Hobbs HH. Expression and characterization of a PNPLA3 protein isoform (I148M) associated with nonalcoholic fatty liver disease. *J Biol Chem*. 2011; 286:37085–93. [PubMed: 21878620]
18. Ruhanen H, Perttila J, Holttä-Vuori M, et al. PNPLA3 mediates hepatocyte triacylglycerol remodeling. *J Lipid Res*. 2014; 55:739–46. [PubMed: 24511104]
19. Pirazzi C, Adiels M, Burza MA, et al. Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro. *J Hepatol*. 2012
20. Pirazzi C, Valenti L, Motta BM, et al. PNPLA3 has retinyl-palmitate lipase activity in human hepatic stellate cells. *Hum Mol Genet*. 2014; 23:4077–85. [PubMed: 24670599]
21. Mondul A, Mancina RM, Merlo A, et al. PNPLA3 I148M Variant Influences Circulating Retinol in Adults with Nonalcoholic Fatty Liver Disease or Obesity. *J Nutr*. 2015
22. Sookoian S, Pirola CJ. Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease. *Hepatology*. 2011; 53:1883–94. [PubMed: 21381068]
23. Holmen OL, Zhang H, Fan Y, et al. Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing total cholesterol and myocardial infarction risk. *Nat Genet*. 2014; 46:345–51. [PubMed: 24633158]
24. Buch S, Stickel F, Trépo E, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. *Nat Genet*. 2015
25. Victor RG, Haley RW, Willett DL, et al. The Dallas Heart Study: a population-based probability sample for the multidisciplinary study of ethnic differences in cardiovascular health. *Am J Cardiol*. 2004; 93:1473–80. [PubMed: 15194016]
26. Browning JD, Szczepaniak LS, Dobbins R, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *Hepatology*. 2004; 40:1387–95. [PubMed: 15565570]
27. Szczepaniak LS, Babcock EE, Schick F, et al. Measurement of intracellular triglyceride stores by H spectroscopy: validation in vivo. *Am J Physiol*. 1999; 276:E977–89. [PubMed: 10329993]
28. Valenti L, Nobili V, Al-Serri A, et al. The APOC3 T-455C and C-482T promoter region polymorphisms are not associated with the severity of liver damage independently of PNPLA3 I148M genotype in patients with nonalcoholic fatty liver. *J Hepatol*. 2011; 55:1409–14. [PubMed: 21777557]
29. Petta S, Miele L, Bugianesi E, et al. Glucokinase regulatory protein gene polymorphism affects liver fibrosis in non-alcoholic fatty liver disease. *PLoS One*. 2014; 9:e87523. [PubMed: 24498332]
30. Valenti L, Alisi A, Galmozzi E, et al. I148M patatin-like phospholipase domain-containing 3 gene variant and severity of pediatric nonalcoholic fatty liver disease. *Hepatology*. 2010; 52:1274–1280. [PubMed: 20648474]
31. Simonen M, Männistö V, Leppänen J, et al. Desmosterol in human nonalcoholic steatohepatitis. *Hepatology*. 2013; 58:976–82. [PubMed: 23447451]
32. Dongiovanni P, Petta S, Mannisto V, et al. Statin use and non-alcoholic steatohepatitis in at risk individuals. *J Hepatol*. 2015; 63:705–12. [PubMed: 25980762]
33. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet*. 2005; 76:887–93. [PubMed: 15789306]
34. Wieckowski MR, Giorgi C, Lebedzinska M, et al. Isolation of mitochondria-associated membranes and mitochondria from animal tissues and cells. *Nat Protoc*. 2009; 4:1582–90. [PubMed: 19816421]
35. Ding Y, Zhang S, Yang L, et al. Isolating lipid droplets from multiple species. *Nat Protoc*. 2013; 8:43–51. [PubMed: 23222457]

