rs641738C>T near *MBOAT7* is positively associated with liver fat, ALT, and histological severity of NAFLD: a meta-analysis

Kevin Teo[1], Kushala W. M. Abeysekera[2], Leon Adams[3,4], Elmar Aigner[5], Jesus M. Banales[6], Rajarshi Banerjee[7], Priyadarshi Basu[8], Thomas Berg[9], Pallav Bhatnagar[10], Stephan Buch[11], Ali Canbay[12], Sonia Caprio[13], Ankita Chatterjee[8], Yii-Der Ida Chen[14], Abhijit Chowdhury[15], Christian Datz[16], Dana de Gracia Hahn[1], Johanna K. DiStefano[17], Jiawen Dong[1], Amedine Duret[1], EU-PNAFLD Investigators*, Connor Emdin[18], Madison Fairey[1], Glenn S Gerhard[19], GOLD Consortium*, Xiuqing Guo[14], Jochen Hampe[11], Matthew Hickman[2], Lena Heintz[20], Christian Hudert[21], Harriet Hunter[1], Matt Kelly[7], Julia Kozlitina[22], Marcin Krawczyk[20,23], Frank Lammert[20], Claudia Langenberg[24], Joel Lavine[25], Lin Li[26], Hong Kai Lim[1], Rohit Loomba[27], Panu K. Luukkonen[28,29,30], Phillip E. Melton[31,32], Trevor A. Mori[3], Nicholette D. Palmer[33], Constantinos A. Parisinos[34], Sreekumar G Pillai[10], Faiza Qayyum[35], Matthias C. Reichert[20], Stefano Romeo[36,37.38], Jerome I. Rotter[14], Yu Ri Im[1], Nicola Santoro[13], Clemens Schafmayer[39], Elizabeth K. Speliotes[40,41], Stefan Stender[35], Felix Stickel[42], Christopher D. Still[43], Pavel Strnad[44], Kent D. Taylor[14], Anne Tybjærg-Hansen[35], Giuseppina Rosaria Umano[13,45], Mrudula Utukuri[1], Luca Valenti[46,47], Lynne E. Wagenknecht[48], Nicholas J. Wareham[24], Richard M. Watanabe[49], Julia Wattacheril[50], Hanieh Yaghootkar[51], Hannele Yki-Järvinen[28,29], Kendra A. Young[52], Jake P. Mann[24]

Affiliations

1. Clinical School of Medicine, University of Cambridge, Cambridge, UK
2. MRC Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, United Kingdom

3. Medical School, Faculty of Health and Medical Sciences, University of Western Australia, Perth, Australia
4. Department of Hepatology, Sir Charles Gairdner Hospital, Perth, WA, Australia
5. First Department of Medicine, Paracelsus Medical University Salzburg, Austria
6. Department on Liver and Gastrointestinal Diseases, Biodonostia Health Research Institute, Donostia University Hospital, University of the Basque Country (UPV/EHU), CIBERehd, Ikerbasque, San Sebastian, Spain
7. Perspectum Diagnostics Ltd., Oxford, UK.
8. National Institute of Biomedical Genomics, Kalyani, India
9. Division of Hepatolgy, Clinic and Polyclinic for Gastroenterology, Hepatology, Infectious Diseases, and Pneumology, University Clinic Leipzig
10. Eli Lilly and Company, Indianapolis IN.
11. Medical Department 1, University Hospital Dresden, Technische Universität Dresden (TU Dresden), Dresden, Germany
12. Gastroenterology, Hepatology and Infectiology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany
13. Yale University, Department of Pediatrics, New Haven, CT
14. The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA
15. Institute of Post Graduate Medical Education And Research, Kolkata, India
16. Department of Internal Medicine, General Hospital Oberndorf, Teaching Hospital of the Paracelsus Medical University Salzburg, Oberndorf, Austria
17. Diabetes and Fibrotic Disease Unit Translational Genomics Research Institute (TGen)
18. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Boston, MA
19. Department of Medical Genetics and Molecular Biochemistry, Lewis Katz School of Medicine at Temple University, Philadelphia, PA
20. Department of Medicine II, Saarland University Medical Center, Saarland University, Homburg, Germany
21. Center for Chronically Sick Children, Charité - Universitätsmedizin Berlin, Berlin, Germany
22. Eugene McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, TX, USA
23. Laboratory of Metabolic Liver Diseases, Department of General, Transplant and Liver Surgery, Laboratory of Metabolic Liver Diseases, Centre for Preclinical Research, Medical University of Warsaw, Warsaw, Poland
24. MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, UK.
25. Department of Pediatrics, Columbia University, New York, USA
26. BioStat Solutions Inc, Frederick, Maryland, USA
27. NAFLD Research Center, Division of Gastroenterology and Epidemiology, University of California at San Diego, La Jolla, California, USA
28. Minerva Foundation Institute for Medical Research, Helsinki, Finland.
29. Department of Medicine, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
30. Yale University School of Medicine, New Haven, Connecticut, US.
31. The Curtin UWA Centre for Genetic Origins of Health and Disease, Faculty of Health and Medical Sciences, Curtin University, The University of Western Australia, Perth, WA, Australia
32. School of Pharmacy and Biomedical Sciences, Faculty of Health Sciences, Curtin University, Perth, WA, Australia
33. Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina, US
34. Institute of Health Informatics, Faculty of Population Health Sciences, University College London, London, UK
35. Department of Clinical Biochemistry, Rigshospitalet Copenhagen University Hospital Copenhagen, Denmark
36. Department of Molecular and Clinical Medicine, University of Gothenburg, Sweden
37. Cardiology Department, Sahlgrenska University Hospital, Gothenburg, Sweden

38. Clinical Nutrition Unit, Department of Medical and Surgical Sciences, University Magna Graecia, Catanzaro, Italy
39. Department of Visceral and Thoracic Surgery, Kiel University, Kiel, Germany
40. Division of Gastroenterology and Hepatology, Department of Medicine, University of Michigan Health System Ann Arbor MI.
41. Department of Computational Medicine and Bioinformatics, University of Michigan Medical School Ann Arbor MI.
42. Department of Gastroenterology and Hepatology, University Hospital of Zurich, Switzerland
43. Geisinger Obesity Institute, Danville, PA 17822, United States
44. Medical Clinic III, University Hospital RWTH Aachen, Aachen, Germany
45. Department of the Woman, the Child, of General and Specialized Surgery, University of Campania Luigi Vanvitelli, Naples Italy
46. Department of Pathophysiology and Transplantation, Università degli Studi di Milano
47. Translational Medicine, Department of Transfusion Medicine and Hematology, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano, Milan, Italy
48. Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, US
49. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA
50. Department of Medicine, Center for Liver Disease and Transplantation, Columbia University College of Physicians and Surgeons, New York Presbyterian Hospital, New York, New York 10032, USA
51. Genetics of Complex Traits, College of Medicine and Health, University of Exeter, Exeter, UK
52. Department of Epidemiology, Colorado School of Public Health, University of Colorado Denver, Aurora, Colorado, USA

*A list of contributors is included in the Acknowledgements

Keywords: *MBOAT7*, fibrosis, NAFLD, triglyceride, diabetes, ALSPAC

Correspondence: Dr Jake P. Mann, University of Cambridge, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, United Kingdom. jm2032@cam.ac.uk, Tel: +44 1223 336792, Fax: +44 1223 330598

Financial support: JPM is supported by a Wellcome Trust Fellowship (216329/Z/19/Z), a European Paediatric Research Society award, and a Children's Liver Disease Foundation grant. The EU-PNAFLD is supported by an EASL Registry Grant. NIH grants: R01HD028016 (SC), R01DK111038 (SC), R01DK114504 (NS), DK091601 (JKD), UL1TR001105 (JK). Supported by the Nonalcoholic Steatohepatitis Clinical Research Network (NASH CRN) of the National Institute of Diabetes and Digestive and Kidney Diseases

(NIDDK) (U01DK061718, U01DK061728, U01DK061731, U01DK061732, U01DK061734, U01DK061737, U01DK061738, U01DK061730, and U01DK061713); the National Center for Advancing Translational Sciences (UL1TR000439, UL1TR000436, UL1TR000006, UL1TR000448, UL1TR000100, UL1TR000004, UL1TR000423, UL1TR000058, and UL1TR001881); and the NIDDK (DK063491 to the Southern California Diabetes Endocrinology Research Center). This study was supported by the German Federal Ministry for Education and Research (BmBF) through the Livers Systems Medicine (LiSyM) project. This work was supported by grants from the Swiss National Funds (SNF no. 310030_169196) and the Swiss Foundation for Alcohol Research (SSA) to FS. This Raine Study was supported by the National Health and Medical Research Council of Australia [grant numbers 403981, 353514 and 572613]. The UK Medical Research Council and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. ALSPAC GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf); This research was specifically funded by grants from MRC and Alcohol Research UK (MR/L022206/1) and NIH (5R01AA018333-05) to KWMA & MH. LV was supported by MyFirst Grant AIRC n.16888, Ricerca Finalizzata Ministero della Salute RF-2016-02364358, Ricerca corrente Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, LV and AG. received funding from the European Union (EU) Programme Horizon

2020 (under grant agreement No. 777377) for the project LITMUS-“ Liver Investigation: Testing Marker Utility in Steatohepatitis”. German Federal Ministry of Education and Research (BMBF LiSyM 031L0051 to F.L.) PL is supported by grants from the Sigrid Jusélius Foundation and the Novo Nordisk Foundation. The Fenland study was funded by grants to the MRC Epidemiology Unit (MC UU12015/1, MC UU 12015/5). RB and MK are employees of and shareholders in Perspectum Diagnostics Ltd. CAP is funded by a Wellcome Trust Clinical PhD Programme (206274/Z/17/Z). J.M. Banales is supported by the Spanish Carlos III Health Institute (ISCIII; PI15/01132, PI18/01075, CIBERehd and Miguel Servet Program CON14/00129) cofinanced by “Fondo Europeo de Desarrollo Regional” (FEDER), and La Caixa Scientific Foundation (HR17-00601). Research from the GUARDIAN Study was supported DK085175 and DK118062 from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). HY is funded by a Diabetes UK RD Lawrence fellowship (17/0005594).

