Gastroenterology 2021;160:1634–1646

Exome-Wide Association Study on Alanine Aminotransferase Identifies Sequence Variants in the *GPAM* and *APOE* Associated With Fatty Liver Disease

Oveis Jamialahmadi,[1] Rosellina Margherita Mancina,[1] Ester Ciociola,[1] Federica Tavaglione,[1,2] Panu K. Luukkonen,[3,4,5] Guido Baselli,[6] Francesco Malvestiti,[7] Dorothée Thuillier,[8] Violeta Raverdy,[8,9] Ville Männistö,[10,11,12] Rosaria Maria Pipitone,[13] Grazia Pennisi,[13] Daniele Prati,[6] Rocco Spagnuolo,[14] Salvatore Petta,[13] Jussi Pihlajamäki,[11,12] François Pattou,[8,9] Hannele Yki-Järvinen,[3,4] Luca Valenti,[6,7] and Stefano Romeo[1,14,15]

[1]Department of Molecular and Clinical Medicine, Institute of Medicine, Sahlgrenska Academy, Wallenberg Laboratory, University of Gothenburg, Gothenburg, Sweden; [2]Clinical Medicine and Hepatology Unit, Department of Internal Medicine and Geriatrics, Campus Bio-Medico University, Rome, Italy; [3]Department of Medicine, University of Helsinki and Helsinki University Hospital, Helsinki, Finland; [4]Minerva Foundation Institute for Medical Research, Helsinki, Finland; [5]Department of Internal Medicine, Yale University, New Haven, Connecticut; [6]Translational Medicine, Department of Transfusion Medicine and Hematology, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico Ca' Granda Ospedale Maggiore Policlinico, Milano, Italy; [7]Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milano, Italy; [8]Univ Lille, Inserm, Lille Pasteur Institute, Centre Hospitalier Universitaire de Lille, European Genomic Institute for Diabetes, U1190 Translational Research in Diabetes, Lille University, Lille, France; [9]Centre Hospitalier Universitaire de Lille, Department of General and Endocrine Surgery, Integrated Center for Obesity, Lille, France; [10]Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Finland; [11]Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Finland; [12]Department of Medicine, Endocrinology and Clinical Nutrition, Kuopio University Hospital, Finland; [13]Section of Gastroenterology and Hepatology, Promozione della Salute, Materno-Infantile, di Medicina Interna e Specialistica di Eccellenza "G. D'Alessandro," University of Palermo, Palermo, Italy; [14]Clinical Nutrition Unit, Department of Medical and Surgical Sciences, University Magna Graecia, Catanzaro, Italy; and [15]Cardiology Department, Sahlgrenska University Hospital, Gothenburg, Sweden

BACKGROUND & AIMS: Fatty liver disease (FLD) is a growing epidemic that is expected to be the leading cause of end-stage liver disease within the next decade. Both environmental and genetic factors contribute to the susceptibility of FLD. Several genetic variants contributing to FLD have been identified in exome-wide association studies. However, there is still a missing hereditability indicating that other genetic variants are yet to be discovered. **METHODS:** To find genes involved in FLD, we first examined the association of missense and nonsense variants with alanine aminotransferase at an exome-wide level in 425,671 participants from the UK Biobank. We then validated genetic variants with liver fat content in 8930 participants in whom liver fat measurement was available, and replicated 2 genetic variants in 3 independent cohorts comprising 2621 individuals with available liver biopsy. **RESULTS:** We identified 190 genetic variants independently associated with alanine aminotransferase after correcting for multiple testing with Bonferroni method. The majority of these variants were not previously associated with this trait. Among those associated, there was a striking enrichment of genetic variants influencing lipid metabolism. We identified the variants rs2792751 in *GPAM/GPAT1*, the gene encoding glycerol-3-phosphate acyltransferase, mitochondrial, and rs429358 in *APOE*, the gene encoding apolipoprotein E, as robustly associated with liver fat content and liver disease after adjusting for multiple testing. Both genes affect lipid metabolism in the liver. **CONCLUSIONS:** We identified 2 novel genetic variants in *GPAM* and *APOE* that are robustly associated with steatosis and liver damage. These findings may help to better elucidate the genetic susceptibility to FLD onset and progression.

Keywords: Nonalcoholic Fatty Liver Disease; NAFLD; Transaminase; Metabolic Associated Fatty Liver Disease; MAFLD.

Fatty liver disease (FLD) is a growing epidemic and it is estimated to become the leading cause of end-stage liver disease within the next 10 years.[1] Although environmental factors are known to play a role, heritability accounts for a large fraction of inter-individual variability in hepatic fat content and FLD susceptibility.[2]

Previous exome-wide, genome-wide, and candidate gene association studies have identified genetic variants in *PNPLA3*, *TM6SF2*, *GCKR*, *MBOAT7* as genetic determinants of FLD.[3–7] A successful strategy in some of these studies was to restrict the analysis to the coding sequence of the genome. By examining only nonsense and missense variations, the number of variants tested is markedly reduced, thereby

Abbreviations used in this paper: **ALT, alanine aminotransferase; APOE, apolipoprotein E; BMI, body mass index; FLD, fatty liver disease; GPAM, glycerol-3-phosphate acyltransferase, mitochondrial; GWAS, genome-wide association study; ICD-10, International Classification of Diseases, 10th edition; LD, linkage disequilibrium; MAF, minor allele frequency; MRI, magnetic resonance imaging; PDFF, proton density fat fraction; SNP, single nucleotide polymorphism.**

Most current article

0016-5085
https://doi.org/10.1053/j.gastro.2020.12.023

WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT

Fatty liver disease (FLD) is becoming the leading cause of end-stage liver disease. Although several genetic variants contributing to FLD have been identified, there is still a missing hereditability.

NEW FINDINGS

We identified 190 genetic variants independently associated with ALT at an exome-wide level, and further identified *GPAM* rs2792751 and *APOE* rs429358 to be robustly associated with liver fat content and liver disease in the UK Biobank and validation cohorts. Our data demonstrate a tight relationship between liver damage and lipid biology.

LIMITATIONS

Smaller sample size of participants with PDFF data compared to those with ALT in the UK Biobank may hinder the detection of more genetic variants associated with FLD.

IMPACT

Our results may help to better elucidate the genetic susceptibility to FLD and the molecular mechanisms underlying the disease onset and progression.

maximizing the odds to find novel genetic determinants of traits. Indeed, an exome-wide study starting from examining alanine aminotransferase (ALT), commonly associated with liver fat, and aspartate aminotransferase, more closely related to liver damage, led to the discovery of a gene that contributes to protection against FLD, namely *HSD17B13*.[8] Consistently, a genome-wide association study (GWAS) of cirrhosis has also identified a protective variant in *MARC1* associated with lower ALT and liver fat.[9] However, there is still a missing hereditability indicating that other genetic variants remain to be discovered.[2]

The hallmark of FLD is the accumulation of neutral fat in hepatocytes, which is associated with a spectrum of liver damage ranging from uncomplicated steatosis to inflammation, ballooning, and fibrosis. Strikingly, among the currently known common genetic determinants of FLD, the vast majority are in genes involved in hepatic lipid handling.[10] *PNPLA3* is involved in lipid droplet remodeling.[3,10] *TM6SF2* in hepatic lipid export,[4,11] and *GCKR*[5] and *MBOAT7*[6,12,13] increase triglyceride synthesis through canonical and noncanonical pathways of de novo lipogenesis, respectively. There is also evidence that rare variants contributing to severe FLD are involved in hepatic lipid export.[14] Mendelian randomization studies also provided evidence that quantitative and qualitative alterations in liver fat content are per se deleterious causing inflammation and fibrosis, although the underlying mechanisms are still not known.[10,15]

In this study, we aimed to identify new genetic determinants of FLD. We started by examining the association between all common missense and nonsense variations and ALT in 425,671 participants of the UK Biobank cohort. We then validated our results by examining 8930 individuals from this study for whom liver fat content measurement was available, and in 3 independent European cohorts comprising a total of 2621 individuals, for whom histologic evaluation of liver damage was available.

Material and Methods

UK Biobank

The UK Biobank is a large-scale national cohort study of approximately 500,000 participants aged between 40 and 69 years who visited 22 assessment centers throughout the UK between 2006 and 2010, with comprehensive baseline assessment and collection of genetic data. The UK Biobank received ethical approval from the National Research Ethics Service Committee North West Multi-Centre Haydock (reference 16/NW/0274).[16] Data used in this study were obtained under application number 37142.

Here, we used the European subset of UK Biobank individuals by adding participants who self-reported as being "Irish" or "any other White background" (after removal of outliers based on first 6 genetic principal components) to the subset of White British ancestry. Next, we excluded individuals with more than 10 putative third-degree relatives, with a mismatch between their self-reported and genetically inferred sex, having putative sex chromosome aneuploidy, who had withdrawn consent, and were identified by the UK Biobank as outliers based on heterozygosity and missingness.[16–18]

Definition of Chronic Liver Disease in the UK Biobank

We used the International Classification of Diseases, 10th edition (ICD-10) codes from hospitalization (data field 41270) and underlying primary and secondary cause of death (data fields 40001 and 40002) to define chronic liver disease (K70.0, K70.1, K70.2, K70.3, K70.4, K70.9, K72.1, K72.9, K73.0, K73.1, K73.2, K73.8, K73.9, K74.0, K74.1, K74.2, K74.6, K76.0, K76.6, K76.7, K76.8, K76.9, I85.0, I85.9) and cirrhosis (K70.3, K70.4, K72.1, K72.9, K74.1, K74.2, K74.6, K76.6, K76.7, I85.0, I85.9). We then excluded individuals with chronic viral hepatitis diagnosis (ICD-10 codes B18.0, B18.1, B18.2, B18.8, B18.9, B19.0, B19.9) from the analyses (Supplementary Table 1).

Liver Fat Content Measurement in the UK Biobank

We used derived magnetic resonance imaging (MRI) liver fat, as measured by proton density fat fraction (PDFF) (data field 22436), to quantify liver fat content in participants from the UK Biobank. The details on liver MRI protocols and analyses can be found elsewhere.[19,20] Briefly, individuals were scanned with a Siemens MAGNETOM Aera 1.5-T MRI scanner using a 6-minute dual-echo Dixon Vibe protocol, providing a water and fat separated volumetric dataset covering neck to knees. A single multi-echo slice was further acquired to analyze the liver PDFF.

Genotyping and Imputation

UK Biobank. UK Biobank participants were genotyped using 2 highly similar (>95% overlap) genotyping arrays: UK BiLEVE (approximately 50,000 individuals) or UK Biobank Axiom arrays (Affymetrix, approximately 450,000 individuals).

Following single nucleotide polymorphism (SNP) and sample quality controls, directly genotyped data were then imputed centrally by the UK Biobank based on the 1000 Genomes Phase 3, UK 10K haplotype, and Haplotype Reference Consortium reference panels.[21] From approximately 97 million variants, we further excluded variants with a minor allele frequency (MAF) <0.01, imputation INFO score <0.8, and Hardy–Weinberg equilibrium $P < 10^{-10}$, which resulted in a set of 9,356,431 variants.