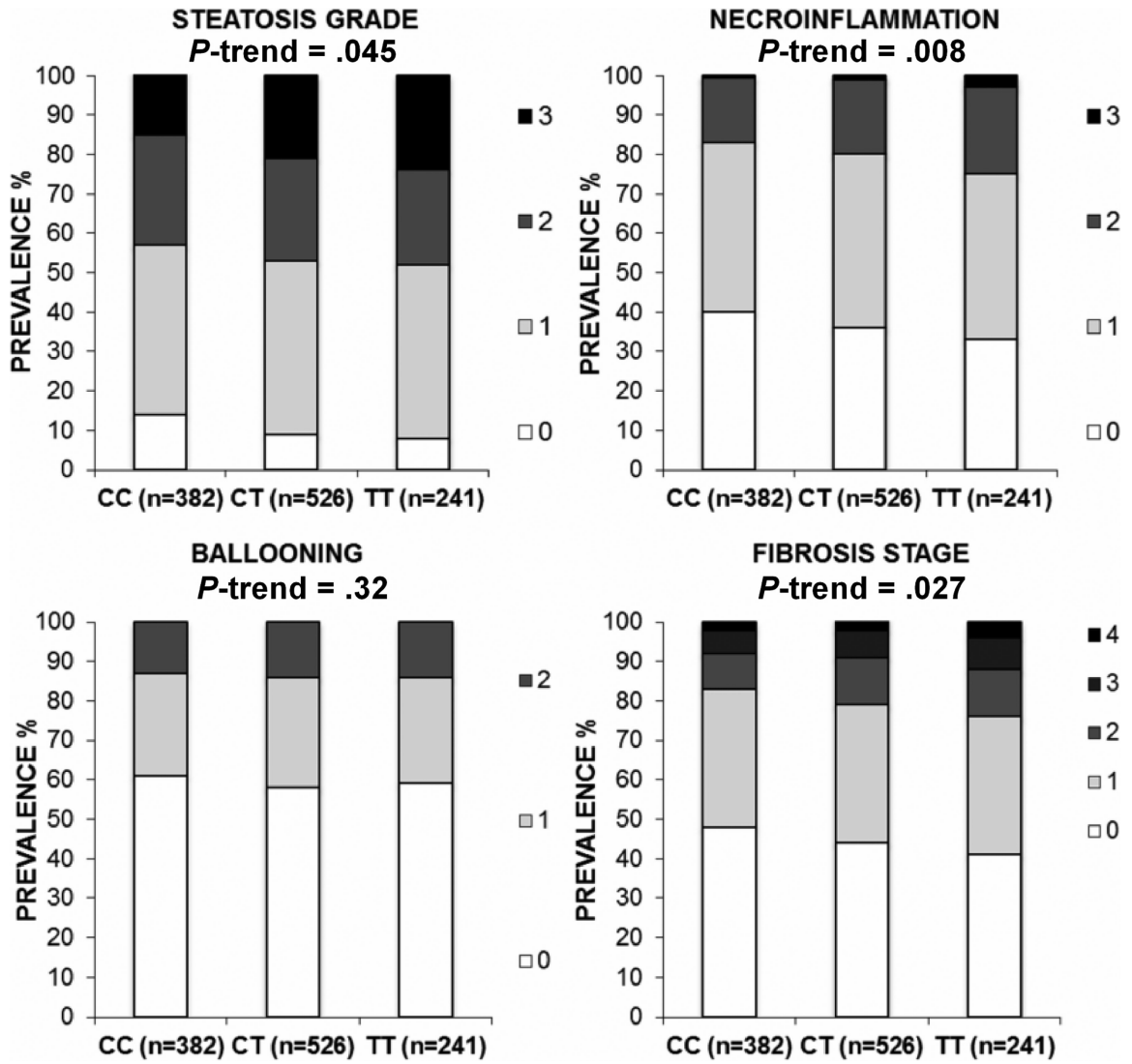
36. Stickel F, Buch S, Lau K, et al. Genetic variation in the PNPLA3 gene is associated with alcoholic liver injury in caucasians. *Hepatology*. 2011; 53:86–95. [PubMed: 21254164]
37. Colombo F, Baldan F, Mazzucchelli S, et al. Evidence of distinct tumour-propagating cell populations with different properties in primary human hepatocellular carcinoma. *PLoS One*. 2011; 6:e21369. [PubMed: 21731718]
38. Gijon MA, Riekhof WR, Zarini S, et al. Lysophospholipid acyltransferases and arachidonate recycling in human neutrophils. *J Biol Chem*. 2008; 283:30235–45. [PubMed: 18772128]
39. Ekstedt M, Hagstrom H, Nasr P, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology*. 2015; 61:1547–54. [PubMed: 25125077]
40. Angulo P, Kleiner DE, Dam-Larsen S, et al. Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology*. 2015; 149:389–97. e10. [PubMed: 25935633]
41. Hirata Y, Yamamori N, Kono N, et al. Identification of small subunit of serine palmitoyltransferase a as a lysophosphatidylinositol acyltransferase 1-interacting protein. *Genes Cells*. 2013; 18:397–409. [PubMed: 23510452]
42. Abecasis GR, Auton A, Brooks LD, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012; 491:56–65. [PubMed: 23128226]
43. D'Souza K, Epand RM. Enrichment of phosphatidylinositols with specific acyl chains. *Biochim Biophys Acta*. 2014; 1838:1501–8. [PubMed: 24120446]
44. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet*. 2014; 46:543–50. [PubMed: 24816252]
45. Gieger C, Geistlinger L, Altmaier E, et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. *PLoS Genet*. 2008; 4:e1000282. [PubMed: 19043545]
46. Anderson KE, Kielkowska A, Durrant TN, et al. Lysophosphatidylinositolacyltransferase-1 (LPIAT1) is required to maintain physiological levels of PtdIns and PtdInsP(2) in the mouse. *PLoS One*. 2013; 8:e58425. [PubMed: 23472195]
47. Rong X, Wang B, Dunham MM, et al. Lpcat3-dependent production of arachidonoyl phospholipids is a key determinant of triglyceride secretion. *Elife*. 2015:4.
48. Loomba R, Quehenberger O, Armando A, et al. Polyunsaturated fatty acid metabolites as novel lipidomic biomarkers for noninvasive diagnosis of nonalcoholic steatohepatitis. *J Lipid Res*. 2015; 56:185–92. [PubMed: 25404585]