ABSTRACT

Background & Aims

A common genetic variant near *MBOAT7* (rs641738C>T) has been previously associated with hepatic fat and advanced histology in non-alcoholic fatty liver disease (NAFLD), however, these findings have not been consistently replicated in the literature. We aimed to establish whether rs641738C>T is a risk factor across the spectrum of NAFLD and characterize its role in the regulation of related metabolic phenotypes through meta-analysis.

Methods

We performed meta-analysis of studies with data on the association between rs641738C>T genotype and: liver fat, NAFLD histology, and serum ALT, lipids, or insulin. These included directly genotyped studies and population-level data from genome-wide association studies (GWAS). We performed random effects meta-analysis using recessive, additive, and dominant genetic models.

Results

Data from 1,047,265 participants (8,303 with liver biopsies) across 42 studies was included in the meta-analysis. rs641738C>T was associated with higher liver fat on CT/MRI (+0.03 standard deviations [95% CI: 0.02 - 0.05]) and diagnosis of NAFLD (OR 1.22 [95% CI 1.08 - 1.39]) in Caucasian adults. The variant was also positively associated with presence of severe steatosis, NASH, and advanced fibrosis (OR: 1.32 [95% CI: 1.06 - 1.63]) in Caucasian adults using a recessive model of inheritance (CC+CT vs. TT). Meta-analysis of data from previous GWAS found the variant to be associated with higher ALT (P_z=0.002) and lower serum triglycerides (P_z=1.5x10^{-4}). rs641738C>T

was not associated with fasting insulin and no effect was observed in children with NAFLD.

Conclusion

Our study validates rs641738C>T near *MBOAT7* as a risk factor for the presence and severity of NAFLD in individuals of European descent.

Abstract word count: 274

Conflict of interest: Connor Emdin reports personal fees from Navitor Pharma and Novartis.

INTRODUCTION

Since the first genome-wide association study (GWAS) of liver fat[1], more than 20 genetic single nucleotide variants (SNVs) have been associated with non-alcoholic fatty liver disease (NAFLD)[2,3]. These studies have deepened our understanding of the condition, its heritability, and its relationship with cardio-metabolic disease.

Rs641738C>T near *MBOAT7* (membrane bound O-acyltransferase domain containing 7) was initially identified as a genome-wide significant risk variant for alcohol-related cirrhosis (odds ratio=1.35, $p=1.03 \times 10^{-9}$)) [4]. It has since been implicated in the pathogenesis of NAFLD[5], hepatocellular carcinoma[6], as well as in fibrosis development in chronic hepatitis B/C[7–9], and primary sclerosing cholangitis[10]. However, unlike variants in *PNPLA3*, *TM6SF2*, and *HSD17B13*, it has not been identified at genome-wide significance for liver fat or ALT[1,11,12].

Rs641738 is located a few hundred base pairs downstream of the 3' untranslated region of *MBOAT7*, which belongs to a family of genes that code for specific acyl donors and acceptors[13]. *MBOAT7* encodes lysophosphatidylinositol acyltransferase 1 (LPIAT1), which contributes to the regulation of free arachidonic acid in cells[14,15]. Rs641738C>T is associated with lower hepatic expression of *MBOAT7* at both the mRNA[16] and protein level[5]. Given its role in inflammatory lipid pathways, most mechanistic work relating to rs641738 has focused on *MBOAT7[17]*.

In NAFLD, the rs641738C>T variant was first demonstrated to be associated with increased hepatic fat content and severity of fibrosis in individuals of European descent[5]. Proton magnetic resonance spectroscopy data from 2736 individuals showed a modest increase in hepatic fat in those with TT-genotype (4.1%) compared to those with CT- (3.6%) or CC-genotype (3.5%, p=.005). Follow-up studies of European subjects corroborated the initial findings, and suggested a role in development of hepatocellular carcinoma[18,19]. However, these results were not replicated in adults of other ancestries[5,20–22] or in children[23].

In addition, biallelic loss of function mutations in *MBOAT7* cause autosomal recessive mental retardation 57 (OMIM #617188) and no liver phenotype has been reported in these patients to date[15,24,25]. However, rare likely pathogenic (coding) variants in *MBOAT7* are associated with HCC in NAFLD[26].

In summary, the association between rs641738C>T and hepatic fat content, as well as its effects on severity of NAFLD, remain unclear. Moreover, the broader metabolic effects of this SNV, including its association with markers of insulin resistance and dyslipidaemia have not been assessed. Understanding the broader metabolic effects of rs641738C>T is important if it were to be investigated as a drug target in NAFLD.

Here, we conducted a large meta-analysis to determine if rs641738C>T influences the development or stage of NAFLD and related traits.

METHODS

Data sources and study selection

Two data sources were included in the meta-analysis: (i) studies which looked at the effect of the variant on traits of interest by genotyping the variant; and (ii) look-up from GWAS of traits of interest.

Published studies were sourced through: Medline, Embase, HuGe Navigator, Web of Science, bioRxiv, and medRxiv. The search terms used were: "(*MBOAT7* or membrane-bound-o-acyltransferase) or (rs641738 or rs626283) or (*TMC4*)". In addition, HuGe Navigator Phenopedia was searched using terms related to liver disease (Supplementary Methods). There were no restrictions on date or language, and the study selection included all original studies including AASLD Liver Meeting and EASL meeting abstracts. The search was completed on 4th March 2020. Reference lists of publications were also reviewed.

A separate search was conducted for all potentially relevant GWAS through: GWAS Catalogue[27], Phenoscanner[28], Type 2 diabetes knowledge portal[29], and Cardiovascular disease knowledge portal[30] (Supplementary Methods).

After removal of duplicates, titles and abstracts were screened for eligibility independently by two authors (investigators), with inclusion/exclusion criteria applied to potentially eligible full texts.

HuGENet guidelines[31] were followed throughout and MOOSE reporting guidelines[32] were used. This study was prospectively registered on PROSPERO Database of Systematic Reviews (CRD42018105507) Available

from:

http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018105507

Inclusion and exclusion criteria

Studies were included if genotyping of rs641738C>T (or rs626283G>C [R^2>0.98 in European and American populations[33]] / rs2576452C>T [R^2=0.92 in Guzman *et al.* [34]], which are in strong linkage disequilibrium with rs641738C>T) was conducted and data on one of the outcomes of interest were reported. Narrative review articles, *in vitro* studies, and investigations involving animals, fish, and invertebrates were excluded. Studies which investigated liver disease of other etiologies were also excluded. There was no restriction on ethnicity or ancestry. Types of studies eligible for inclusion were: case-control, cohort, genome-wide association studies, systematic reviews, and meta-analyses. Pre-print and abstract publications were eligible for inclusion provided that sufficiently detailed information could be obtained for assessment of inclusion/exclusion criteria, through availability of individual participant-level data, where possible. Several studies reported on the same cohort (or patient sample) in more than one article. In these instances, data only from the larger of the overlapping cohorts were included in analyses. A full list of overlapping cohorts and articles is in Supplementary Table 1.

Outcomes of interest

Given the strong evidence base for fibrosis stage as a prognostic marker in NAFLD[35], the primary outcome for the meta-analysis was the association of

rs641738C>T with presence of advanced fibrosis (F3-4 versus F0-2) in individuals with NAFLD.

Dichotomous secondary outcomes were: radiological diagnosis of NAFLD (or hepatic steatosis); presence of severe steatosis on liver biopsy (S3 versus S1-2); presence of non-alcoholic steatohepatitis (NASH) on liver biopsy (NASH versus non-alcoholic fatty liver (NAFL)); presence of any fibrosis (F1-4 versus F0); presence of hepatocellular carcinoma (HCC) in patients with NAFLD (NAFLD-HCC versus NAFLD).