Next, variant annotation was performed with snpEff[22] and Variant Effect Predictor,[23] and mutations resulting in a premature stop codon (stop gained), loss of a start (start loss) or stop codon (stop loss), disruption of canonical splice dinucleotides (splice acceptor and splice donor variants), insertion or deletion of a frameshift (frameshift variant), insertion or deletion causes a frameshift (inframe insertion and deletion), and codons producing a different amino acid (missense) were considered for association analyses (final set of 33,926 variants).

French cohort. A total of 1331 individuals from the biological Atlas of Severe Obesity (ABOS, ClinicalTrials.gov, Number NCT01129297) were included in the analyses. Briefly, all individuals were severely obese adults who fulfilled the criteria for weight loss surgery. More than 90% of this cohort were of European ancestry. Individuals were excluded if they had high current alcohol intake (daily alcohol consumption ≥20 g/d for women and ≥30 g/d for men); a history of excessive drinking for a period longer than 2 years at any time in the past 20 years; a long-term consumption of hepatotoxic drugs, positive screening for chronic liver diseases, including positive testing for hepatitis B surface antigen and hepatitis C virus antibodies; evidence of genetic hemochromatosis, or were aged younger than 18 years. Biopsies were managed at Lille University Hospital Biobank (CRB/CIC1403, brief registration number: BB0033–00030). Genotyping was performed by using KASPar chemistry according to manufacturer instructions by Genoscreen (https://www.genoscreen.fr/en/). All participants gave written informed consent to participate in the study.

Italian cohort. A total of 466 individuals from the Gastrointestinal and Liver Unit of the Palermo University Hospital, Palermo, Italy, were included in the study.[24] Briefly, individuals with high alcohol intake (men, >30 g/d; women, >20 g/d), viral and autoimmune hepatitis, or other causes of liver disease were excluded. The diagnosis of nonalcoholic steatohepatitis was based on the presence of steatosis with lobular necroinflammation and ballooning or fibrosis. The study was approved by the Ethics Committees of Palermo University Hospital (Palermo). All participants gave written informed consent to participate to the study. Participants were genotyped for the variants rs2792751 *GPAM* and rs429358 *APOE* by TaqMan SNP Genotyping Assays (ThermoFisher Scientific, Waltham, MA). All genotypes were performed in duplicate with 100% concordance rate.

Finnish cohort. In this study, 512 individuals from the Northern Savo Hospital District, Kuopio, Finland,[25] and 312 from the Hospital District of Helsinki and Uusimaa, Finland,[26] were included. Briefly, individuals with high alcohol intake (men, >30 g/d; women, >20 g/d), viral and autoimmune hepatitis, or other causes of liver disease were excluded. The diagnosis of nonalcoholic steatohepatitis was based on the presence of steatosis with lobular necroinflammation and ballooning or fibrosis. The study was approved by the Ethics Committees of the Northern Savo Hospital District in Kuopio (Finland) and the ethics committee of the Hospital District of Helsinki and Uusimaa (Finland). All participants gave written informed consent to participate to the study. Participants were genotyped for the rs2792751 *GPAM* and the rs429358 *APOE* by TaqMan SNP Genotyping Assays (ThermoFisher Scientific, Waltham, MA). All genotypes were performed in duplicate with 100% concordance rate.

Genetic Association Analysis

UK Biobank. The association between imputed variant dosages, and ALT and PDFF was performed using a linear mixed-effects model implemented in BOLT-LMM, version 2.3.4.[27,28] Similarly, the association analyses for categorical traits (ie, chronic liver disease and cirrhosis) were performed by a logistic mixed-effects model implemented in SAIGE to correct for the relatedness among participants, population stratification, and the imbalanced case to control ratio.[29] For other continuous traits (ie, glucose, cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, and triglycerides), a linear regression analysis was performed. All continuous traits were rank-based inverse normal transformed before the analysis. Furthermore, all analyses were carried out under an additive genetic model and adjusted for age, sex, body mass index (BMI), the first 10 principal components of ancestry, and genotyping array.

For sensitivity analyses, genome-wide association analysis of European individuals was performed in BOLT-LMM using the full set of imputed common variants (9,356,431), adjusted for the same above-mentioned covariates. Similarly, PLINK 2 was used for the corresponding analysis in a subset of unrelated white British individuals.[21]

For both SAIGE and BOLT-LMM, a subset of high-quality and directly genotyped variants was used to estimate genetic relationship matrix. Variants on long-range linkage disequilibrium (LD) and major histocompatibility complex regions or those with missingness >0.01, MAF <0.01, or Hardy–Weinberg equilibrium $P < 10^{-15}$ were excluded. Finally, LD pruning with a windows size of 500,000 base pairs and pairwise $r^2 < 0.1$ resulted in a set of 146,883 markers.[30]

To identify independent signals from association analysis, we performed LD clumping procedure (PLINK parameters: *–clump-p1 1.47e-6 –clump-r2 0.2 –clump-kb 1000*, after excluding individuals with third-degree or closer relatives based on the pairwise kinship coefficients provided by UK Biobank),[21,30] so those variants with $P < 1.47 \times 10^{-6}$, $r^2 > 0.2$, and within 1 Mb of the index variant were clumped together. To further examine the statistically independence of variants identified by LD clumping, we performed approximate stepwise model selection in conditional and joint multiple-SNP analysis implemented in Genome-wide Complex Trait Analysis (GCTA),[31] with an LD window of 1 Mb and using 50,000 randomly selected unrelated European individuals from the UK Biobank as reference sample for LD structure.

We examined the association of statistically independent variants with liver related traits using previously reported associations ($P < 1.47 \times 10^{-6}$) in 3 comprehensive and curated genome-wide association summary statistics databases: GWAS Catalog,[32] PhenoScanner,[33,34] and openGWAS.[35]

Replication cohorts. The association between the *GPAM* rs2792751 and the *APOE* rs429358 variants and liver disease within 3 cohorts (ie, French, Italian, and Finnish) was evaluated under an additive genetic model by ordinal logistic regression analysis (severity of liver steatosis, inflammation, ballooning, and fibrosis) adjusted for age, sex, BMI, recruitment center (for the Finnish cohort), and number of *PNPLA3* I148M mutant alleles. The liver histologic grading for FLD was assessed by a local pathologist and by using Kleiner score.[36] An inverse-variance meta-analysis of 3 studies was performed with fixed and random-effects model (to capture the heterogeneity across 3 cohorts) in *"meta"* package (http://cran.r-project.org/web/packages/meta/index.html) in R, version 3.6.1.

Gene-Set Enrichment Analysis

The set of genes within exome-wide independently significant signals were used to perform gene ontology (biological processes, http://geneontology.org/), Kyoto Encyclopedia of Genes and Genomes biological pathways, disease (DisGeNET), and cell line/tissue (ARCHS4 and GTEx) enrichment analysis using Enrichr tool.[37–42] Nominal and adjusted *P* values and combined scores (log *P* value × *z* score) were reported.

Assessment of Bias in the Genome-Wide Analyses

Genomic control (calculated as the ratio between observed and median χ^2 statistics, Supplementary Figure 1) showed an inflation in association analysis results ($\lambda = 1.84$). We hypothesized that this inflation was due to the selection of genetic variants that may be more likely to result in functional changes. Therefore, we repeated the genome-wide analyses by using a total of 9,356,431 imputed common variants (Supplementary Figure 2*A*). Genomic control showed a reduction in the inflation ($\lambda = 1.43$, Supplementary Figure 2*B*) that, however, was still elevated. Because genomic control is not able to differentiate between confounding bias (eg, population stratification or cryptic relatedness) and polygenicity, we calculated LD score regression intercept on this set of total imputed common variants (n = 9,356,431).[43,44] The LD score regression intercept suggested (1.079; SE = 0.0449) most of the inflation was most probably due to the polygenicity. To further assess the potential confounding, we repeated the same analysis in the subset of unrelated White British individuals from the UK Biobank (linear regression analysis adjusted for genomic principal components) and observed a similar finding (LD score regression intercept: 1.050; SE = 0.0356).[21] Notably, this was consistent with previous findings showing the tendency of LD score regression intercept to increase with sample size and SNP heritability and we observed similar attenuation ratio between linear mixed-effect model and principal components adjusted linear regression analysis (0.07 and 0.064, respectively).[28] Moreover, by using the LD score regression analysis, we estimated the heritability of ALT in Europeans to be 0.134 (SE = 0.0125).

Linkage Disequilibrium Score Regression Analysis

We estimated heritability and confounding bias in our GWAS results with LD score regression analysis (https://github.com/bulik/ldsc/)[43] using the baseline LD model (version 2.2; https://data.broadinstitute.org/alkesgroup/LDSCORE/), which contains 97 annotations, including functional annotations and MAF-/LD-dependent architectures.[44,45] We excluded variants within the HLA region on chromosome 6 (26–34 Mb) and set LDSC parameter chisq-max to an arbitrary large number (99999) to keep high-effect associations.

Liver Transcriptomic Analyses

These analyses were conducted in 125 obese individuals from Milan, Italy, who underwent percutaneous liver biopsy performed during bariatric surgery at the Fondazione IRCCS Ca' Granda.[46] Briefly, individuals with high alcohol intake (men, > 30 g/d; women 20 g/d), viral and autoimmune hepatitis, or other causes of liver disease were excluded. Liver biopsy was performed by needle gauge. Total RNA was isolated using RNeasy mini-kit (Qiagen, Hulsterweg, Germany). RNA was sequenced in paired-end Mode (read length 150nt) using the Illumina HiSeq 4000 (Novogene, Hong Kong, China). Reads count (ENSEMBL human transcript reference assembly, version 75) was performed using RSEM package.[47] Counts were normalized using DESeq2 package.[48] To identify differentially expressed pathways, pre-ranked gene-set enrichment analysis was performed on differentially expressed or significantly correlated genes.[49,50]

Results

Exome-Wide Association Study on Alanine Aminotransferase in Europeans from the UK Biobank

To identify novel genetic predictors associated with FLD, we first examined the association between 33,926 missense and nonsense common variants (MAF > 1%) and ALT in European participants from the UK Biobank cohort after excluding those with available measurement of liver fat content by PDFF (Supplementary Table 2). The analyses were done by using a linear mixed-effects model under an additive genetic model. Following LD clumping and conditional and joint multiple-SNP analysis, we identified 190 statistically independent variants exceeding Bonferroni correction threshold ($P < 1.47 \times 10^{-6}$, Figure 1, Supplementary Figure 1, and Supplementary Table 3).[30,31] Genomic control was highly inflated indicating potential confounding bias (see Methods for more details). However, LD score regression analyses showed that inflation was mostly due to polygenicity of the trait rather than population stratification or relatedness.