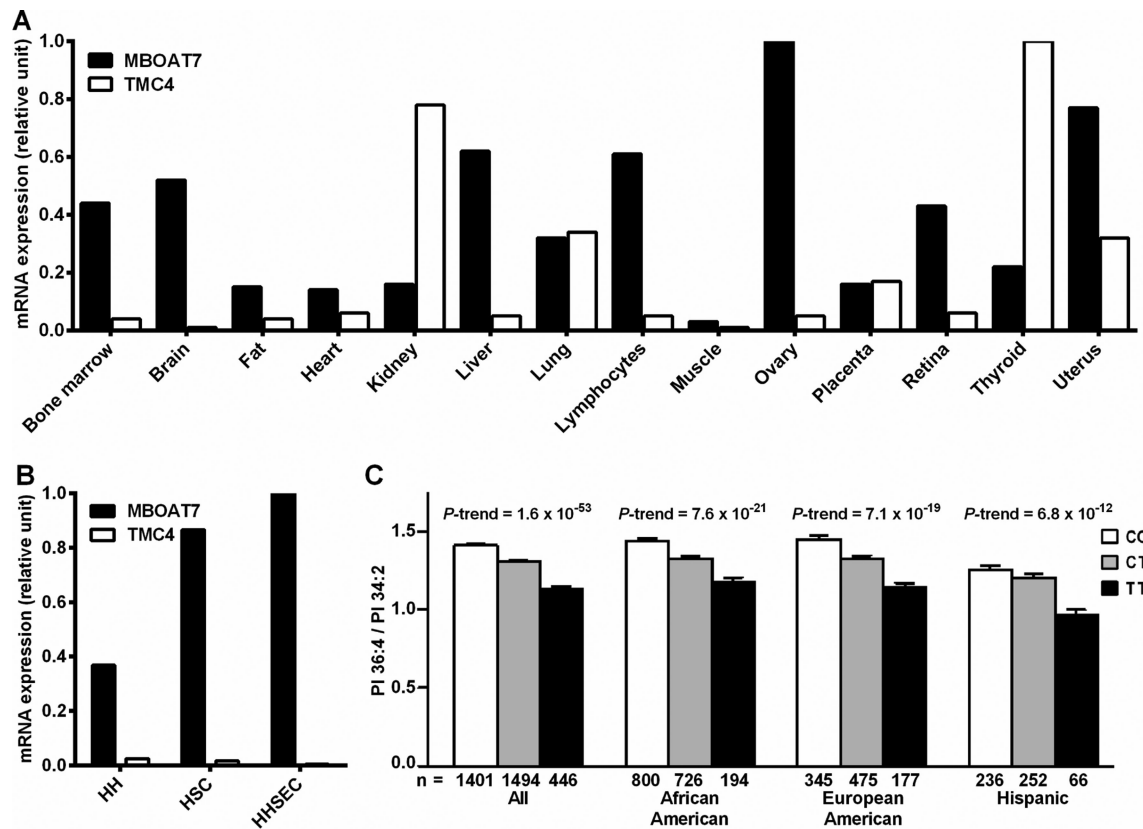
**Fig.1.**

The *MBOAT7/TMC4* rs641738 variant is associated with increased hepatic triglyceride content in the Dallas Heart Study. The association was tested by linear regression analysis adjusted for age, gender, ethnicity, BMI, IFG/T2DM, *PNPLA3*I148M and *TM6SF2*E167K genotype. Data are shown as median value. Numbers in brackets represent interquartile range. Genetic analyses were performed by using an additive model. *P*-value represents the significance of a linear trend in median trait value among genotypes.