Continuous secondary outcomes were: quantitative, radiological hepatic fat content; total cholesterol; high-density lipoprotein cholesterol (HDL); low-density lipoprotein cholesterol (LDL); triglycerides (TG).

Data collection

Details on the recruitment of controls and cases were obtained from each study and, where necessary, clarified by discussion with the study's authors. In particular, it was noted when cases and controls were not recruited from the same population or clinics.

Hepatic steatosis or NAFLD (as diagnosis) was evaluated as a dichotomous variable where radiological assessment (liver ultrasound, controlled attenuation parameter [CAP, with cut-off >240m/s], CT, MRI) were used. Hepatic fat content was collected as a continuous variable from CT, MRS, MRI, PDFF. Non-invasive assessment of hepatic fat content was also assessed using semi-quantitative scoring in the Fenland cohort, as previously described[36], and using CAP.

Individual participant-level histology data were extracted according to the NASH Clinical Research Network scoring system[37] and, where not otherwise diagnosed by a pathologist's assessment, NASH was defined using the Fatty Liver Inhibition of Progression (FLIP) algorithm[38]. The above data were collected for each genotype separately (CC, CT, and TT).

Participant demographics and characteristics meta-data were collected from each study, including: sex, age, ethnicity, presence of type 2 diabetes, body mass index (BMI). Where possible, individual patient-level data was obtained. The authors of 58 studies were contacted for additional data or clarification, of whom 48 replied. Data from 15 potentially relevant studies could not be included, which are listed in the Supplementary Methods.

Cohorts with genome-wide data

Relevant GWASs were identified as described above. Data were included where densely imputed genotyping results were available for rs641738C>T or rs626283C>G in all with >0.98 call rate. These cohorts have been described elsewhere and detailed description of the quality control processes for genome-wide data is available in their original descriptions (Table 1), but in brief, single nucleotide polymorphisms (SNP) with Hardy-Weinberg equilibrium p-values $<1x10^{-6}$ were excluded prior to imputation. Association analysis was performed for variants with mean allele frequencies >0.01 and with minimum imputation quality of >0.3.

Liver fat data from the UK BioBank was extracted under Application ID 9914 ('Determining the Outcomes of People with Liver Disease').

Additional details regarding the Avon Longitudinal Study of Parents and Children (ALSPAC) [39–41] and data extracted from the UK BioBank cohort is found in the Supplementary Methods.

Patient and Public Involvement

Patients and public were not directly involved in the planning of this meta-analysis; however, many of the original studies included were co-produced with patient and public involvement.

Study quality assessment

Two reviewers independently assessed risk of bias in each study by applying the Cochrane Risk of Bias in Cohort Studies tool[42].

Statistical Analysis

Genotype frequencies for each study were assessed for Hardy-Weinberg equilibrium using chi-squared test.

Due to the unclear effect of this variant on liver disease and previous studies using multiple different models of inheritance, genetic association analyses were performed using additive, dominant, and recessive models for each outcome.

For dichotomous outcomes, the effect statistic was calculated as an odds ratio between groups.

For analysis of diagnosis of liver fat, a sensitivity analysis was performed by excluding studies where there was a risk of confounding due to differences between cases and controls.

For analysis of effect on liver fat (continuous quantitative liver fat data from CT, MRI, MRS, or PDFF and semi-quantitative data using ultrasound or CAP), data were inverse normalized and an additive genetic model (coding the number of T alleles as 0, 1, and 2) was used with linear regression, adjusted for age, sex, and (where available) principal components of genetic ancestry. For other continuous variables where raw data was available (i.e. triglycerides, HDL, LDL, ALT, and total cholesterol), effect summary was calculated as a mean difference between CC and TT groups.

For meta-analysis of GWAS summary statistics, pooled effect summaries were calculated where traits had been logarithmically transformed and an additive model had been used in analyses. Beta regression coefficients were pooled using inverse-variance pooling with Fisher's z-transformation. Meta-analysis was performed using random effects throughout using DerSimonian-Laird method for estimation of tau^2.

Summary statistics were reported with 95% confidence intervals (CI). Data from paediatric and adult studies were analyzed separately. Sub-analysis was performed using only studies conducted in Caucasian populations (self-reported white, Non-Finnish-, or Finnish-European ethnicity) where data were available from at least four studies. This sub-analysis was selected due to initial identification of this variant in Caucasian individuals and further sub-analysis by ethnicity may be affected by differences in linkage disequilibrium between genetic ancestries. For studies including cohorts of multiple ethnicities, where available, data were analysed separately for each ethnicity. Heterogeneity between groups was described using the Q statistic, tau^2, and I^2. Where there was substantial heterogeneity (I^2>50%, P_Q<0.05) that

remained after sub-group analysis by non-Caucasian population, univariable random-effects meta-regression was performed using: proportion of female participants, mean age, proportion with type 2 diabetes, and mean BMI. In addition, for the analysis of hepatic fat as a quantitative trait, meta-regression was performed for the imaging modality used (e.g. MRI-PDFF, MRS).

For dichotomous outcomes: p-value<0.017 (i.e. p<0.05/3) was considered statistically significant due to testing outcomes against three genetic association models. For other outcomes, where only a single genetic model was used, p-value<0.05 was considered significant.

Bias was assessed using Egger's test and funnel plots where more than 10 studies were included. Where Egger's test suggested bias (p<0.05), a funnel plot was generated and missing studies were imputed using Duval & Tweedie's trim-and-fill procedure[43].

Analysis was performed using STATAv14 for Windows (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP), GraphPad Prism (v8.0 for Mac, GraphPad Software, La Jolla California, USA), and R 3.6.1[44,45]. Script used in analyses in R is available in the Supplementary Data.

RESULTS

Database search identified 1120 articles (Supplementary Fig. 1), of which 43 articles were included: 42 primary studies (Table 1 & Supplementary Tables 2-3) and one meta-analysis (Supplementary Table 4).

Study; Age group	Study design and sample size (N)	Liver biopsy (N)	Features and participant characteristics	Outcomes
Abeysekera 2019[46]; Adult	Population-based cohort N=2,919	NA	GWAS data from the ALSPAC birth cohort study with steatosis diagnosed using CAP >240kPa	NAFLD diagnosis, liver fat (CAP)
Adams 2013[47]; Paediatric	Population-based cohort N=948	NA	GWAS of adolescents with hepatic steatosis measured by US	NAFLD diagnosis, biochemistry
Caussy 2019[48]; Adult	Family, twin, & sibling-paired study N=104	NA	Community-based family case-control study with NAFLD diagnosed by MRI-PDFF	NAFLD diagnosis, liver fat, biochemistry
Chambers 2011[49]; Adult	Population-based cohorts N=61,089	NA	GWAS of adults from 21 population-based cohorts for concentration of serum liver enzymes	ALT (GWAS)
Chatterjee 2018[50]; Adult	Hospital-based case-control N=354	132	Exome-wide association study of adults with NAFLD diagnosed by LB or MRI	NAFLD diagnosis, liver fat, histology
Chen 2020[51]; Adult	Hospital-based cohort N=19,598	NA	GWAS of adults undergoing elective surgery with linkage to electronic health records	Biochemistry (GWAS)
Di Costanzo 2018[52]; Adult	Hospital-based case-control N=445	NA	Sample from adult obesity clinic with NAFLD diagnosed using US	NAFLD diagnosis, biochemistry
Di Costanzo 2019[53]; Paediatric	Hospital-based case-control N=230	NA	Sample from paediatric obesity clinic with NAFLD diagnosed using MRI hepatic fat fraction	NAFLD diagnosis, liver fat, biochemistry
Di Sessa 2018[54]; Paediatric	Hospital-based case-control N=1,002	NA	Sample from paediatric obesity clinic with NAFLD diagnosed using US	NAFLD diagnosis, biochemistry
DiStefano 2015[55]; Adult	Hospital-based case-control N=1,868	1868	GWAS of adults undergoing bariatric surgery with NAFLD diagnosed by LB	NAFLD diagnosis, histology
Donati 2017[6]; Adult	Hospital-based case-control N=1073 (Italian)	1123	Cohorts from adult liver/obesity clinics with NAFLD diagnosed using US and/or LB	NAFLD diagnosis, histology, HCC