Validation of the Alanine Aminotransferase Association and Enrichment Analyses

To validate our findings, we examined the association of the 190 statistically independent variants with liver-related traits in 3 public GWAS summary statistics databases: GWAS Catalog,[32] PhenoScanner,[33,34] and openGWAS.[35] We observed that 36 variants (19%) were associated with multiple liver-related traits. The majority of the previously

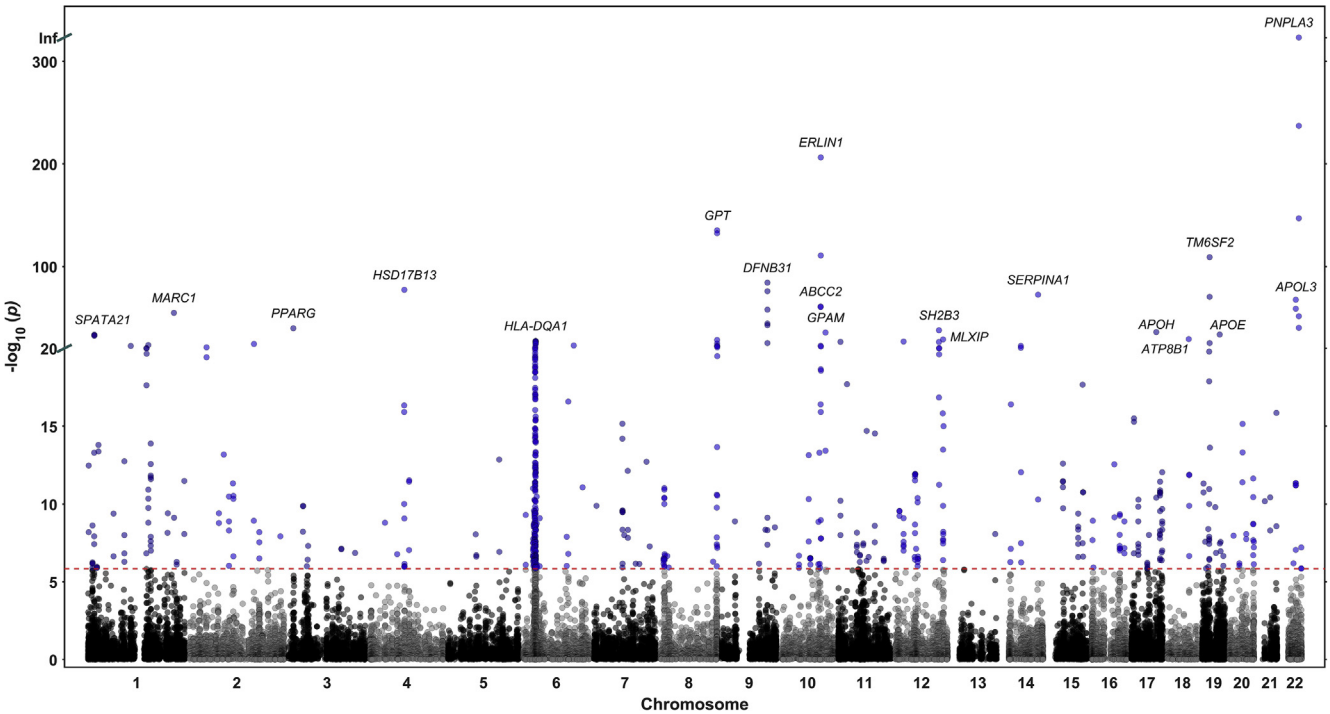


Figure 1. Manhattan plot of exome-wide association study of ALT in 425,671 European participants from the UK Biobank. A total of 33,926 nonsense and missense variants were used in the analyses. *P* values were calculated by using a linear mixed-effects model. *Red dashed line* represents the exome-wide significance threshold based on Bonferroni correction for multiple testing. *X-axis* shows chromosome number.

reported variants were associated with lipoproteins (Figure 2 and Supplementary Table 4).

To gain insight into the function of the genetic variants identified, we performed gene-set enrichment analysis in multiple genetic, biologic, tissue expression, and gene–disease association databases (after exclusion of genes in the major histocompatibility complex region). Interestingly, gene ontology biological processes showed an enrichment of processes involved in lipid homeostasis and triglyceride metabolism (Figure 3 and Supplementary Table 5). Kyoto Encyclopedias of Genes and Genomes (KEGG) showed an enrichment of pathways involved in bile secretion and insulin signaling pathways. Moreover, this set of genes was overrepresented in metabolic liver disease (DisGeNET) and mostly expressed in liver and hepatocytes (in GTEx and ARCHS4).

Association of Genetic Variants with Liver Fat Content in the UK Biobank

To refine the association identified by examining ALT, we examined the association of the 190 independent genetic variants in Europeans from the UK Biobank with available liver fat measurement by MRI-PDFF (n = 8930) by using the linear mixed-effects model. Among the 190 variants, 8 exceeded the threshold ($P = 2.63 \times 10^{-4}$) after Bonferroni correction (Supplementary Table 6). Among these, 4 were already well-known genes associated with FLD, namely *PNPLA3*, *TM6SF2*, *MARC1*, and *MBOAT7*(*TMC4*), and 1 was an independent genetic variant in *TM6SF2* (rs187429064) that had previously been found to associate with lipoproteins and diabetes (Supplementary Table 6). Three novel genetic variants were identified: rs429358 in *APOE* (which encodes apolipoprotein E), rs2792751 in *GPAM* (glycerol-3-phosphate acyltransferase, mitochondrial) and rs3128853 in *OR12D2* (olfactory receptor family 12 subfamily D member 2).

Next, we examined the association between these 8 genetic variants and chronic liver disease and cirrhosis. As expected, genetic variants in *PNPLA3*, *MBOAT7*, *TM6SF2*, and *MARC1* were associated with both traits. Genetic variants in *GPAM* ($P = .001$ and $P = .051$, respectively, encoding for p.Val43Ile) and *APOE* ($P = .003$ and $P = .014$, respectively, encoding for p.Cys112Arg), but not the sequence variant in *OR12D2*, were associated with chronic liver disease and cirrhosis (Supplementary Table 6). *GPAM* and *APOE* genetic variants were also associated with circulating lipoprotein levels (see Table 1).

Next, we examined the population attributable fraction (PAF) of the top 6 variants in *PNPLA3, TM6SF2, MBOAT7, MARC1, GPAM*, and *APOE* associated with PDFF for chronic liver disease and cirrhosis. The total PAF was >20% for chronic liver disease and >30% for cirrhosis. As expected the largest PAF was conferred by the *PNPLA3* variant. However, the second largest PAF was conferred by *APOE* and *MARC1* (Supplementary Table 7).

After excluding carriers of mutations involved in hemochromatosis (n = 13,259), namely 2,835 homozygotes for C282Y sequence variant (rs1800562) and 10,424 compound heterozygotes C282Y/H63D sequence variants

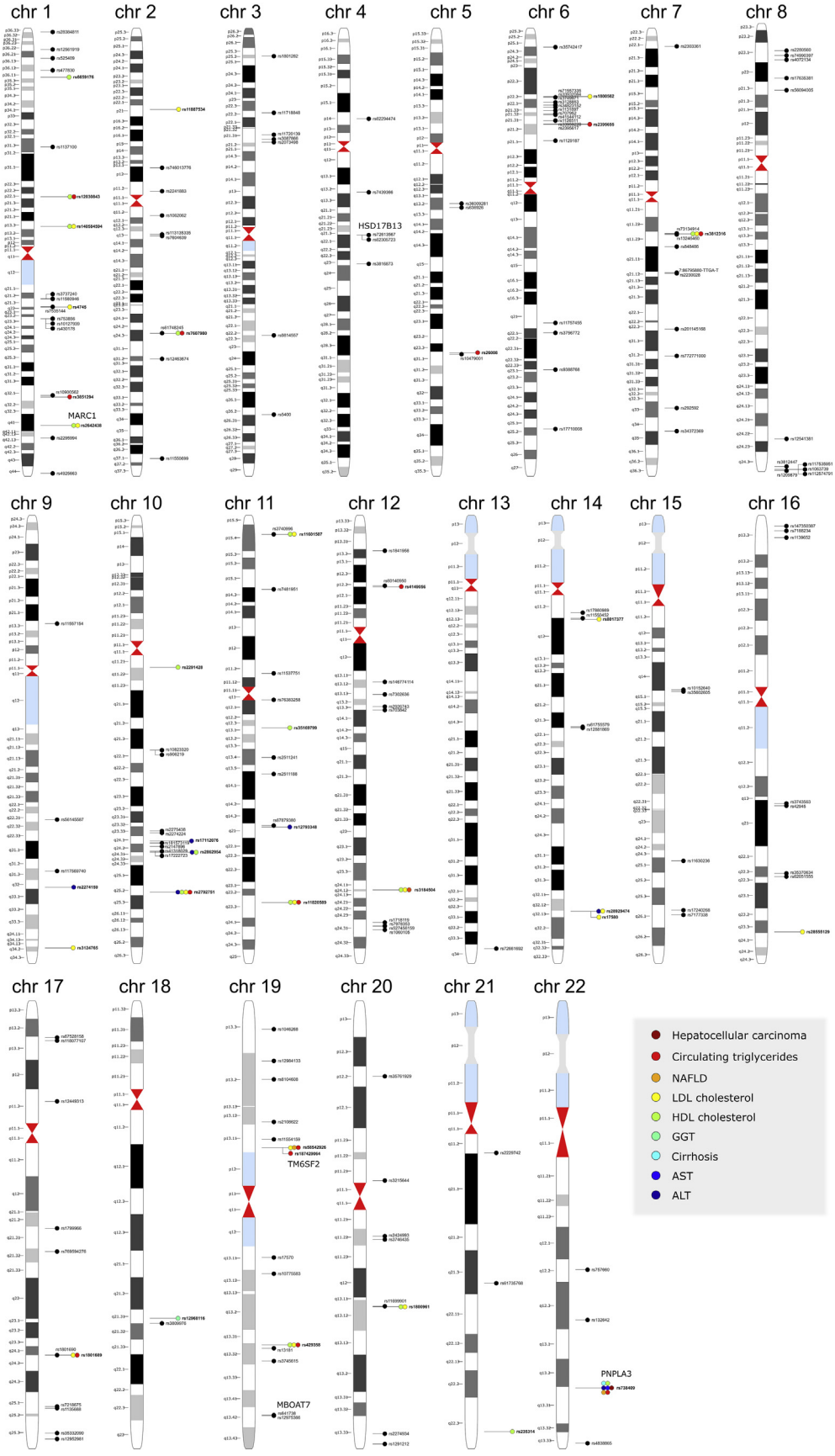


Figure 2. *Chromosomal ideogram* of 190 statistically independent genetic variants detected in exome-wide association study. Each of the 190 independent genetic variants identified by the exome-wide genetic association analyses is shown in its chromosomal location. Genetic variants previously associated in publicly available summary statistics database with liver-related traits are color-coded as displayed in the legend. AST, aspartate aminotransferase; GGT, gamma glutamyltransferase; NAFLD, nonalcoholic fatty liver disease.

(rs1800562/ rs1799945), the association of variants in *PNPLA3*, *TM6SF2*, *MBOAT7*, *MARC1*, *GPAM*, and *APOE* with ALT, PDFF, cirrhosis, and chronic liver disease was virtually identical (data not shown).