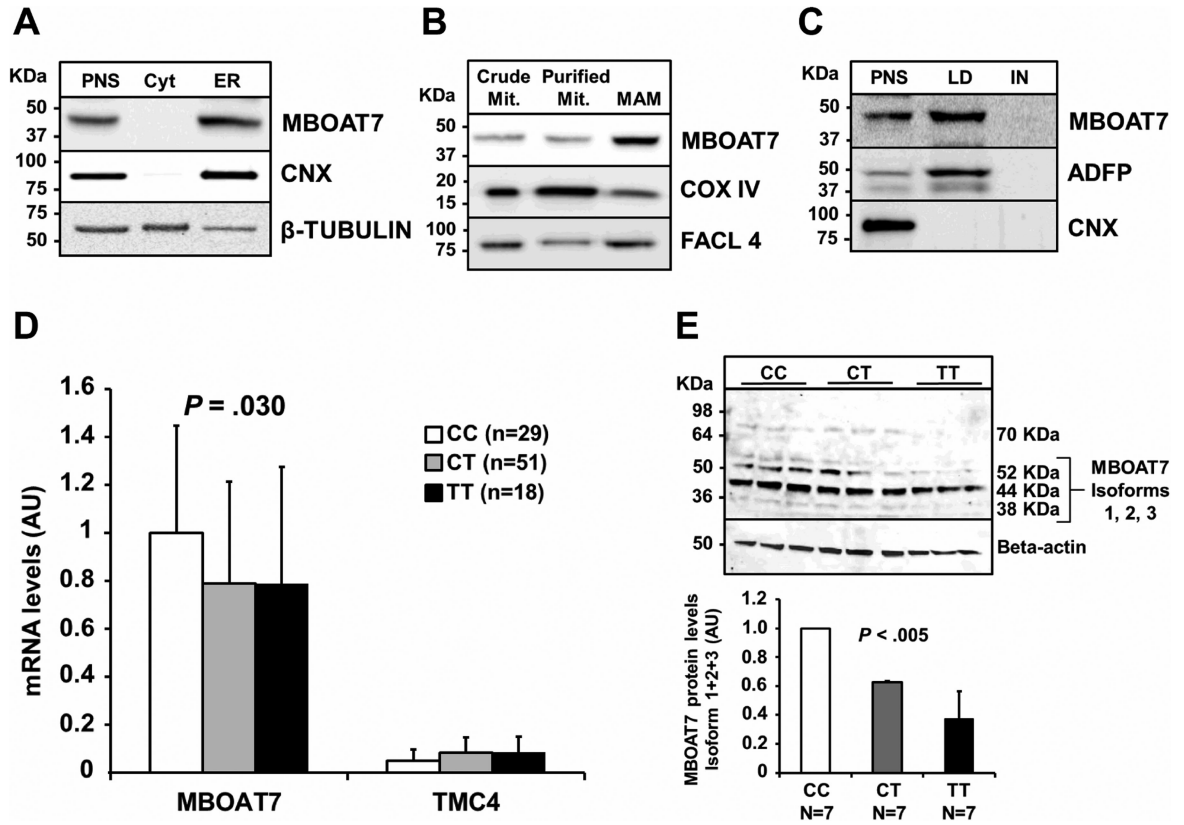


**Fig.2.** In the Liver Biopsy cross-sectional Cohort, the *MBOAT7/TMC4* rs641738 variant is associated with the severity of histological damage. The histological damage was evaluated by the different components of NAFLD activity score (NAS) and with hepatic fibrosis stage. CC: homozygotes for the C allele; CT: heterozygotes; TT: homozygotes for the T allele. The association between the *MBOAT7/TMC4* rs641738 variant and the components of the NAS has been tested by multivariate ordinal regression analysis adjusted for age, gender, BMI, presence of IFG/T2DM, number of *PNPLA3* I148M alleles, presence of the *TM6SF2* E167K variant, and indication of liver biopsy (severe obesity vs. nonalcoholic fatty liver with increased liver enzymes). Genetic analyses were calculated by using an additive model, except for *TM6SF2* where a dominant model was used due to few homozygotes for the 167K mutant allele. *P*-values represent the significance of a trend in the prevalence of more severe degree of histological damage among genotypes.

**Fig.3.**

MBOAT7 but not TMC4 is highly expressed in human liver and the rs641738 variant is associated with lower circulating levels of Phosphatidylinositol 36:4/34:2 ratio. (A) Distribution of MBOAT7 and TMC4 mRNA in 14 human tissues. (B) MBOAT7 and TMC4 mRNA expression in human primary hepatic cell types. Gene expression level was assessed by qPCR. The tissue or the cell line with the highest CT value was assigned the value of 1. (C) Ratio of phosphatidylinositol species containing different degree of saturation stratified by the rs643718 C>T genotype in the combined Dallas Heart Study (DHS) cohort and stratified by the three ethnic groups composing the study.