	N=358 (UK)			
Dongiovanni 2018 (DHS)[56]; Adult	Population-based cohort N=2,675	NA	GWAS of adults from population-based cohort with NAFLD diagnosed using MRS	NAFLD diagnosis, liver fat, biochemistry
Dongiovanni 2018 (LBC)[56]; Adult	Hospital-based case-control N=1,515 (LBC)	1515	Samples from adult liver/obesity clinics and bariatric surgery with NAFLD diagnosed using LB	NAFLD diagnosis, histology, biochemistry
Gurdasani 2019[57]; Adult	Population-based cohorts N=13,116	NA	GWAS of adults from 4 population-based cohort studies, including concentration of serum liver enzymes and biochemistry	Biochemistry (GWAS)
Guzman 2018[34]; Adult	Clinical trial-based case-control N=148	NA	Sample of adults with type 2 diabetes who had taken part in one of two interventional diabetes trials, with NAFLD diagnosed on MRI	Liver fat
Hudert 2019[23]; Paediatric	Hospital-based case-control N=270	70	Cases from paediatric liver clinic with diagnosis of NAFLD using LB	NAFLD diagnosis, liver fat, histology
Kanai 2018[58]; Adult	Population-based cohort N=134,182	NA	GWAS of adults from a population cohort study with biobank, including concentration of serum liver enzymes and biochemistry	Biochemistry (GWAS)
Karajamaki 2019[59]; Adult	Population-based case-control N=935	NA	Population-level case-control study with recruitment of cases with hypertension; NAFLD diagnosed using US	NAFLD diagnosis, biochemistry
Kawaguchi 2018[60]; Adult	Mixed hospital- and population-based case-control N=8,571	936	Sample of adults from liver clinics with NAFLD diagnosed by LB	NAFLD diagnosis, histology, HCC, biochemistry
Koo 2020[61]; Adult	Hospital-based case-control N=461	461	Sample of adults from hospital clinics with NAFLD diagnosed by LB	NAFLD diagnosis, histology, biochemistry
Krawczyk 2016[62]; Adult	Hospital-based case-control N=84	84	Samples from adults undergoing bariatric surgery with NAFLD diagnosed using LB	Histology, biochemistry
Krawczyk 2018[63]; Adult	Hospital-based case-control N=515	295	Sample of adults from liver clinics with NAFLD diagnosed by US or LB	Histology, liver fat (CAP), biochemistry
Lin 2018[64]; Paediatric	Population-based cohort N=819	NA	Cohort of children recruited from schools with diagnosis of NAFLD using US	NAFLD diagnosis, biochemistry

Luukkonen 2020[65]; Adult	Hospital-based case-control N=793	236	Sample of adults undergoing bariatric surgery with NAFLD diagnosed by LB or MRS	NAFLD diagnosis, liver fat, histology, biochemistry
Mann 2018a[66]; Adult	Population-based cohort N=10,934	NA	Cohort study of adults from population with NAFLD diagnosed by US	NAFLD diagnosis, liver fat (US), biochemistry
Mann 2018b[67]; Paediatric	Hospital-based case-control N=379	306	Sample from paediatric obesity and liver clinics with diagnosis of NAFLD using LB or US	NAFLD diagnosis, histology, biochemistry
Middelberg 2011[68]; Adult & paediatric	Population-based cohort N=11,693	NA	Combination of two cohort studies: adolescent twins and siblings; and adult twins, both population-based, with GWAS including concentration of serum liver enzymes and biochemistry	Biochemistry (GWAS)
Moon 2019[69]; Adult	Population-based cohort N=6,949	NA	GWAS of adults from a population cohort study with biobank, including concentration of serum liver enzymes and biochemistry	Biochemistry (GWAS)
Prins 2017[70]; Adult	Population-based cohort N=9,731	NA	GWAS of adults from a population cohort study with biobank, including concentration of serum liver enzymes and biochemistry	Biochemistry (GWAS)
Reichert 2019[71]; Adult	Hospital-based case-control N= 54	NA	Sample of adults from liver clinics with NAFLD cirrhosis diagnosed by LB, or US/MRI/CT	HCC
Seidelin 2020[72]; Adult	Population-based cohort N=7511	NA	Hepatic steatosis measured by CT, part of the Copenhagen General Population Study	Liver fat
Sookoian 2018[20]; Adult	Hospital-based case-control N=634	372	Samples from adult liver clinics and bariatric surgery patients with NAFLD diagnosed by LB or US	NAFLD diagnosis, histology
Speliotes 2011[1]; Adult	Population-based cohorts N=4,244	NA	GWAS of adults from 4 population-based cohorts for hepatic fat content as determined by CT	Liver fat
Strnad 2018[73]; Adult	Hospital-based case-control N=672	672	Sample of adults from liver clinics with NAFLD diagnosed by LB	Histology
Tabassum 2019[74]; Adult	Population-based cohorts N=2,181	NA	GWAS of adults from 2 population-based cohorts with data on serum triglycerides	Biochemistry (GWAS)
UKBB 2019[75]; Adult	Population-based N=14,440 (liver fat); N=77,464 (coded fibrosis); N= 353,596	NA	GWAS of hepatic steatosis measured by MRI from the UK BioBank; adults with coded diagnosis of NAFLD and cirrhosis; and concentration of serum	Liver fat, biochemistry (GWAS), fibrosis (coding data)

	(biochemistry)		biochemistry available from GWAS	
Umano 2018[22]; Paediatric	Hospital-based case-control N=878	NA	Sample from paediatric obesity clinic with NAFLD diagnosed using MRI hepatic fat fraction	Liver fat, biochemistry
Verma 2017[76]; Adult	Hospital-based cohort N=31,466	NA	GWAS of adult hospital patients with linkage to electronic health records	Biochemistry (GWAS)
Viitasalo 2016[77]; Paediatric	Population-based cohort N=467	NA	Population cohort of children recruited from schools with measurement of ALT	Biochemistry
Wattacheril 2017[78]; Paediatric	Cases-only N=208	208	GWAS of Hispanic boys with NAFLD diagnosed by LB	Histology
Willer 2013[79]; Adult	Population-based cohorts N=347,532	NA	GWAS of adults from 45 population-based cohorts with data on serum lipids	Biochemistry (GWAS)
Young 2019[80]; Adult	Population-based cohorts N=3,555	NA	GWAS of adults from 5 population-based cohorts with data on ALT	Biochemistry (GWAS)
Zusi 2019[81]; Paediatric	Hospital-based case-control N=510	NA	Sample from paediatric obesity clinic with diagnosis of NAFLD using US	NAFLD diagnosis, biochemistry

Table 1. **Characteristics of studies included in the meta-analysis.** Further characteristics available in Supplementary Tables 2-3. ALSPAC, Avon Longitudinal Study of Parents and Children; CAP, controlled attenuation parameter; CT, computerized tomography; DHS, Dallas Heart Study; GWAS, genome-wide association study; LB, liver biopsy; LBC, Liver Biopsy Cohort; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopy; NA, not applicable; US, ultrasound.

In total, 1,047,265 individuals (5,711 children) were included in the meta-analysis. Most studies were in adults (32/42, 76%) and in individuals from

predominantly Caucasian populations (26/42, 62%). Of the 42 included studies, 14 studies (8,303 participants, hereof 584 children) reported data on liver histology.

Studies were generally of high quality, though in four studies[23,52,60,61] (three in adults and one in children) the control group was recruited from a different population or sample to the cases (Supplementary Table 2).

One previous meta-analysis was included[82], which used data from 5 case-control studies to assess the effect of rs641738C>T on diagnosis of NAFLD. The meta-analysis included 2560 cases and 8738 controls and found no evidence of an association between this variant and diagnosis of NAFLD (Supplementary Table 4).

Liver fat, NAFLD, and severe steatosis in adults

Nine studies (29,916 participants) reported data on hepatic fat as a continuous variable assayed by CT or MRI. On meta-analysis, rs641738C>T was associated with higher liver fat in studies in Caucasian populations, with a per T-allele change of β 0.034 (95% CI 0.018, 0.051), $P_z=5.7x10^{-5}$) standard deviations in inverse-normalized liver fat (Figure 1), whilst no difference was observed in non-Caucasian populations.

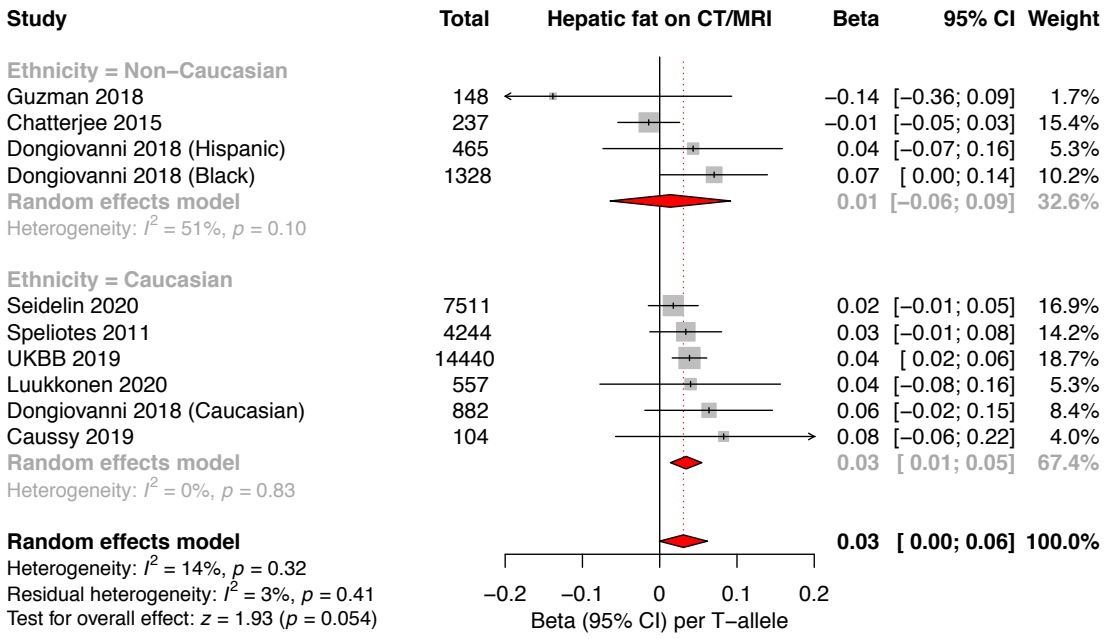

Figure 1. ***The effect of rs641738C>T on liver fat.*** *Data from 29,916 individuals with CT or MRI liver fat. rs641738C>T positively associated with liver fat in Caucasian populations, where data represents standard deviation change in normalized liver fat per T-allele. CI, confidence interval; UKBB, UK BioBank.*