Replication of GPAM rs2792751 and APOE rs429358 Association With Liver Damage in Independent Liver Biopsy Cohorts

To replicate the genetic association with liver fat content, we genotyped *GPAM* rs2792751 and *APOE* rs429358 in 3 independent cohorts comprising 2621 individuals at risk for FLD from Europe (466 from Italy, 1331 from France, and 824 from Finland) in whom liver biopsy was available. In these cohorts, we examined the association with severity of liver disease by using an ordinal logistic regression analysis after adjustment for age, sex, BMI, and *PNPLA3* rs738409. The *GPAM* rs2792751 was associated with higher severity of liver steatosis in the Italian ($P = .017$) and French ($P = .006$) cohorts, but not in the Finnish cohort (Figure 4*A*). *APOE* rs429358 was associated with a lower severity of liver steatosis in the Finnish cohort ($P = .015$), was borderline for the French cohort ($P = .052$) and was not significant for the Italian cohort (Figure 4*B*). No other consistent associations with liver inflammation, ballooning, and fibrosis were present for both genetic variants (Supplementary Table 8). Finally, we performed a meta-analysis using both fixed- and random-effects models, and showed that the association with severity of liver steatosis was present only for *GPAM* rs2792751 ($P = .002$ and $P = .041$ for fixed- and random-effects models, respectively), while *APOE* rs429358 conferred protection ($P = .002$ for both fixed- and random-effects models) (Figure 4).

Genetic Risk Score in the UK Biobank and Measurement of Liver Fat Content

Next, we generated an unweighted genetic risk score and examined liver fat content in the UK Biobank using: all the 6 top variants (GRS1); the top 6 variants except for the variants in *APOE* and *GPAM* (GRS2); and only the *GPAM* and *APOE* variants (GRS3) (Figure 5 and Supplementary Table 9). GRS1 had the strongest genetic association with liver fat content ($P = 5.72 \times 10^{-84}$) explaining almost 28% of this trait variation. As expected, the strength of the association and the variation explained was lowest when using only the *APOE* and *GPAM* variants (GRS3).

Liver Transcriptomic Analyses and Functional Prediction of GPAM rs2792751 and APOE rs429358

To gain insight into the mechanisms underlying the genetic association of *GPAM* and *APOE* with liver disease, we examined the liver transcriptome from an independent cohort of 125 obese individuals from Northern Italy from whom liver biopsy was collected during bariatric surgery[46] (Figure 6 and Supplementary Figure 3). Gene enrichment analyses, comparing *GPAM* rs2792751 minor allele (T) carriers with noncarriers, showed an up-regulation of lipid metabolism, namely oxidative phosphorylation, fatty acid metabolism, and adipogenesis, and a down-regulation of the inflammatory response (Figure 6*A*). The corresponding analyses for *APOE* rs429358 showed an up-regulation of the inflammatory response and a reduction in oxidative phosphorylation, fatty acid metabolism, and adipogenesis (Figure 6*B*).

Next, we examined the hepatic messenger RNA expression levels of *GPAM* and *APOE* stratified by rs2792751 and rs429358, respectively. No differences in RNA expression of these genes were found among genotypes (Supplementary Figure 3). To understand the consequences of the amino acid substitution at the protein level, we examined a total of 12 prediction tools. In silico predictions showed that the amino acid substitutions in *GPAM* and *APOE* were benign (Supplementary Table 10).

Discussion

In this study, we identified 2 novel genetic variants in *GPAM* and *APOE* that were robustly associated with liver damage and steatosis, both in individuals from the general population-based UK Biobank cohort and in those at risk for liver disease.

ALT levels are commonly used in clinical practice as a marker of liver damage and the liver plays a pivotal role in lipid and lipoprotein metabolism. We started our analyses by examining the association between ALT and common nonsense and missense variants in participants of the UK Biobank cohort. We found 190 genetic variants that independently associated with ALT at the exome-wide level. This number is considerably larger than found in previous studies that examined variation in the entire genome and used smaller sample sizes,[51] highlighting the advantage of using an exome-wide approach in large cohorts.

In publicly available databases, only 1 of 5 variants identified was associated with liver-related traits in GWAS summary statistics results. Among those already known, most variants were associated with lipid metabolism and lipoproteins as, for example, the loss of function sequence variant in *TM6SF2* inducing hepatic lipid retention by interfering with lipoprotein secretion.[4,11] In line with this, rare variants in the lipoprotein secretion apparatus also are known to cause FLD progression.[14] This demonstrates that there is a tight relationship between liver damage and lipoprotein metabolism.

Furthermore, gene-set enrichment analysis showed that the novel set of genes identified was mostly enriched in lipid biologic processes. By using a Mendelian randomization approach with variants associated with FLD, we showed that quantitative/qualitative alterations in liver fat content per se are deleterious to the liver.[10,15] Results of the present work strongly reinforce the concept that biology of lipids is tightly related to liver damage. In our exome-wide analyses, genomic control was highly inflated, indicating potential confounding bias. However, LD score regression analyses showed that inflation was mostly due to polygenicity of the trait, rather than population stratification or relatedness.

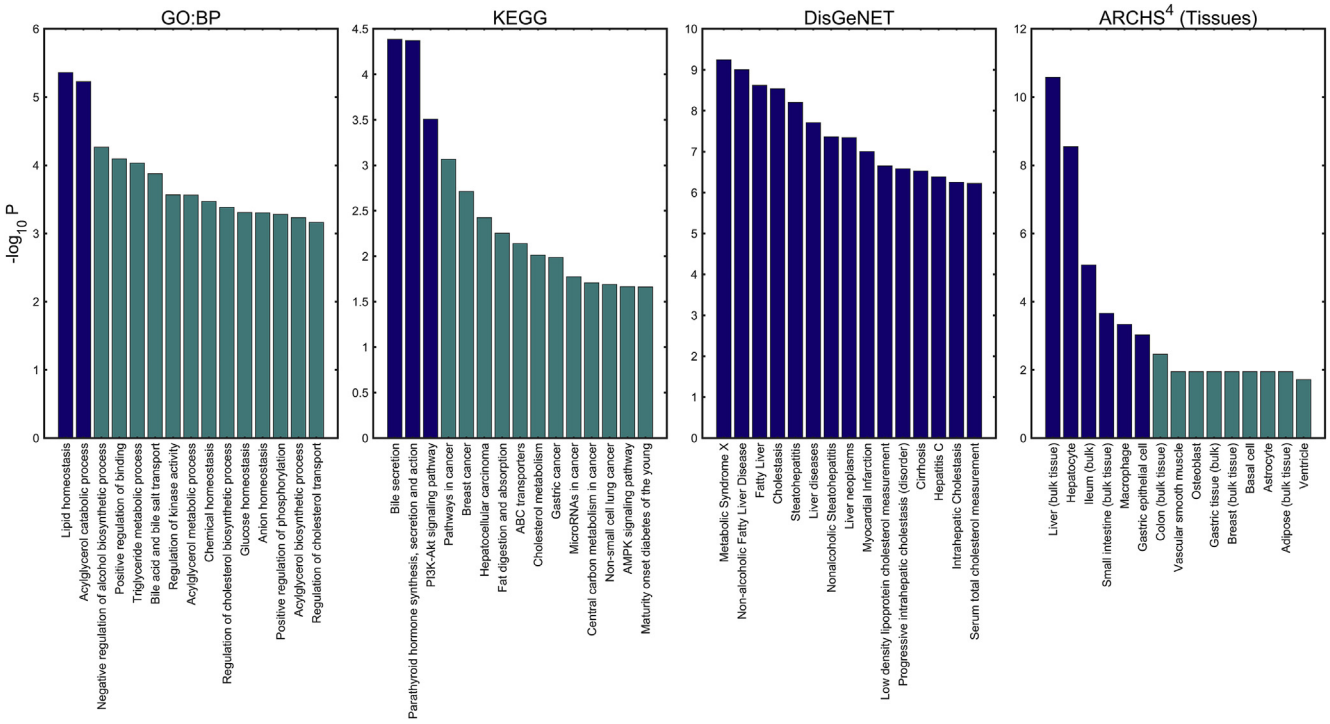


Figure 3. Gene-set enrichment analysis of the genes identified by the exome-wide association study. The set of genes within exome-wide independently significant signals was used to perform gene ontology (biologic processes), Kyoto Encyclopedias of Genes and Genomes (biologic pathways), disease (DisGeNET), and tissue (ARCHS4) enrichment analysis using Enrichr tool.[37–42] Only the first top 15 terms for each analysis are shown here. *Navy blue* represents terms statistically significant after multiple comparison. *Cyan* represents nominal significant *P* values.

To pinpoint the variants with the strongest genetic impact to liver fat content, we examined the 190 genetic variants in 8930 participants of the UK Biobank for whom liver fat measurement was available, and in the overall cohort against chronic liver disease and cirrhosis. Genetic variants in *GPAM* and *APOE* were consistently associated with increased liver fat content and remarkably with the risk of cirrhosis and chronic liver disease. Next, we examined the PAF of the top 6 variants in *PNPLA3, TM6SF2, MBOAT7, MARC1, GPAM*, and *APOE* associated with PDFF for chronic liver disease and cirrhosis. The total PAF was >20% for chronic liver disease and >30% for cirrhosis. As expected the largest PAF was conferred by the *PNPLA3* variant. However, the second largest PAF was conferred by *APOE* and *MARC1*. The reason for this is that these 2 variants are protective with the major allele representing the risk allele. Therefore, because PAF is a function of the risk conferred (odds ratio) and the allele frequency of the risk allele, despite the odds ratio being lower than other variants (eg, TM6SF2), the resulting PAF is higher because of the high frequency of the allele conferring the risk.

We confirmed the association of *GPAM* rs2792751 and *APOE* rs429358 with steatosis severity in a meta-analysis of 3 independent European cohorts of individuals at risk for liver disease in which liver biopsy was available. The association with liver steatosis severity was not statistically significant in all the individual replication cohorts for these 2 loci. However, the direction of the association, namely at risk for *GPAM* and protective for the *APOE* was always conserved, except for the Finnish cohort for *GPAM*, in which it was neutral. This suggests that these results may be due to a lack of power of some individual replication cohorts to detect the association. We did not observe any association between *GPAM* rs2792751 and *APOE* rs429358 with severity of liver inflammation, ballooning, and fibrosis, which again may be due to a lack of power of the study to detect these associations. Further genetic studies in large cohorts with liver biopsy available are needed to assess the effect of these variants on liver inflammation, ballooning, and fibrosis.

Finally, we generated an unweighted genetic risk score and examined liver fat content in the UK Biobank using all (GRS1) or a subset (GRS2 and GRS3) of the 6 top variants. GRS1 had the strongest genetic association with liver fat content, explaining almost 28% of this trait variation. When using only *APOE* and *GPAM* (GRS3), the strength of the association and the variation explained was lower. Of note, in GRS3, there was an additive effect on liver fat content of each of the single alleles of *GPAM* and *APOE*, suggesting that they affect liver fat content through independent pathways.

A recent GWAS examining individuals at risk for FLD identified a novel genetic variant, the rs62021874, in *PYGO1* associated with this disease.[7] We did not find any association between this variant and ALT, PDFF, chronic liver disease, and cirrhosis. This could be due to the absence in the UK Biobank of an interaction between the genetic variant and adiposity or other risk factors for FLD that may be required to uncover the association.