Abbreviations: MBOAT7, Membrane Bound O-Acyltransferase domain containing 7; TMC4, Transmembrane Channel-like 4; HH, Human Hepatocytes; HSC, Hepatic Stellate Cells; HHSEC, Human Hepatic Sinusoidal Endothelial Cell.



**Fig.4.** Endogenous MBOAT7 is attached to the ER, MAM, LD and *MBOAT7/TMC4*rs641738 variant decreases MBOAT7 mRNA expression and synthesis level in liver biopsies of obese patients.

(A) Post-nuclear supernatant (PNS), cytosol and endoplasmic reticulum (ER) were obtained by a differential centrifugation. (B) Crude mitochondria were used to isolate purified mitochondria and mitochondria-associated membranes (MAM) by Percoll density gradient. (C) Post-nuclear supernatant was used to isolate lipid droplets (LD) by sucrose gradient. (D) Hepatic mRNA levels of MBOAT7 and TMC4 in 98 severely obese stratified by the rs643718 C>T genotype (CC n=29; CT n=51; TT n=18). Data are expressed as fold increase as compared to the protective CC genotype. p=0.03 for the effect of the rs641738 T allele on MBOAT7 expression (additive model). Demographic, anthropometric and clinical characteristics of these patients are shown in supplementary Table 7. (E) Western blotting analysis and quantification of MBOAT7 protein levels in obese patients carrying the three *MBOAT7/TMC4* different genotypes, of whom hepatic gene expression was available (Supplemental Table 7). Seven samples for each rs641738 genotype (CC/CT/TT) were collected from liver biopsies and pooled according to genotype. All reactions (25 µg of lysates) were performed in triplicate in the same gel. Data are expressed as fold increase of the three MBOAT7 isoforms 1-3 (molecular weight: 52 KDa, 44 KDa and 38 KDa), as compared to the protective CC genotype. p<0.005 (additive model; it represents the difference in the trait examined across *MBOAT7* genotypes).