A similar trend was observed using CAP and semi-quantitative ultrasound to assess steatosis severity in 12,224 adults (β 0.02 (95% CI -0.002, 0.04), Supplementary Figure 2).

rs641738C>T was also associated with NAFLD as a trait (OR 1.20 (95% CI 1.08, 1.32) using a recessive model of inheritance (Figure 2) but not using additive or dominant models (Supplementary Table 5). The effect was only observed in studies of Caucasian populations. The association remained after excluding three studies where there was a lack of similarity between cases

and controls (OR 1.20 (95% CI 1.08, 1.33), I^2 =18%, P_Q=.26) using a recessive model of inheritance.

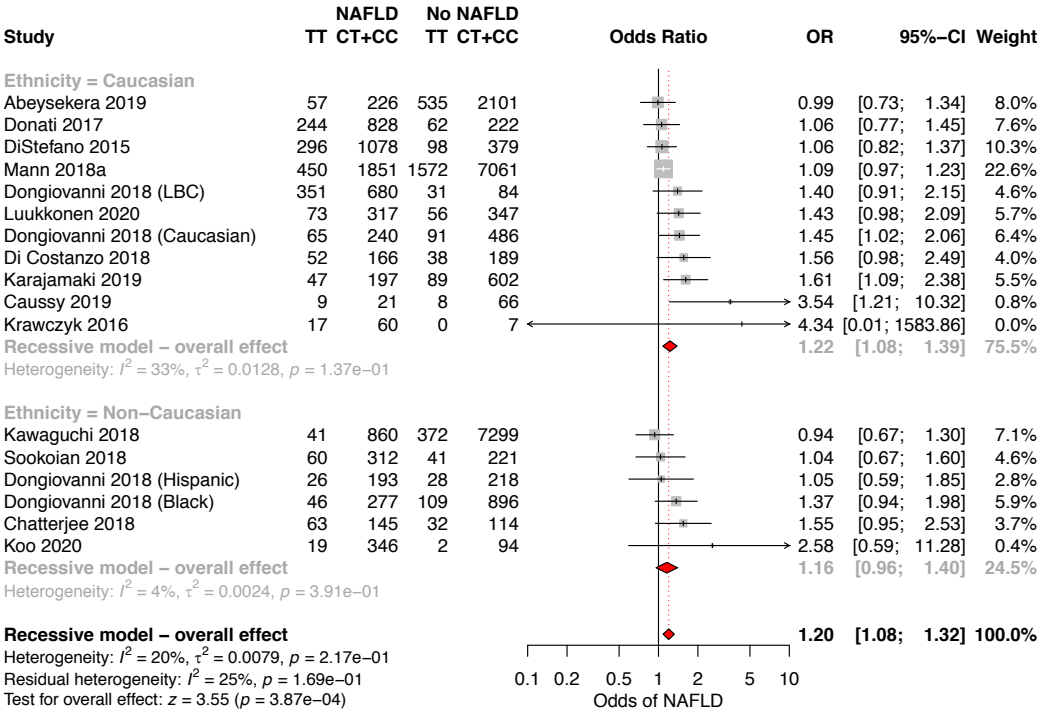

Figure 2. ***rs641738C>T is associated with higher odds of diagnosis of NAFLD****. Data from 33,263 adults (9,713 cases and 23,550 controls) with radiologically defined steatosis for presence versus absence of NAFLD. CI, confidence interval; LBC, Liver Biopsy Cohort; OR, odds ratio.*

However, Egger's test suggested evidence of study distribution (publication) bias (p=0.01) and when using the Trim and Fill method to account for this bias, the positive association remained but was attenuated (OR 1.13 (95% CI 1.01, 1.26), P_Z=0.026, Supplementary Figure 3).

In patients with NAFLD, rs641738C>T was associated with the presence of severe steatosis (S1-2 vs. S3) on liver biopsy (OR 1.32 (95% CI 1.10, 1.59],

Supplementary Figure 4), which was again due to a positive association in Caucasian individuals only (Table 2).

Outcome	Genetic Model	Population	No. studies	Heterogeneity		Effect summary			
				I2 (%)	PQ	OR	95% CI		Pz
NAFLD diagnosis (control vs NAFLD)	Recessive	Overall	17	20	0.27	1.20	1.08	1.32	$4.0x10^{-4}$
NAFLD diagnosis (control vs NAFLD)	Recessive	Non-Caucasian	6	4	0.39	1.16	0.96	1.40	0.12
NAFLD diagnosis (control vs NAFLD)	Recessive	Caucasian	11	33	0.19	1.22	1.08	1.39	0.0021
Severe steatosis (S1-2 vs S3)	Recessive	Overall	8	0	0.85	1.32	1.10	1.59	0.0032
Severe steatosis (S1-2 vs S3)	Recessive	Non-Caucasian	2	0	0.78	1.22	0.56	2.65	0.63
Severe steatosis (S1-2 vs S3)	Recessive	Caucasian	6	0	0.63	1.33	1.10	1.61	0.0034
NASH (NAFL vs NASH)	Recessive	Overall	9	0	0.60	1.20	1.04	1.39	0.013
NASH (NAFL vs NASH)	Recessive	Non-Caucasian	4	0	0.78	1.24	0.85	1.80	0.26
NASH (NAFL vs NASH)	Recessive	Caucasian	5	32	0.26	1.22	1.00	1.50	0.05
Any fibrosis (F0 vs F1-4)	Recessive	Overall	9	40	0.11	1.32	1.08	1.62	0.0073
Any fibrosis (F0 vs F1-4)	Recessive	Non-Caucasian	3	18	0.30	1.77	0.92	3.43	0.09
Any fibrosis (F0 vs F1-4)	Recessive	Caucasian	6	43	0.13	1.21	1.06	1.40	0.007
Advanced fibrosis (F0-2 vs F3-4)	Recessive	Overall	7	0	0.65	1.28	1.04	1.57	0.019
Advanced fibrosis (F0-2 vs F3-4)	Recessive	Non-Caucasian	2	0	0.64	0.96	0.50	1.85	0.90
Advanced fibrosis (F0-2 vs F3-4)	Recessive	Caucasian	5	0	0.52	1.32	1.06	1.63	0.011
HCC (NAFLD vs NAFLD-HCC)	Dominant	Overall	4	0	0.65	1.64	1.18	2.27	0.0031

Table 2. ***Summary of results in adults from meta-analyses for dichotomous outcomes****. Meta-analyses were performed using random*

effects with subgroup analysis for Caucasian and Non-Caucasian populations. Additive, recessive, and dominant genetic models were tested for all outcomes. Results using a recessive model of inheritance (CC+CT vs. TT) are shown for all outcomes, except for HCC, where a dominant model (CC vs. CT+TT) is shown. Due to use of three genetic models, critical p-value for effect summary is P_z<0.017. Full results (with all genetic models) are in Supplementary Table 5. CI, confidence interval; HCC, hepatocellular carcinoma; NAFL, non-alcoholic fatty liver; NASH, non-alcoholic steatohepatitis; OR, odds ratio.

Histological NASH in adults

Data from nine studies (6,155 participants) found that rs641738C>T had a positive association with the presence of NASH on biopsy in adults (OR 1.20 (95% 1.04, 1.39), P_z=0.013 Supplementary Figure 5).

Fibrosis in adults

Liver biopsy data on presence of advanced fibrosis was available from seven studies (6,211 adults). Our primary outcome, presence of advanced fibrosis in adults (stage F0-2 versus stage F3-4), was positively associated with rs641738C>T in Caucasian populations (OR 1.32 (95% 1.06, 1.63), P_z=0.011), though the effect was borderline when non-Caucasian populations were included (Figure 3). In addition, two studies used ICD-codes (International Statistical Classification of Diseases and Related Health Problems) in the UKBB cohort to identify individuals with NAFLD and advanced fibrosis or cirrhosis[51,83]. Both found positive associations below

genome-wide significance: for example, using an additive model of inheritance Emdin *et al.* found the association between rs641738C>T and cirrhosis as β 1.15 (SE 0.13), using an additive genetic model.