Table 1. Characteristics of the UK Biobank Individuals Stratified by *GPAM* rs2792751 and *APOE* rs429358

	GPAM rs2792751					*APOE* rs429358				
Trait	CC	CT	TT	β/OR	*P*-value	TT	TC	CC	β/OR	*P*-value
N	229,741	172,362	32,498			309,463	114,662	10,476		
Age, *y*	57±8	57±8	57±8	-0.002	0.366	57±8	57±8	57±8	-0.014	9.62E-07
Male gender, n (%)	105,541 (46)	79,022 (46)	14,824 (46)	0.997	0.500	141,932 (46)	52,599 (46)	4,856 (46)	1.009	0.157
BMI, kg/m^2	27.4±4.8	27.4±4.7	27.3±4.7	-0.009	2.1E-04	27.4±4.8	27.4±4.8	27.1±4.7	-0.024	1.29E-16
ALT, *U/L*[a]	20.1 (11.8)	20.2 (12.1)	20.4 (12.5)	0.028	2.9E-36	20.3 (12.1)	19.9 (11.5)	19.8 (11.3)	-0.033	2.6E-34
PDFF, %	2.3 (3.1)	2.4 (3.1)	2.6 (4.3)	0.083	7.6E-9	2.5 (3.4)	2.2 (2.6)	2.2 (2.3)	-0.120	4.8E-11
CLD, n (%)	3,166 (1.38)	2,566 (1.49)	482 (1.48)	1.068	1.3E-3	4534 (1.47)	1,563 (1.36)	117 (1.12)	0.928	2.97E-3
Cirrhosis, n (%)	941 (0.41)	761 (0.44)	144 (0.44)	1.076	0.051	1373 (0.44)	434 (0.38)	39 (0.37)	0.893	0.014
Glucose, *mmol/L*	5.12±1.21	5.12±1.21	5.12±1.21	-2.59E-06	0.999	5.12±1.22	5.11±1.19	5.1±1.12	0.001	0.703
Cholesterol, *mmol/L*	5.69±1.14	5.73±1.15	5.75±1.15	0.024	6.1E-23	5.65±1.12	5.84±1.17	5.99±1.2	0.160	<1E-308
HDL cholesterol, *mmol/L*	1.44±0.38	1.46±0.385	1.47±0.385	0.034	2.1E-54	1.46±0.383	1.43±0.379	1.41±0.376	-0.084	3.66E-217
LDL cholesterol, *mmol/L*	3.56±0.868	3.58±0.872	3.59±0.871	0.018	5.8E-14	3.51±0.855	3.69±0.889	3.81±0.915	0.190	<1E-308
Triglycerides, *mmol/L*	1.5 (1.11)	1.49 (1.1)	1.47 (1.09)	-0.015	4.7E-11	1.48 (1.09)	1.52 (1.14)	1.56 (1.21)	0.068	1.08E-133

NOTE. Continuous traits are shown as mean and standard deviation or median (interquartile range) as appropriate. Gender is shown as number and proportion. *P*-values were calculated by using linear/logistic mixed-effects model (for ALT, PDFF, CLD, and cirrhosis), or linear or logistic (for sex) regression analysis adjusted for age, sex, BMI, the first 10 principal components of ancestry, and genotyping array (except for age, sex, and BMI, for which the trait under analysis was excluded from covariates). Continuous traits were rank-based inverse normal transformed before regression analyses.
ALT, alanine aminotransferase; BMI, body mass index; CLD, chronic liver disease; HDL, high-density lipoprotein; LDL, low-density lipoprotein; OR, odds ratio; PDFF, proton density fat fraction.
[a]in a subset of individuals without available PDFF.

The rs2792751 encodes for a missense (p.Val43Ile) variation in *GPAM.* The minor allele of this variant (43Ile) was associated with higher liver fat content. *GPAM* encodes an enzyme that catalyzes the initial step of triglyceride synthesis by esterification of a fatty acid onto the glycerol back bone. *GPAM* is highly expressed in the liver and adipose tissue and it is nutritionally up-regulated by insulin signaling.[52,53] *Gpam* knockout mice have lower hepatic triglyceride content,[54] whereas liver overexpression of *Gpam* in rodents results in higher hepatic triglyceride content.[54,55] Transcriptomic analyses in individuals stratified by the *GPAM* rs2792751 minor allele in our study were consistent with up-regulation of lipid metabolism in the liver and no differences in *GPAM* expression were observed across genotypes. According to this interpretation, the up-regulation of triglyceride catabolism pathways

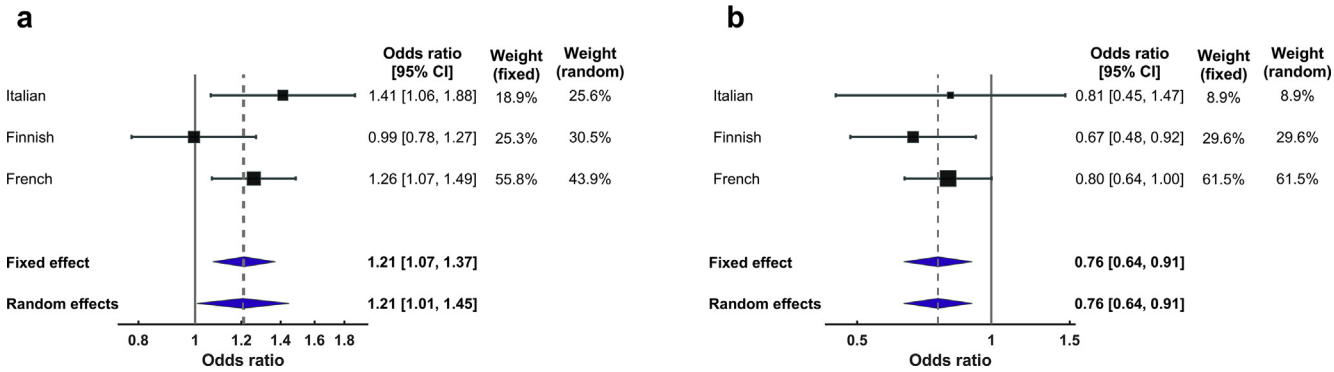


Figure 4. Forest plot of association and meta-analysis for (*A*) *GPAM* rs2792751 and (*B*) *APOE* rs429358 with severity of steatosis in 3 replication cohorts: French, Italian, and Finnish. The association was tested by an ordinal logistic regression analysis under an additive genetic model adjusted by age, sex, BMI, recruitment center (only for the Finnish cohort), and number of *PNPLA3* I148M mutant alleles. Pooled effect estimates were calculated using inverse-variance–weighted fixed- and random-effects meta-analysis.

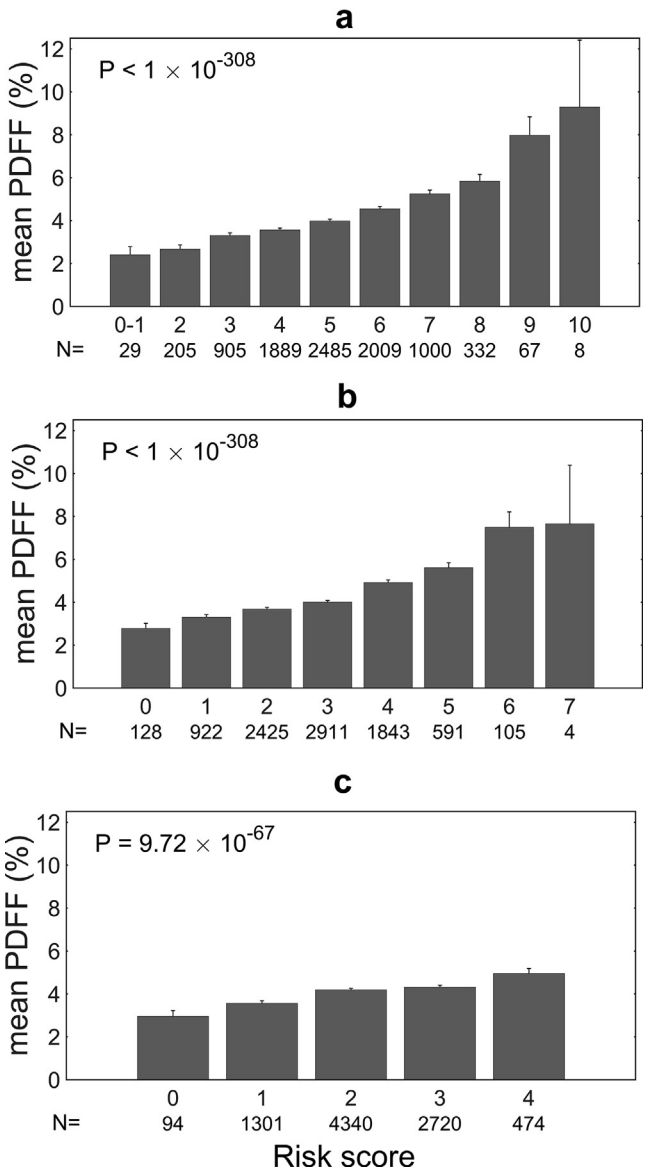


Figure 5. The relationship between PDFF and genetic risk scores. (*A*) GRS1: *PNPLA3* rs738409, *TM6SF2* rs58542926, *MBOAT7* rs641738, *MARC1* rs2642438, *GPAM* rs2792751 and *APOE* rs429358; (*B*) GRS2: *PNPLA3* rs738409, *TM6SF2* rs58542926, *MBOAT7* rs641738 and *MARC1* rs2642438; (*C*) GRS3: *GPAM* rs2792751 and *APOE* rs429358. Each genetic risk score was calculated as the sum of risk alleles. Data are shown as mean PDFF and SEM for the risk scores indicated. *P* values were calculated by a linear regression model adjusted for adjusted for age, sex, BMI, the first 10 principal components of ancestry, and genotyping array.

(peroxisome, fatty acid metabolism, and oxidative phosphorylation) observed in the transcriptomic analysis may possibly represent a protective feedback mechanism to counteract excess lipid accumulation. Moreover, in silico analyses were in line with the amino acid substitution being benign. We therefore speculate that the amino acid substitution induced by rs2792751 leads to an increase in the hepatic triglyceride content with a compensatory increase in the triglyceride utilization to counteract the excess in lipid accumulation. Alternatively, the genetic variant may be in linkage with other noncoding variants that may affect expression levels of neighboring genes.

The rs429358 encodes for a missense p.Cys112Arg in *APOE*, which defines the *APOE4* allele. Our results show that this variant is associated with lower liver fat content and it protects against FLD. Transcriptomic analyses found a down-regulation of the triglyceride metabolism in the liver that may be adaptive to the reduction of hepatic triglycerides content in carriers of the variant. Carriers of the minor allele also had higher circulating triglycerides and low-density lipoprotein cholesterol. *APOE* has a major role in lipid fluxes between tissues during fasting and refeeding.[56] It promotes the clearance of circulating triglycerides and the transfer of lipids into the muscle and adipose tissue after the action of lipoprotein lipase.[57] It could be speculated that *APOE* rs429358 impedes the clearance of circulating lipoproteins and possibly the reuptake of lipids in the liver, or by influencing the efflux of cholesterol in hepatocytes.[57] In silico analyses indicated that this amino acid substitution was benign and there were no differences in *APOE* expression levels between genotypes. Of note, this genetic variant is well known to be associated with higher risk of Alzheimer's disease[57] and dyslipidemia.[58] A link between liver and neurologic diseases is unknown. However, we previously identified *MBOAT7* as a locus involved in FLD.[6] Rare nonsense mutations in homozygosity of this gene result in severe neurologic development delay.[59–61] Functional studies are needed to understand the mechanisms underlying the link between liver and brain disease.