Abbreviations: CNX: calnexin; COX IV, cytochrome c oxidase, Complex IV; FAFL4, long-chain acyl-CoA synthetase; ADFP, Adipose differentiation-related protein; IN, Infranatant, MBOAT7, Membrane Bound O-Acyltransferase domain containing 7; TMC4, Transmembrane Channel-like 4.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 1**

Estimated risk of hepatic steatosis (hepatic TG content > 5.5%) in 2736 participants of the Dallas Heart Study per minor (T) allele of *MBOAT7/TMC4* rs641738 C>T.

Ethnic Group	Hepatic steatosis		
	OR	95% CI	P-value
All	1.20	1.05 - 1.37	0.006
African American	1.17	0.97 - 1.43	0.110
European American	1.37	1.09 - 1.72	0.007
Hispanic American	0.99	0.72 - 1.37	0.97

OR: odds ratio; CI: confidence interval

Odds ratios and confidence intervals were computed using logistic regression models, assuming an additive effect of rs641738 (by coding the genotypes 0, 1, 2 for CC, CT, and TT, respectively) adjusted for ethnicity, age, gender, BMI, IFG/T2DM, *PNPLA3* I148M and *TM6SF2* E167K genotypes.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**  
Independent predictors of liver damage (presence of steatosis, NASH, and fibrosis stage F2-F4) in 1149 patients of the Liver Biopsy cross-sectional Cohort.

	Steatosis			NASH			Fibrosis F2-F4		
	OR	95% CI	P-value	OR	95% CI	P-value	OR	95% CI	P-value
Age, years	1.01	0.99-1.02	0.32	1.01	1.00-1.02	0.17	1.02	1.01-1.04	<0.001
Gender, F	0.44	0.28-0.69	<0.001	0.66	0.51-0.86	0.002	0.66	0.47-0.91	0.011
BMI, Kg/m <sup>2</sup>	0.96	0.94-0.98	0.001	1.01	0.99-1.03	0.13	0.98	0.95-0.99	0.014
IFG/T2DM, yes	1.43	0.86-2.45	0.17	1.93	1.41-2.64	<0.001	2.95	2.07-4.23	<0.001
<i>PVAL3</i> , number of 1148M alleles	1.90	1.37-2.68	<0.001	1.70	1.41-2.04	<0.001	1.63	1.32-2.02	<0.001
<i>TMSF2</i> , E167K	1.64	0.85-3.50	0.14	1.62	1.12-2.39	0.010	1.41	0.92-2.12	0.11
rs641738, number of T alleles	1.42	1.07-1.91	0.015	1.18	1.00-1.40	0.050	1.30	1.06-1.70	0.012

OR: odds ratio; CI: confidence interval; BMI: body mass index, IFG: impaired fasting glucose; T2DM: type 2 diabetes mellitus.

Logistic regression models were used to test the association of demographic, clinical, and genetic factors with liver damage. In addition to predictors shown above, models were adjusted for statin use and indication for liver biopsy (severe obesity vs. increased liver enzymes in NAFLD. Genetic analyses were calculated by using an additive model, except for *TMSF2* where a dominant model was used due to few homozygotes for the 167K mutant allele.

The reference group was defined by the absence of steatosis, NASH and fibrosis (FO-F1). NASH diagnosis was based on the criteria reported in the Methods section.