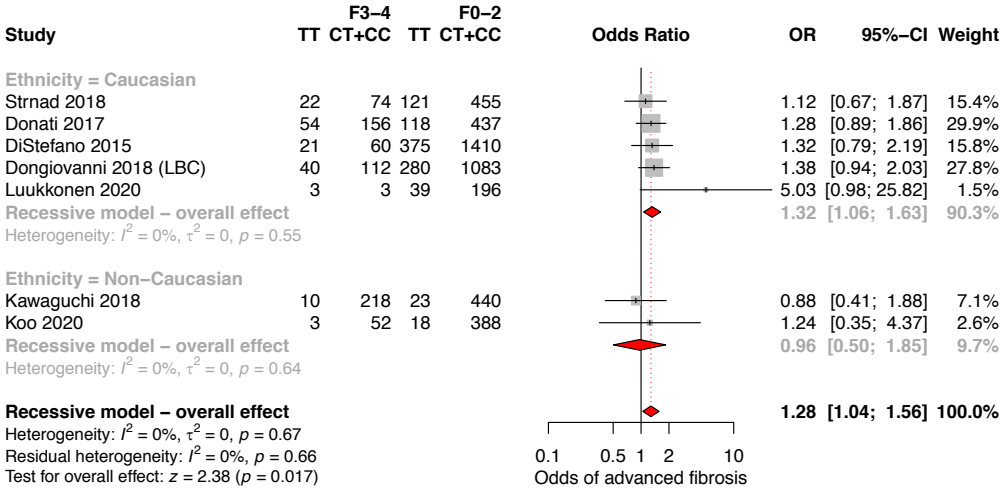

Figure 3. ***rs641738C>T is associated with higher odds of advanced fibrosis in Caucasian populations with NAFLD****. Data from 6,211 adults (828 cases and 5,383 controls) with biopsy-proven NAFLD comparing advanced fibrosis (F3-4) versus F0-2, using a recessive model of inheritance (CC+CT vs. TT). CI, confidence interval; LBC, Liver Biopsy Cohort; OR, odds ratio.*

Presence of any fibrosis (F0 versus F1-4) was also positively associated with rs641738C>T overall (OR 1.32 (95% 1.08, 1.62) and in Caucasian populations as a sub-group (Supplementary Figure 6).

Development of hepatocellular carcinoma

Four cohorts (2,328 participants, 228 cases of NAFLD-HCC) reported on development of HCC in patients with NAFLD. rs641738C>T was associated

with increased odds of HCC in NAFLD only when using a dominant model (CC vs. CT+TT) of inheritance (OR 1.64 (95% CI 1.18, 2.27), P_z=0.003, Figure 4).

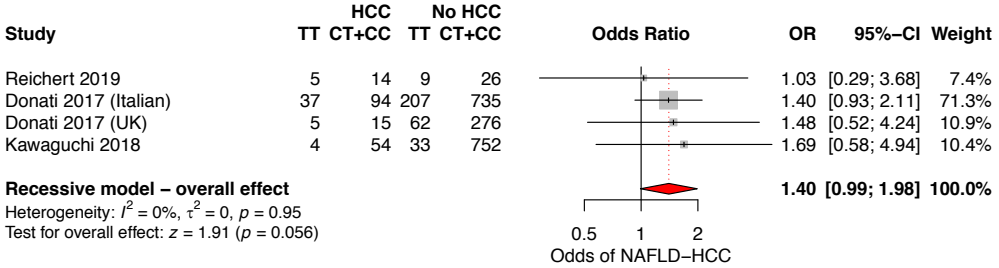

Figure 4. ***rs641738C>T is associated with higher odds of NAFLD-HCC in****. Data from 2,328 adults with NAFLD assessing for the presence versus absence of HCC, using a dominant model of inheritance (CC vs. CT+TT).*

Effect on alanine aminotransferase (ALT)

Data from GWAS using log-transformed ALT (609,794 participants) were available for meta-analysis to investigate the role of rs641738C>T on ALT. The variant showed a positive association with ALT (β 0.004 (95% CI 0.002, 0.007), P_z=0.002), which on sub-analysis was observed in Caucasian populations but not in non-Caucasian populations (Figure 5 & Supplementary Table 6).

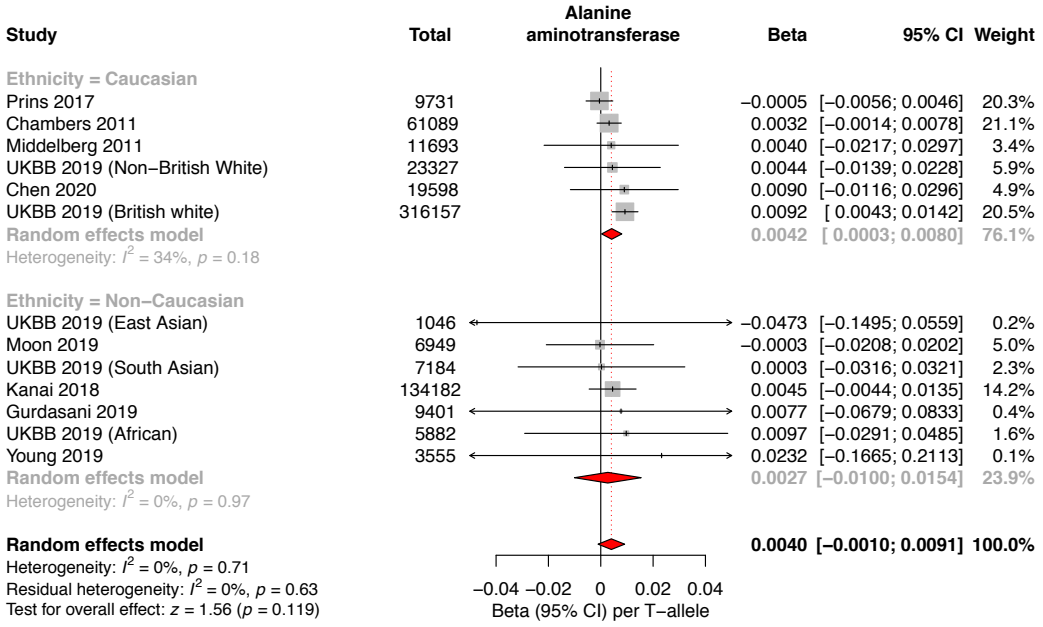

Figure 5. ***rs641738C>T is positively associated with alanine aminotransferase (ALT) in Caucasian populations in genome-wide association studies (GWAS)****. Meta-analysis of GWAS summary statistics from 609,794 participants for the association between rs641738C>T on logarithmically-transformed ALT using linear regression. CI, confidence interval; UKBB, UK BioBank.*

Additionally, in the UKBB cohort, rs641738C>T was associated with a small, but statistically significant (P=2.0x10^{-8}) increase in un-transformed ALT: 0.18 IU/L higher ALT per T-allele in this variant (Supplementary Table 7).

In the remaining cohort and case-control studies included in the meta-analysis (15,208 adults), rs641738C>T was not found to be associated with a change in ALT: mean difference between CC vs. TT 0.2 IU/L (95% CI -0.4, 0.8), Supplementary Table 8).

Effect on serum lipids and insulin

Data from GWAS using log-transformed serum triglycerides (850,241 participants) found that rs641738C>T was associated with lower triglycerides (β -0.01 (95% CI -0.018, -0.006), P_z=1.5x10^{-4}), which on sub-analysis was observed in Caucasian populations but not in non-Caucasian populations (Supplementary Fig. 7). Similar findings were obtained from meta-analysis of mean difference in triglycerides between CC vs. TT -0.03 mmol/L (95% CI -0.05, -0.01), Supplementary Table 8).

Data from GWAS (852,409 participants) found rs641738C>T to be positively associated with total cholesterol (β 0.007 (95% CI 0.003, 0.01), P_z=2.3x10^{-4}), in Caucasian populations but not in non-Caucasian populations (Supplementary Fig. 8). A borderline positive association was also observed in between rs641738C>T and high-density lipoprotein (HDL) cholesterol (β 0.009 (95% CI 0.001, 0.02), P_z=0.02), Supplementary Table 6). There was no effect observed on low-density lipoprotein (LDL) or fasting insulin levels (Supplementary Tables 7 & 8).

Effect of rs641738C>T on paediatric NAFLD

Data from ten studies (5,711 children) was used in the meta-analysis. rs641738C>T was not significantly associated with the diagnosis of NAFLD, liver fat content, or stage of liver histology (Supplementary table 9). The variant was observed to be associated with lower serum triglycerides in non-Caucasian children comparing CC vs. TT: -0.11 mmol/L (95% CI -0.14, -0.08).

DISCUSSION

Identification of genetic variants associated with NAFLD has the potential to inform pre-clinical research and our understanding of hepatic metabolism. In this meta-analysis we have validated rs641738C>T near *MBOAT7* as a risk factor for the full spectrum of NAFLD in Caucasian adults.