The hepatic transcriptome analyses showed a dissociation between hepatic lipogenesis and inflammation/fibrosis across both *GPAM* and *APOE* genotypes. Moreover, in our replication cohort, we did not observe any association between these variants and histologic evidence of liver inflammation/fibrosis. This may indicate the presence of pleiotropy of these genetic variants on inflammatory/fibrosis pathways that can confound the association with the liver inflammation and fibrosis. Alternatively, the lack of association between the genetic variants and inflammation/fibrosis may be due to a low statistical power of these cohorts to detect an association with these traits. Consistently with this hypothesis, the association with inflammation and fibrosis was directionally consistent with the effect on hepatic steatosis. Further human genetic studies are warranted to confirm that genetic variants in *GPAM* and *APOE* are associated with liver inflammation and fibrosis.

In conclusion, we identified 2 novel genetic variants in *GPAM* and *APOE* associated with liver fat content. Functional studies are needed to understand the mechanisms underlying the genetic associations.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at https://doi.org/10.1053/j.gastro.2020.12.023.

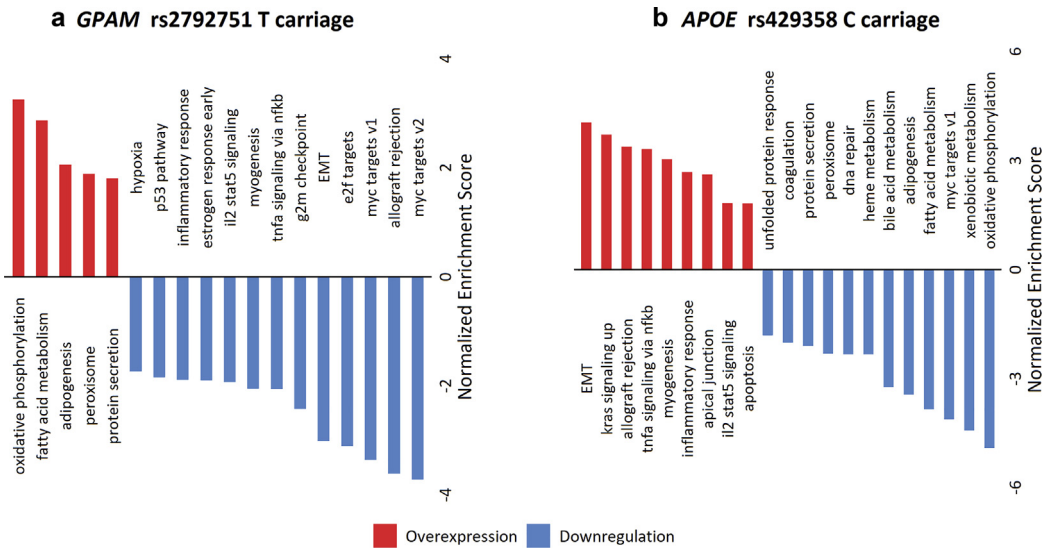


Figure 6. Liver transcriptomic enrichment analysis of 125 obese individuals from Milan, Italy. Pathway enrichment analysis in carriers vs noncarriers for (*A*) *GPAM* rs2792751 and (*B*) *APOE* rs429358.

References

1. Estes C, Anstee QM, Arias-Loste MT, et al. Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016-2030. J Hepatol 2018;69:896–904.
2. Trepo E, Valenti L. Update on NAFLD genetics: from new variants to the clinic. J Hepatol 2020;72:1196–1209.
3. **Romeo S, Kozlitina J**, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40:1461–1465.
4. 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;46:352–356.
5. **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.
6. **Mancina RM, Dongiovanni P**, Petta S, et al. The MBOAT7-TMC4 variant rs641738 Increases risk of nonalcoholic fatty liver disease in individuals of European descent. Gastroenterology 2016;150:1219–1230.e6.
7. Anstee QM, Darlay R, Cockell S, et al. Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically characterised cohort. J Hepatol 2020;73:505–515.
8. Abul-Husn NS, Cheng X, Li AH, et al. A protein-truncating HSD17B13 variant and protection from chronic liver disease. N Engl J Med 2018;378:1096–1106.
9. Prill CA, Haas ME, Khera AV, et al. A missense variant in Mitochondrial Amidoxime Reducing Component 1 gene and protection against liver disease. PLoS Genet 2020;16:e1008629.
10. Romeo S, Sanyal A, Valenti L. Leveraging human genetics to identify potential new treatments for fatty liver disease. Cell Metab 2020;31:35–45.
11. Prill S, Caddeo A, Baselli G, et al. The TM6SF2 E167K genetic variant induces lipid biosynthesis and reduces apolipoprotein B secretion in human hepatic 3D spheroids. Sci Rep 2019;9:11585.
12. Meroni M, Dongiovanni P, Longo M, et al. Mboat7 down-regulation by hyper-insulinemia induces fat accumulation in hepatocytes. EBioMedicine 2020;52:102658.
13. **Tanaka Y, Shimanaka Y, Caddeo A**, et al. LPIAT1/MBOAT7 depletion increases triglyceride synthesis fueled by high phosphatidylinositol turnover. Gut 2021;70:180–193.
14. **Pelusi S, Baselli G**, Pietrelli A, et al. Rare pathogenic variants predispose to hepatocellular carcinoma in nonalcoholic fatty liver disease. Sci Rep 2019;9:3682.
15. **Dongiovanni P, Stender S, Pietrelli A**, et al. Causal relationship of hepatic fat with liver damage and insulin resistance in nonalcoholic fatty liver. J Intern Med 2018;283:356–370.
16. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLOS Medicine 2015;12:e1001779.
17. **Canela-Xandri O, Rawlik K, Tenesa A**. An atlas of genetic associations in UK Biobank. Nat Genet 2018;50:1593–1599.
18. **Tachmazidou I, Hatzikotoulas K, Southam L**, et al. Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. Nat Genet 2019;51:230–236.
19. Linge J, Borga M, West J, et al. Body composition profiling in the UK Biobank Imaging Study. Obesity 2018;26:1785–1795.
20. Linge J, Whitcher B, Borga M, et al. Sub-phenotyping metabolic disorders using body composition: an individualized, nonparametric approach utilizing large data sets. Obesity 2019;27:1190–1199.

21. **Bycroft C, Freeman C, Petkova D**, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018;562:203–209.
22. Cingolani P, Platts A, Wang IL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin) 2012;6:80–92.
23. McLaren W, Gil L, Hunt SE, et al. The Ensembl variant effect predictor. Genome Biol 2016;17:122.
24. 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.
25. Simonen M, Männistö V, Leppänen J, et al. Desmosterol in human nonalcoholic steatohepatitis. Hepatology 2013;58:976–982.
26. **Luukkonen PK, Zhou Y**, Sädevirta S, et al. Hepatic ceramides dissociate steatosis and insulin resistance in patients with non-alcoholic fatty liver disease. J Hepatol 2016;64:1167–1175.
27. Loh PR, Tucker G, Bulik-Sullivan BK, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat Genet 2015;47:284–290.
28. Loh PR, Kichaev G, Gazal S, et al. Mixed-model association for biobank-scale datasets. Nat Genet 2018;50:906–908.
29. Zhou W, Nielsen JB, Fritsche LG, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. Nat Genet 2018;50:1335–1341.
30. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 2015;4:7.
31. Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 2012;44:369–375; S1–S3.
32. **Buniello A, MacArthur JAL**, Cerezo M, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 2019;47:D1005–D1012.
33. Kamat MA, Blackshaw JA, Young R, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 2019;35:4851–4853.
34. Staley JR, Blackshaw J, Kamat MA, et al. PhenoScanner: a database of human genotype-phenotype associations. Bioinformatics 2016;32:3207–3209.
35. **Hemani G, Zheng J, Elsworth B**, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018;7:e34408.
36. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–1321.
37. Chen EY, Tan CM, Kou Y, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 2013;14:128.
38. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90–W97.
39. Lachmann A, Torre D, Keenan AB, et al. Massive mining of publicly available RNA-seq data from human and mouse. Nat Commun 2018;9:1366.
40. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580–585.
41. Piñero J, Bravo À, Queralt-Rosinach N, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. Nucleic Acids Res 2017;45:D833–D839.
42. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. Protein Sci 2019;28:1947–1951.
43. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015;47:291–295.
44. Gazal S, Finucane HK, Furlotte NA, et al. Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. Nat Genet 2017;49:1421–1427.
45. **Finucane HK, Bulik-Sullivan B**, Gusev A, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet 2015;47:1228–1235.
46. **Baselli GA, Dongiovanni P**, Rametta R, et al. Liver transcriptomics highlights interleukin-32 as novel NAFLD-related cytokine and candidate biomarker. Gut 2020;9:1855–1866.
47. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 2011;12:323.
48. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
49. **Subramanian A, Tamayo P**, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–15550.
50. Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1:417–425.
51. Yuan X, Waterworth D, Perry JR, et al. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. Am J Hum Genet 2008;83:520–528.
52. Cha JY, Repa JJ. The liver X receptor (LXR) and hepatic lipogenesis. The carbohydrate-response element-binding protein is a target gene of LXR. J Biol Chem 2007;282:743–751.
53. Yoshikawa T, Shimano H, Amemiya-Kudo M, et al. Identification of liver X receptor-retinoid X receptor as an activator of the sterol regulatory element-binding protein 1c gene promoter. Mol Cell Biol 2001;21:2991–3000.
54. Neschen S, Morino K, Hammond LE, et al. Prevention of hepatic steatosis and hepatic insulin resistance in mitochondrial acyl-CoA:glycerol-sn-3-phosphate

acyltransferase 1 knockout mice. Cell Metab 2005; 2:55–65.
55. Linden D, William-Olsson L, Ahnmark A, et al. Liver-directed overexpression of mitochondrial glycerol-3-phosphate acyltransferase results in hepatic steatosis, increased triacylglycerol secretion and reduced fatty acid oxidation. FASEB J 2006; 20:434–443.
56. Blum CB. Dynamics of apolipoprotein E metabolism in humans. J Lipid Res 1982;23:1308–1316.
57. Yassine HN, Finch CE. APOE alleles and diet in brain aging and Alzheimer's disease. Front Aging Neurosci 2020;12:150.
58. Marais AD. Apolipoprotein E in lipoprotein metabolism, health and cardiovascular disease. Pathology 2019; 51:165–176.
59. Johansen A, Rosti RO, Musaev D, et al. Mutations in MBOAT7, encoding lysophosphatidylinositol acyltransferase I, lead to intellectual disability accompanied by epilepsy and autistic features. Am J Hum Genet 2016; 99:912–916.
60. **Heidari E**, Caddeo A, **Zarabadi K**, et al. Identification of novel loss of function variants in MBOAT7 resulting in intellectual disability. Genomics 2020;112:4072–4077.
61. Dursun A, Yalnızoğlu D, Özgül RK, et al. Clinical highlights of a very rare phospolipid remodeling disease due to MBOAT7 gene defect. Am J Med Genet B Neuropsychiatr Genet 2020;183:3–4.

Author names in bold designate shared co-first authorship.

Received September 23, 2020. Accepted December 11, 2020.