**Table 3**  
Association of rs641738 with plasma concentrations of phosphatidylinositol (PI) species in DHS.

Metabolite	N	rs641738					P-value
		CC	CT	TT	Beta (SE)		
PI 32:1	3341	0.050 (0.031-0.087)	0.058 (0.034-0.102)	0.064 (0.041-0.131)	0.15 (0.02)	$1.5 \times 10^{-9}$	
PI 34:1	3341	0.31 (0.23-0.41)	0.33 (0.25-0.45)	0.37 (0.27-0.51)	0.19 (0.03)	$1.4 \times 10^{-14}$	
PI 34:2	3341	0.48 (0.37-0.63)	0.496 (0.38-0.66)	0.561 (0.43-0.72)	0.19 (0.02)	$2.0 \times 10^{-14}$	
PI 36:1	3249	0.25 (0.18-0.35)	0.27 (0.19-0.37)	0.29 (0.21-0.40)	0.14 (0.03)	$4.1 \times 10^{-8}$	
PI 36:2	3341	1.00 (0.79-1.25)	1.04 (0.83-1.33)	1.12 (0.92-1.45)	0.20 (0.02)	$3.6 \times 10^{-16}$	
PI 36:3 11	3341	0.30 (0.23-0.39)	0.31 (0.24-0.40)	0.33 (0.25-0.45)	0.09 (0.02)	0.0004	
PI 36:4	3341	0.64 (0.49-0.85)	0.63 (0.48-0.82)	0.60 (0.46-0.80)	-0.08 (0.02)	0.0013	
PI 38:3 11	3341	0.444 (0.346-0.613)	0.436 (0.336-0.583)	0.423 (0.318-0.545)	-0.15 (0.02)	$1.9 \times 10^{-10}$	
PI 38:4	3341	3.38 (2.56-4.51)	3.39 (2.51-4.52)	3.28 (2.43-4.61)	-0.03 (0.03)	0.21	
PI 38:5	3341	0.52 (0.36-0.73)	0.50 (0.34-0.69)	0.48 (0.34-0.66)	-0.10 (0.02)	$2.6 \times 10^{-5}$	
PI 38:6	3341	0.069 (0.045-0.098)	0.072 (0.046-0.102)	0.080 (0.052-0.113)	0.15 (0.02)	$9.2 \times 10^{-10}$	
PI 40:3	3341	0.33 (0.11-0.69)	0.31 (0.12-0.67)	0.30 (0.11-0.59)	-0.02 (0.03)	0.50	
PI 40:4	3249	0.049 (0.034-0.078)	0.051 (0.035-0.078)	0.053 (0.036-0.085)	0.06 (0.03)	0.026	
PI 40:6	3247	0.14 (0.10-0.20)	0.15 (0.10-0.21)	0.17 (0.11-0.24)	0.14 (0.03)	$5.8 \times 10^{-8}$	
PI 32:1 / Total PI	3341	0.0062 (0.0038 - 0.01)	0.0068 (0.0041 - 0.011)	0.0077 (0.005 - 0.0127)	0.16 (0.03)	$1.3 \times 10^{-10}$	
PI 34:1 / Total PI	3341	0.036 (0.029 - 0.046)	0.039 (0.031 - 0.049)	0.044 (0.035 - 0.053)	0.21 (0.03)	$2.7 \times 10^{-16}$	
PI 34:2 / Total PI	3341	0.056 (0.045 - 0.07)	0.059 (0.047 - 0.073)	0.065 (0.051 - 0.08)	0.21 (0.02)	$1.7 \times 10^{-17}$	
PI 36:1 / Total PI	3249	0.03 (0.022 - 0.04)	0.032 (0.024 - 0.042)	0.035 (0.025 - 0.044)	0.12 (0.03)	$1.5 \times 10^{-6}$	
PI 36:2 / Total PI	3341	0.12 (0.1 - 0.14)	0.13 (0.1 - 0.15)	0.13 (0.11 - 0.16)	0.21 (0.03)	$2.7 \times 10^{-16}$	
PI 36:3 11 / Total PI	3341	0.036 (0.028 - 0.044)	0.036 (0.028 - 0.045)	0.038 (0.031 - 0.05)	0.09 (0.02)	$3.4 \times 10^{-4}$	
PI 36:4 / Total PI	3341	0.075 (0.06 - 0.094)	0.074 (0.058 - 0.091)	0.071 (0.055 - 0.086)	-0.14 (0.03)	$8.7 \times 10^{-8}$	
PI 38:3 11 / Total PI	3341	0.054 (0.042 - 0.069)	0.052 (0.041 - 0.066)	0.05 (0.039 - 0.061)	-0.18 (0.02)	$2.0 \times 10^{-13}$	
PI 38:4 / Total PI	3341	0.41 (0.33 - 0.49)	0.4 (0.33 - 0.48)	0.39 (0.32 - 0.46)	-0.07 (0.03)	0.0058	
PI 38:5 / Total PI	3341	0.064 (0.042 - 0.085)	0.06 (0.042 - 0.081)	0.058 (0.041 - 0.076)	-0.12 (0.03)	$1.6 \times 10^{-6}$	
PI 38:6 / Total PI	3341	0.0082 (0.0056 - 0.011)	0.0085 (0.0059 - 0.0113)	0.0091 (0.0063 - 0.0126)	0.15 (0.02)	$4.6 \times 10^{-9}$	
PI 40:3 / Total PI	3341	0.038 (0.014 - 0.078)	0.036 (0.014 - 0.081)	0.033 (0.013 - 0.07)	-0.02 (0.03)	0.41	