A two-stage GWAS initially identified rs641738C>T as a genome-wide significant locus for alcohol-related cirrhosis[4]. *MBOAT7* was a potentially interesting target as an enzyme involved in (phospho)lipid metabolism, conceptually similar to other SNVs at GWAS-significance in alcoholic and non-alcoholic liver disease, namely *TM6SF2* and *PNPLA3*. Later studies found the variant to influence the full spectrum of fatty liver disease, from steatosis to NASH, to fibrosis, cirrhosis and HCC[5,18]. However, these associations have not been consistently replicated in the literature[20]. We conducted a meta-analysis to firmly establish the association of

Main findings

We found that rs641738C>T was associated with higher liver fat content, higher ALT, and with higher odds of NAFLD diagnosis, severe steatosis, NASH, fibrosis, and HCC, particularly in Caucasian adults and in the homozygous 'TT' genotype. The effects sizes of rs641738C>T reported here are small compared to those of *PNPLA3* p.I148M and *TM6SF2* p.E167K, the two strongest steatogenic variants[3]. Also, the magnitude of change in alanine aminotransferase is small relative to that associated with variants in *PNPLA3*, *HSD17B13*, *MARC1*, and *TM6SF2*. The marginal positive effect on

hepatic triglyceride content may suggest this variant acts through alterations in the composition of hepatic lipid, as well as quantity[18]. This is consistent with pre-clinical data on lipotoxicity, where the composition of hepatic fats influence development of NASH. On the other hand, a recent Mendelian randomization study using these variables as instruments to assess causality of fatty liver in determining fibrosis has shown the effect of steatosis highly correlates with fibrosis in all the genetic variables indicating that quantity of lipid rather than quality may be more important[56]. Functional studies are needed to understand the relationship between quality/quantity of fat and hepato-toxic/-protective mechanism in causing progression of disease.

The function of this variant is still relatively poorly understood and there is conflicting evidence as to whether rs641738C>T is associated with changes in hepatic expression of *MBOAT7*. Results from the GTEx Consortium show a strong negative association with T-allele[16], which is supported by data from Schadt *et al.[84]*. *MBOAT7* protein expression correlated with mRNA in liver biopsies from Mancina *et al.[5]* but this finding was not replicated by Sookoian *et al.[20]*. *MBOAT7* encodes LPIAT1, a 6 transmembrane domain protein involved in acyl-chain remodeling of membranes that influence intracellular membrane composition and circulating phosphatidylinositols[85]. Further, recent metabolite profiling data implicates *MBOAT7* as the causal gene for this SNV[66]. Moreover, TMC4 was found with a low expression in the liver[5] that is consistent with no mechanistic data supporting its role in NAFLD.

The hypothesis that *MBOAT7* is the causal gene underlying the association with liver disease at the locus is supported by the observation that mice deficient for *MBOAT7* have altered hepatic concentrations of polyunsaturated phosphatidylinositol[85]. Similarly, metabolite data from humans is strongly suggestive that rs641738C>T reduces MBOAT7 function[86]. In addition, two independent groups have found that loss of *MBOAT7* (but not *TMC4)* increases the severity of NAFLD in mice fed a high-fat diet[87,88].

These analyses suggest that rs641738C>T impacts the severity of NAFLD through a recessive model of inheritance, though some analyses (e.g. liver fat and ALT) were suggestive of a role using an additive genetic model. Other genetic variants are known to impact on all-cause mortality in a recessive manner, notably variants that perturb *HFE[83].* Further mechanistic work is required to understand the extent to which the halo-insufficient state affects hepatocyte function.

We found no evidence of rs641738C>T on insulin resistance: the key driver of hepatic steatosis, as determined by unaltered fasting insulin concentrations. GWAS meta-analyses of type 2 diabetes have implicated p.I148M in *PNPLA3* and p.E167K in *TM6SF2* as significant risk loci (albeit with very modest effect size as compared to their effects on liver disease)[89] and Mendelian randomization studies indicate a causal role in determining insulin resistance mediated by the degree of liver damage[56,75]. Similarly, these two variants are associated with reduced risk of coronary artery disease; whilst our

analysis did find lower serum triglycerides to be associated with this variant it has not been associated with lower rates of cardiovascular disease[90].

A strength of this meta-analysis is the large number of individuals with liver biopsy-derived phenotypic data as well as use of population-based GWAS data. The larger number of included studies and participants is likely to account for the different conclusions reached in this study compared to the previous meta-analysis by Xia *et al.[82]*.

Limitations and quality of evidence

An important practical consideration is the population frequency of this variant in different ethnicities. The mean allelic frequency of the effect (T) allele is highly variable: from 0.24 in East Asians compared to 0.53 in those of South Asian ancestry[91]. Moreover, the majority of studies included in this meta-analysis used self-reported ethnicity, rather than genetic ancestry[92].

Though this analysis did include data from individuals of multiple ethnicities (and genetic ancestries) we only found evidence of an effect of this variant in Caucasian individuals. This is consistent with the initial discovery however, due to differences in patterns of linkage disequilibrium, we cannot exclude the possibility that a different nearby locus is associated with liver-related phenotypes in individuals of other genetic ancestries.

We found significant differences between adult and paediatric histological analyses. Whilst there were fewer clinical events (e.g. with advanced fibrosis)

in children, the analyses did not show a trend congruous with those in adults. Paediatric NAFLD has a different histological phenotype to that of adults (with prominent periportal inflammation) and it is therefore plausible that this is a true lack of association in children with NAFLD.

The magnitude of effect observed across all associations is small in comparison to other well-established variants. The clinical relevance of rs738409C>G in *PNPLA3* has been validated with hard end-points[93] but large cohorts will be required to prospectively demonstrate the clinical risk associated with this variant near *MBOAT7*.

Though there was minimal heterogeneity across included studies, there was evidence of publication bias but the effect on diagnosis of NAFLD appeared to persist after attempting to account for this. Also of note, the numbers of individuals with NAFLD and HCC were comparatively low. It was also unique in only demonstrating an effect in the dominant, rather than recessive, model of inheritance. Further work in this area may improve the accuracy of effect estimates.

Conclusions

rs641738C>T near *MBOAT7* is positively associated with liver fat, ALT, and histological severity in Caucasian adults with NAFLD, but negatively associated with serum triglycerides and with relatively small effect sizes throughout. These data validate this locus as significant in the pathogenesis of NAFLD.

ACKNOWLEDGEMENTS

The authors are grateful to the Raine Study participants and their families, and to the Raine Study research staff for cohort coordination and data collection. The authors gratefully acknowledge the following institutes for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA), Curtin University, the Raine Medical Research Foundation, the UWA Faculty of Medicine, Dentistry and Health Sciences, the Telethon Kids Institute, the Women and Infants Research Foundation (King Edward Memorial Hospital) and Edith Cowan University).

We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

This study has been conducted using data from the Fenland study. The authors gratefully acknowledge the help of the MRC Epidemiology Unit Support Teams, including Field, Laboratory and Data Management Teams.

The authors are grateful to the members of the EU-PNAFLD Registry, including Anita Vreugdenhil, Anna Alisi, Piotr Socha, Wojciech Jańczyk, Ulrich Baumann, Sanjay Rajwal, Indra van Mourik, Florence Lacaille, Myriam Dabbas, Deirdre A. Kelly, Quentin M. Anstee and the late Valerio Nobili. We would also like to thank Naga Chalasani for his helpful comments. This

research has made use of the UK Biobank resource under project number 9914.

REFERENCES

[1] Speliotes EK, Yerges-armstrong LM, Wu J, Hernaez R, Lauren J, Palmer CD, et al. Genome-Wide Association Analysis Identifies Variants Associated with Nonalcoholic Fatty Liver Disease That Have Distinct Effects on Metabolic Traits. PLoS Genet 2011;7:e1001324.

[2] Romeo S, Sanyal A, Valenti L. Leveraging Human Genetics to Identify Potential New Treatments for Fatty Liver Disease. Cell Metab 2020;31:35–45.

[3] Sookoian S, Pirola CJ, Valenti L, Davidson NO. Genetic pathways in nonalcoholic fatty liver disease: Insights from systems biology. Hepatology 2020. https://doi.org/10.1002/hep.31229.

[4] Buch S, Stickel F, Trépo E, Way M, Herrmann A, Nischalke HD, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. Nat Genet 2015;47:1443–8.

[5] Mancina RM, Dongiovanni P, Petta S, Pingitore P, Meroni M, Rametta R, et al. The MBOAT7-TMC4 Variant rs641738 Increases Risk of Nonalcoholic Fatty Liver Disease in Individuals of European Descent. Gastroenterology 2016;150:1219–30e6.

[6] Donati B, Dongiovanni P, Romeo S, Meroni M, McCain M, Miele L, et al. MBOAT7 rs641738 variant and hepatocellular carcinoma in non-cirrhotic individuals. Sci Rep 2017;7:4492.