Correspondence
Address correspondence to: Stefano Romeo, MD, PhD, Department of Molecular and Clinical Medicine, University of Gothenburg, Bruna Stråket 16, Gothenburg, 41345 Gothenburg, Sweden. e-mail: stefano.romeo@wlab.gu.se; or Luca Valenti, MD, Department of Pathophysiology and Transplantation, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico Milano, Università degli Studi di Milano, Via Francesco Sforza, 35, Milano, Italy 20122. e-mail: luca.valenti@unimi.it.

Acknowledgments
The authors thank the staff and the participants of the UK Biobank study. This research has been conducted using the UK Biobank resource (application 37142).

CRediT Authorship Contributions
Oveis Jamialahmadi, PhD (Conceptualization: Equal; Data curation: Lead; Formal analysis: Lead; Investigation: Equal; Methodology: Equal; Software: Lead; Validation: Lead; Visualization: Lead; Writing – original draft: Lead; Writing – review & editing: Equal)

Rosellina Margherita Mancina, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Ester Ciociola, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Federica Tavaglione, MD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Panu K. Luukkonen, MD, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Funding acquisition: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Guido Baselli, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Francesco Malvestiti, MSc (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Dorothee Thuillier, MSc (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Violeta Raverdy, MD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Ville Männistö, MD, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Rosaria Maria Pipitone, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Grazia Pennisi, MD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Daniele Prati, MD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Rocco Spagnuolo, MD, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Salvatore Petta, MD, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Validation: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Jussi Pihlajamäki, MD, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Funding acquisition: Equal; Investigation: Equal; Methodology: Equal; Validation: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Francois Pattou, MD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Validation: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Hannele Yki-Järvinen, MD, PhD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Validation: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Luca Vittorio Carlo Valenti, MD (Conceptualization: Lead; Data curation: Lead; Formal analysis: Lead; Funding acquisition: Lead; Investigation: Lead; Methodology: Lead; Project administration: Lead; Resources: Lead; Supervision: Lead; Validation: Lead; Visualization: Lead; Writing – original draft: Lead; Writing – review & editing: Lead)

Stefano Romeo, MD, PhD (Conceptualization: Lead; Data curation: Lead; Formal analysis: Lead; Funding acquisition: Lead; Investigation: Lead; Methodology: Lead; Project administration: Lead; Resources: Lead; Supervision: Lead; Validation: Lead; Visualization: Lead; Writing – original draft: Lead; Writing – review & editing: Lead).

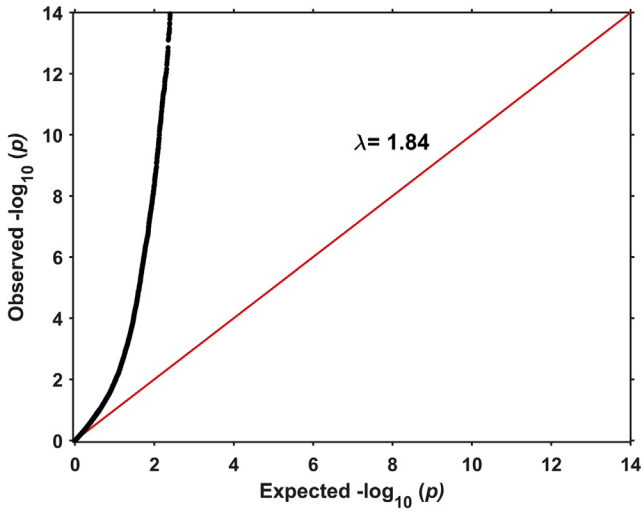
Conflicts of interest
These authors disclose the following: Stefano Romeo[1] has served as a consultant for AstraZeneca, Celgene, Sanofi, Amgen, Akcea Therapeutics, Camp4, Medacorp, and Ambys in the last 3 years. Stefano Romeo[1] has received research grants from AstraZeneca in the last 2 years. Luca Valenti reports having received speaking fees from MSD, Gilead, AlfaSigma, and AbbVie, having served as a consultant for Gilead, Pfizer, Astra Zeneca, Novo Nordisk, and having received research grants from Gilead. Hannele Yki-Järvinen has served as a consultant for Gilead, Novo Nordisk, Hamni, Eli Lilly, and MSD. The remaining authors disclose no conflicts.

Funding
Stefano Romeo was supported by project grants from the Swedish Research Council (Vetenskapsrådet [VR], 2016-01527), the Swedish state under the agreement between the Swedish government and the county councils (the ALF agreement) (SU 2018-04276), the Swedish Diabetes Foundation (DIA 2017-205), the Swedish Heart-Lung Foundation (20120533), the Wallenberg Academy Fellows from the Knut and Alice Wallenberg Foundation (KAW 2017.0203), the Astra Zeneca Agreement for Research, Grant SSF ITM17-0384 Swedish Foundation for Strategic Research, and Novo Nordisk Project Grants in Endocrinology & Metabolism - Nordic Region 2020. Luca Valenti was supported by the following projects grants: the MyFirst Grant AIRC n.16888, Ricerca Finalizzata Ministero della Salute RF-2016-02364358, Ricerca Corrente Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, the European Union Programme Horizon 2020 (under grant agreement no. 777377) for the project LITMUS "Liver Investigation: Testing Marker Utility in Steatohepatitis", Fondazione IRCCS Ca' Granda "Liver BIBLE" PR-0391, and Fondazione IRCCS Ca' Granda core COVID-19 Biobank (RC100017A). Kuopio Obesity Surgery Study (PI JP) was supported by the Finnish Diabetes Research Foundation, Kuopio University Hospital Project grant (EVO/VTR grants 2005-2020), and the Academy of Finland grant (contract No 138006). Panu K. Luukkonen was supported by the Novo Nordisk, Sigrid Jusélius, and Instrumentarium Science Foundations.

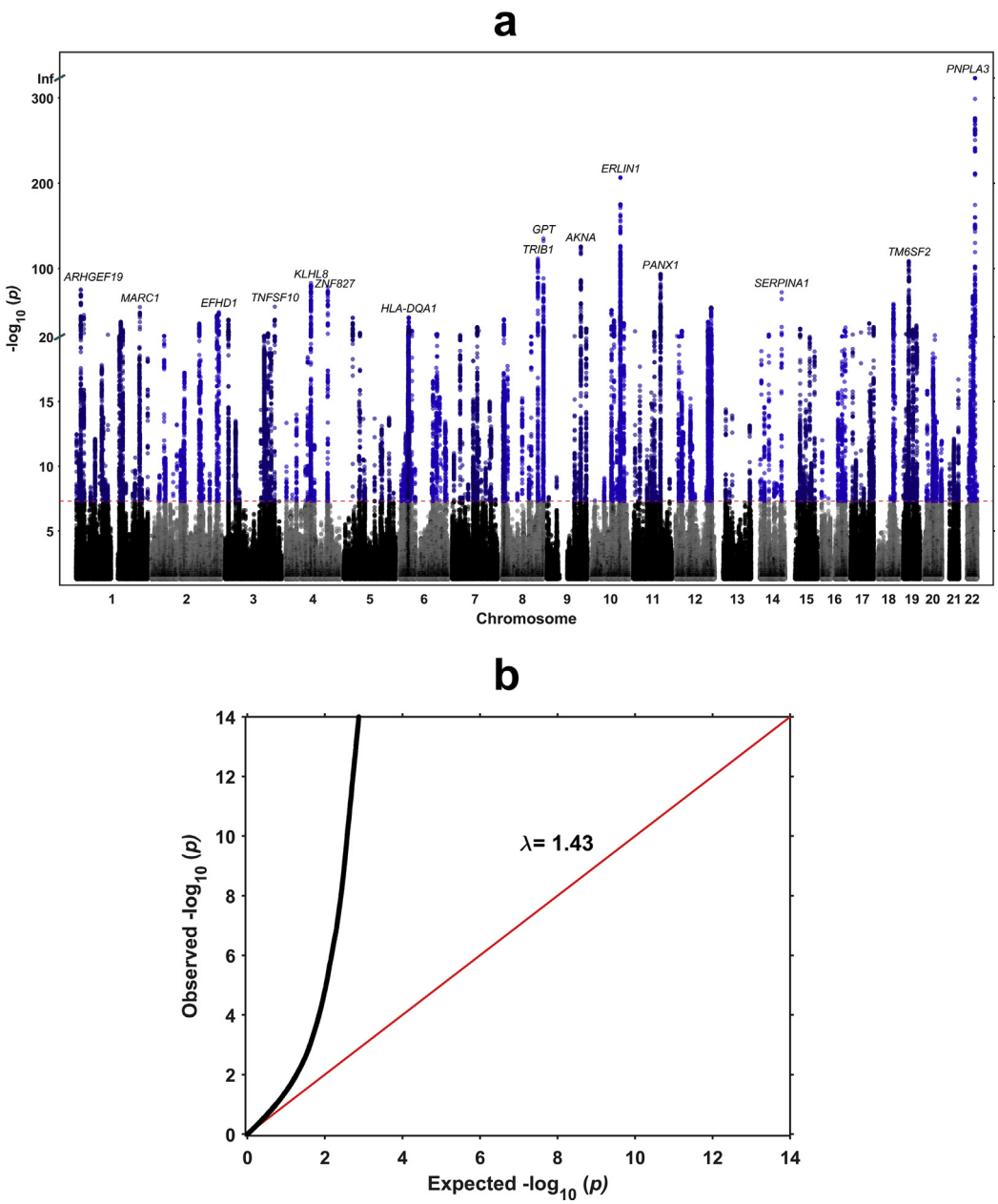
References

1. Naj AC, Jun G, Beecham GW, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet 2011;43:436–441.
2. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res 2001;11:863–874.
3. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–249.
4. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310–315.
5. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics 2015;31:761–763.
6. Shihab HA, Gough J, Cooper DN, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat 2013;34:57–65.
7. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res 2009;19:1553–1561.
8. Dong C, Wei P, Jian X, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum Mol Genet 2015;24:2125–2137.
9. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res 2011;39:e118.
10. Schwarz JM, Rödelsperger C, Schuelke M, et al. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 2010;7:575–576.
11. Choi Y, Sims GE, Murphy S, et al. Predicting the functional effect of amino acid substitutions and indels. PLoS One 2012;7:e46688.
12. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet 2016;99:877–885.