rs641738						
Metabolite	N	CC	CT	TT	Beta (SE)	P-value
PI 40:4 / Total PI	3249	0.006 (0.0042 - 0.0088)	0.0062 (0.0043 - 0.009)	0.0064 (0.0044 - 0.0087)	0.04 (0.03)	0.10
PI 40:6 / Total PI	3247	0.018 (0.011 - 0.024)	0.018 (0.012 - 0.025)	0.02 (0.014 - 0.027)	0.11 (0.03)	$1.8 \times 10^{-5}$
PI 36:3 i1 / PI 34:2	3341	0.63 (0.5 - 0.78)	0.61 (0.49 - 0.78)	0.61 (0.48 - 0.76)	-0.12 (0.03)	$3.0 \times 10^{-6}$
PI 36:4 / PI 34:2	3341	1.4 (1.1 - 1.7)	1.3 (1 - 1.5)	1.1 (0.9 - 1.3)	-0.38 (0.02)	$1.6 \times 10^{-53}$
PI 38:3 i1 / PI 36:2	3341	0.46 (0.37 - 0.56)	0.42 (0.34 - 0.52)	0.37 (0.3 - 0.45)	-0.39 (0.02)	$5.1 \times 10^{-62}$
PI 38:4 / PI 34:1	3341	10.9 (7.5 - 16.1)	10 (6.9 - 14.7)	8.8 (5.8 - 13.3)	-0.15 (0.03)	$1.1 \times 10^{-9}$
PI 38:5 / PI 36:2	3341	0.54 (0.39 - 0.73)	0.49 (0.34 - 0.66)	0.44 (0.31 - 0.58)	-0.28 (0.02)	$8.1 \times 10^{-30}$
PI 38:6 / PI 38:3 i1	3341	0.15 (0.1 - 0.22)	0.16 (0.11 - 0.23)	0.19 (0.12 - 0.27)	0.26 (0.02)	$4.2 \times 10^{-29}$
PI 40:3 / PI 34:1	3341	0.99 (0.36 - 2.48)	0.91 (0.34 - 2.21)	0.74 (0.27 - 1.78)	-0.11 (0.03)	$5.4 \times 10^{-6}$
PI 40:4 / PI 38:3 i1	3249	0.11 (0.08 - 0.15)	0.12 (0.09 - 0.16)	0.13 (0.1 - 0.17)	0.20 (0.02)	$1.3 \times 10^{-16}$
PI 40:6 / PI 38:3 i1	3247	0.33 (0.21 - 0.46)	0.36 (0.24 - 0.5)	0.43 (0.28 - 0.55)	0.24 (0.02)	$5.5 \times 10^{-24}$

Values are median (interquartile range). Concentrations and ratios of concentrations of PI species were compared across genotypes using linear regression, adjusted for age, gender, and ancestry. Additive model was assumed for the genotype. Metabolite concentrations were transformed using an inverse normal transformation prior to analysis to achieve approximate normality. Beta coefficients are in s.d. units. i: denotes isomers of lipid subclass