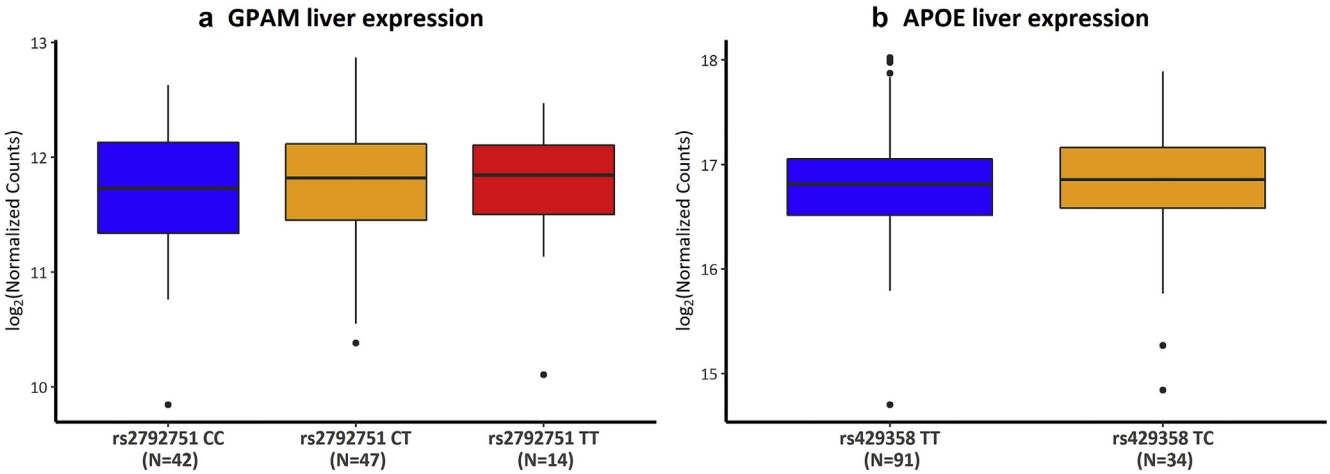
e13. Liu X, Wu C, Li C, et al. dbNSFP v3.0: A one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs. Hum Mutat 2016;37:235–241.



Supplementary Figure 1. QQ plot for exome-wide association study of ALT in European participants from the UK Biobank.



Supplementary Figure 2. a) Manhattan and b) QQ plots for genome-wide association study of ALT in European participants from the UK Biobank.



Supplementary Figure 3. Hepatic mRNA expression levels of a) *GPAM* rs2792751 and b) *APOE* rs429358 stratified by genotype in 125 obese individuals from Milan, Italy.

Supplementary Table 1. International Classification of Diseases, 10th Edition, Codes for Chronic Liver Disease and Cirrhosis in UK Biobank

ICD-10 code	Diagnosis
Inclusion critera	
K70.0	Alcoholic fatty liver
K70.1	Alcoholic hepatitis
K70.2	Alcoholic fibrosis and sclerosis of liver
K70.3	**Alcoholic cirrhosis of liver**
K70.4	**Alcoholic hepatic failure**
K70.9	Alcoholic liver disease, unspecified
K72.1	**Chronic hepatic failure**
K72.9	**Hepatic failure, unspecified**
K73.0	Chronic persistent hepatitis, not elsewhere classified
K73.1	Chronic lobular hepatitis, not elsewhere classified
K73.2	Chronic active hepatitis, not elsewhere classified
K73.8	Other chronic hepatitis, not elsewhere classified
K73.9	Chronic hepatitis, unspecified
K74.0	Hepatic fibrosis
K74.1	**Hepatic sclerosis**
K74.2	**Hepatic fibrosis with hepatic sclerosis**
K74.6	**Other and unspecified cirrhosis of liver**
K76.0	Fatty (change of) liver, not elsewhere classified
K76.6	**Portal hypertension**
K76.7	**Hepatorenal syndrome**
K76.8	Other specified diseases of liver
K76.9	Liver disease, unspecified
I85.0	**Esophageal varices with bleeding**
I85.9	**Esophageal varices without bleeding**
Exclusion criteria	
B18.0	Chronic viral hepatitis B with delta-agent
B18.1	Chronic viral hepatitis B without delta-agent
B18.2	Chronic viral hepatitis C
B18.8	Other chronic viral hepatitis
B18.9	Chronic viral hepatitis, unspecified
B19.0	Unspecified viral hepatitis with coma
B19.9	Unspecified viral hepatitis without coma

NOTE. International Classification of Diseases 10th edition (ICD-10) codes from hospitalization, underlying primary and secondary cause of death (data-fields 41270, 40001 and 40002) were used to define chronic liver disease and cirrhosis. Diagnoses in bold were used for the definition of cirrhosis.

Supplementary Table 2. Characteristics of Participants From UK Biobank Stratified by the Presence of Liver Fat Content Measured by Proton Density Fat Fraction

Trait	Overall	No PDFF	PDFF
n	434,601	425,671	8930
Age, *y*	57 ± 8	57 ± 8	56 ± 8
Sex, *male*, n (%)	199,387 (45.9)	195,106 (45.8)	4281 (47.9)
BMI, kg/m^2	27.4 ± 4.76	27.4 ± 4.8	26.7 ± 4.3
ALT, *U/L*	20.1 (12)	20.2 (11.9)	19.6 (11.6)
PDFF, %	2.4 (3.1)	—	2.4 (3.1)
CLD, n (%)	6,214 (1.4)	6,147 (1.4)	67 (0.8)
Cirrhosis, n (%)	1,846 (0.4)	1,831 (0.4)	15 (0.2)

NOTE. Continuous traits are shown as mean ± SD. Second and third columns represent individuals without and with available PDFF, respectively.
CLD, chronic liver disease; PDFF, proton density fat fraction.

Supplementary Table 7. Population Attributable Fraction Estimates of Top Genetic Variants Associated With Chronic Liver Disease and Cirrhosis

SNP	Gene	OR	MAF	PAF, %	Trait
rs738409	*PNPLA3*	1.309	0.216	6.26	CLD
rs58542926	*TM6SF2*	1.319	0.075	2.34	—
rs641738	*MBOAT7*	1.060	0.440	2.60	—
rs2642438	*MARC1*	0.926	0.297	5.35	—
rs2792751	*GPAM*	1.068	0.273	1.82	—
rs429358	*APOE*	0.928	0.156	6.16	—
Combined PAF	—	—	—	22.26	—
rs738409	*PNPLA3*	1.563	0.216	10.83	Cirrhosis
rs58542926	*TM6SF2*	1.496	0.075	3.59	—
rs641738	*MBOAT7*	1.095	0.440	4.01	—
rs2642438	*MARC1*	0.914	0.297	6.23	—
rs2792751	*GPAM*	1.076	0.273	2.02	—
rs429358	*APOE*	0.893	0.156	9.19	—
Combined PAF	—	—	—	31.15	—

NOTE. PAF was calculated by taking into account the frequency of risk alleles (calculated in all Europeans in UK Biobank) and ORs (from logistic mixed-effects model) using the following equation[e1]: $PAF = \frac{AF\,(OR - 1)}{1 + AF\,(OR - 1)}$ where AF is the frequency of risk allele (major allele for *MARC1* rs2642438 and *APOE* rs429358, and minor allele for the remaining variants). Combined PAF was also calculated as: $PAF_{combined} = 1 - \prod_{i=1}^{all\ SNPs} (1 - PAF_i)$
CLD, chronic liver disease; OR, odds ratio; PAF, population attributable fraction.

Supplementary Table 8. Association and Meta-Analysis Between *GPAM* rs2792751 and *APOE* rs429358 With Severity Of Inflammation, Ballooning, and Fibrosis in 3 Replication Cohorts: French, Italian, and Finnish

	GPAM rs2792751			*APOE* rs429358			
Cohort	OR	95% CI	*P* value	OR	95% CI	*P* value	Trait
Italian	0.96	0.71–1.29	.786	1.21	0.66–2.23	.540	Inflammation
Finnish	1.02	0.77–1.35	.901	0.75	0.52–1.10	.140	—
French	1.31	1.08–1.58	.005	0.93	0.71–1.21	.596	—
Fixed-effect model	1.15	1.00–1.32	.046	0.90	0.73–1.11	.315	—
Random-effects model	1.12	0.91–1.37	.298	0.90	0.73–1.11	.315	—
Italian	1.41	1.06–1.87	.019	0.67	0.37–1.19	.172	Ballooning
Finnish	1.05	0.74–1.48	.791	0.93	0.60–1.44	.729	—
French	0.95	0.69–1.30	.764	0.88	0.56–1.34	.575	—
Fixed-effect model	1.15	0.96–1.38	.138	0.88	0.56–1.34	.575	—
Random-effects model	1.14	0.89–1.45	.305	0.84	0.64–1.11	.226	—
Italian	1.15	0.89–1.48	.290	1.40	0.83–2.37	.208	Fibrosis
Finnish	1.00	0.79–1.27	.997	0.86	0.63–1.18	.348	—
French	1.12	0.91–1.37	.265	1.03	0.78–1.35	.811	—
Fixed-effect model	1.09	0.96–1.24	.205	1.01	0.84–1.20	.948	—
Random-effects model	1.09	0.96–1.24	.205	1.01	0.82–1.25	.915	—

NOTE.The association was tested by an ordinal logistic regression analysis under an additive genetic model adjusted by age, sex, BMI, center of recruitment (only for Finnish cohort), and number of *PNPLA3* I148M mutant allele. Pooled effect estimates were calculated using fixed- and random-effects meta-analysis models.
CI, confidence interval; OR, odds ratio.

Supplementary Table 9. The Association Between Genetic Risk Scores and Proton Density Fat Fraction in UK Biobank

Genetic risk score	β	SE	*P* value	*P* value[a]	R^2
GRS-1[b]	.125	.006	9.51E-67	5.72E-84	0.276
GRS-2[c]	.137	.008	1.34E-53	3.03E-69	0.271
GRS-3[d]	.097	.011	2.54E-15	1.64E-17	0.251

NOTE. Unweighted genetic risk scores were calculated by summing the number of risk alleles. For *PNPLA3* rs738409, *TM6SF2* rs58542926, *MBOAT7* rs641738, and *GPAM* rs2792751 minor allele was considered as the risk allele, while it was the major allele for *MARC1* rs2642438, and *APOE* rs429358. *P* values were calculated using a linear regression analysis.
[a]Adjusted for age, sex, BMI, the first 10 principal components of ancestry, and genotyping array.
[b]GRS-1: *PNPLA3* rs738409, *TM6SF2* rs58542926, *MBOAT7* rs641738, *GPAM* rs2792751 *MARC1* rs2642438 and *APOE* rs429358.
[c]GRS-2: *PNPLA3* rs738409, *TM6SF2* rs58542926, *MBOAT7* rs641738, and *MARC1* rs2642438.
[d]GRS-3: *GPAM* rs2792751 and *APOE* rs429358.

Supplementary Table 10. In Silico Prediction of the *GPAM* rs2792751 and *APOE* rs429358 Effects on Protein Function

Tool	Prediction (raw score)		Range
	GPAM rs2792751	*APOE* rs429358	
SIFT[e2]	Tolerated (1.0)	Tolerated (1.0)	—
PolyPhen[e3]	Benign (0)	Benign (0)	—
CADD PHRED[e4]	1.54 (0.031)[a]	16.65 (1.64)[a]	–6.46 to 18.30
DANN_score[e5]	0.14[a]	0.22[a]	0 to 1
FATHMM[e6]	Tolerated (1.78)	Tolerated (–0.24)	—
LRT[e7]	Neutral (0.005)	Neutral (0.15)	—
MetaLR[e8]	Tolerated (0)	Tolerated (0)	—
MetaSVM[e8]	Tolerated (–0.99)	Tolerated (–1.01)	—
MutationAssessor[e9]	Neutral (–1.59)	Neutral (–1.47)	—
MutationTaster[e10]	P (1), harmless	P (1), harmless	0 to 1
PROVEAN[e11]	Neutral (0.15)	Neutral (4.36)	—
REVEL[e12]	0.04[a]	0.23[a]	0 to 1

NOTE. All predictions were extracted from dbNSFP database (version 4.1a).[e13]
[a]The larger the score the more likely the SNP has damaging effect.