[7] Thabet K, Asimakopoulos A, Shojaei M, Romero-Gomez M, Mangia A, Irving WL, et al. MBOAT7 rs641738 increases risk of liver inflammation and transition to fibrosis in chronic hepatitis C. Nat Commun 2016;7:12757.

[8] Thabet K, Chan HLY, Petta S, Mangia A, Berg T, Boonstra A, et al. The membrane-bound O-acyltransferase domain-containing 7 variant rs641738 increases inflammation and fibrosis in chronic hepatitis B. Hepatology 2017;65:1840–50.

[9] Ezzikouri S, Elfihry R, Chihab H, Elmessaoudi-Idrissi M, Zaidane I, Jadid FZ, et al. Effect of MBOAT7 variant on hepatitis B and C infections in Moroccan patients. Sci Rep 2018;8:12247.

[10] Freund C, Wahlers A, Begli NH, Leopold Y, Klöters-Plachky P, Mehrabi A, et al. The MBOAT7 rs641738 variant is associated with an improved outcome in primary sclerosing cholangitis. Clin Res Hepatol Gastroenterol 2020. https://doi.org/10.1016/j.clinre.2019.12.006.

[11] Abul-Husn NS, Cheng X, Li AH, Xin Y, Schurmann C, Stevis P, et al. A Protein-Truncating *HSD17B13* Variant and Protection from Chronic Liver Disease. N Engl J Med 2018;378:1096–106.

[12] Anstee QM, Darlay R, Cockell S, Meroni M, Govaere O, Tiniakos D, et al. Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically-characterised cohort. J Hepatol 2020. https://doi.org/10.1016/j.jhep.2020.04.003.

[13] Matsuda S, Inoue T, Lee HC, Kono N, Tanaka F, Gengyo-Ando K, et al. Member of the membrane-bound O-acyltransferase (MBOAT) family encodes a lysophospholipid acyltransferase with broad substrate specificity. Genes Cells 2008;13:879–88.

[14] Gijón MA, Riekhof WR, Zarini S, Murphy RC, Voelker DR. Lysophospholipid acyltransferases and arachidonate recycling in human neutrophils. J Biol Chem 2008;283:30235–45.

[15] Johansen A, Rosti RO, Musaev D, Sticca E, Harripaul R, Zaki M, et al. Mutations in MBOAT7, Encoding Lysophosphatidylinositol Acyltransferase I, Lead to Intellectual Disability Accompanied by Epilepsy and Autistic Features. Am J Hum Genet 2016;99:912–6.

[16] The Gtex Consortium. The Genotype-Tissue Expression (GTEx) project. Nat

Genet 2013;45:580–5.
[17] Meroni M, Dongiovanni P, Longo M, Carli F, Baselli G, Rametta R, et al. Mboat7 down-regulation by hyper-insulinemia induces fat accumulation in hepatocytes. EBioMedicine 2020;52:102658.
[18] Luukkonen PK, Zhou Y, Hyötyläinen T, Leivonen M, Arola J, Orho-Melander M, et al. The MBOAT7 variant rs641738 alters hepatic phosphatidylinositols and increases severity of non-alcoholic fatty liver disease in humans. J Hepatol 2016;65:1263–5.
[19] Krawczyk M, Rau M, Schattenberg JM, Bantel H, Pathil A, Demir M, et al. Combined effects of the *PNPLA3* rs738409, *TM6SF2* rs58542926, and *MBOAT7* rs641738 variants on NAFLD severity: a multicenter biopsy-based study. J Lipid Res 2017;58:247–55.
[20] Sookoian S, Flichman D, Garaycoechea ME, Gazzi C, Martino JS, Castaño GO, et al. Lack of evidence supporting a role of TMC4-rs641738 missense variant - MBOAT7- intergenic downstream variant - In the Susceptibility to Nonalcoholic Fatty Liver Disease. Sci Rep 2018;8:5097.
[21] Koo BK, Joo SK, Kim D, Bae JM, Park JH, Kim JH, et al. Additive effects of PNPLA3 and TM6SF2 on the histological severity of non-alcoholic fatty liver disease. J Gastroenterol Hepatol 2018;33:1277–85.
[22] Umano GR, Caprio S, Di Sessa A, Chalasani N, Dykas DJ, Pierpont B, et al. The rs626283 variant in the MBOAT7 gene is associated with insulin resistance and fatty liver in Caucasian obese youth. Am J Gastroenterol 2018;113:376–83.
[23] Hudert CA, Selinski S, Rudolph B, Bläker H, Loddenkemper C, Thielhorn R, et al. Genetic determinants of steatosis and fibrosis progression in paediatric non-alcoholic fatty liver disease. Liver Int 2019;39:540–56.
[24] Yalnızoğlu D, Özgül RK, Oğuz KK, Özer B, Yücel-Yılmaz D, Gürbüz B, et al. Expanding the phenotype of phospholipid remodelling disease due to MBOAT7 gene defect. J Inherit Metab Dis 2019;42:381–8.
[25] Jacher JE, Roy N, Ghaziuddin M, Innis JW. Expanding the phenotypic spectrum of MBOAT7-related intellectual disability. Am J Med Genet B Neuropsychiatr Genet 2019;180:483–7.
[26] Pelusi S, Baselli G, Pietrelli A, Dongiovanni P, Donati B, McCain MV, et al. Rare Pathogenic Variants Predispose to Hepatocellular Carcinoma in Nonalcoholic Fatty Liver Disease. Sci Rep 2019;9:3682.
[27] MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res 2017;45:D896–901.
[28] Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 2019;35:4851–3.
[29] Type 2 diabetes knowledge portal n.d. http://www.type2diabetesgenetics.org/ (accessed April 3, 2020).
[30] Cardiovascular Disease Knowledge Portal n.d. http://www.broadcvdi.org/ (accessed April 3, 2020).
[31] Little J, Higgins J (editors). The HuGENet™ HuGE Review Handbook, version 1.0. Centers for Disease Control and Prevention; 2006.
[32] Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of Observational Studies. JAMA 2000.
[33] Arnold M, Raffler J, Pfeufer A, Suhre K, Kastenmüller G. SNiPA: an interactive, genetic variant-centered annotation browser. Bioinformatics 2015;31:1334–6.
[34] Guzman CB, Duvvuru S, Akkari A, Bhatnagar P, Battioui C, Foster W, et al. Coding variants in PNPLA3 and TM6SF2 are risk factors for hepatic steatosis and elevated serum alanine aminotransferases caused by a glucagon receptor antagonist. Hepatology Communications 2018;2:561–70.
[35] Taylor RS, Taylor RJ, Bayliss S, Hagström H, Nasr P, Schattenberg JM, et al.

Association Between Fibrosis Stage and Outcomes of Patients With Nonalcoholic Fatty Liver Disease: A Systematic Review and Meta-Analysis. Gastroenterology 2020. https://doi.org/10.1053/j.gastro.2020.01.043.
[36] De Lucia Rolfe E, Brage S, Sleigh A, Finucane F, Griffin SJ, Wareham NJ, et al. Validity of ultrasonography to assess hepatic steatosis compared to magnetic resonance spectroscopy as a criterion method in older adults. PLoS One 2018;13:87–99.
[37] Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–21.
[38] Bedossa P, Burt AA, Gouw AHA, Lackner C, Schirmacher P, Terracciano L, et al. Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. Hepatology 2014;60:565–75.
[39] Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort Profile: the “children of the 90s”--the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol 2013;42:111–27.
[40] Northstone K, Lewcock M, Groom A, Boyd A, Macleod J, Timpson N, et al. The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. Wellcome Open Res 2019;4:51.
[41] Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81.
[42] Cochrane Collaboration. Tool to assess risk of bias in cohort studies 2017.
[43] Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455–63.
[44] R Core Team. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing 2019.
[45] Harrer M, Cuijpers P, Furukawa TA, Ebert DD. Doing Meta-Analysis in R: A Hands-on Guide. PROTECT Lab Erlangen 2019.
[46] Abeysekera KWM, Fernandes GS, Hammerton G, Portal AJ, Gordon FH, Heron J, et al. Prevalence of steatosis and fibrosis in young adults in the UK: a population-based study. Lancet Gastroenterol Hepatol 2020;5:295–305.
[47] Adams LA, White SW, Marsh JA, Lye SJ, Connor KL, Maganga R, et al. Association between liver-specific gene polymorphisms and their expression levels with nonalcoholic fatty liver disease. Hepatology 2013;57:590–600.
[48] Caussy C, Bhargava M, Villesen IF, Gudmann NS, Leeming DJ, Karsdal MA, et al. Collagen Formation Assessed by N-Terminal Propeptide of Type 3 Procollagen Is a Heritable Trait and Is Associated With Liver Fibrosis Assessed by Magnetic Resonance Elastography. Hepatology 2019;70:127–41.
[49] Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, Van Der Harst P, et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet 2011;43:1131–8.
[50] Chatterjee A, Das K, Singh P, Mondal D, Ghosh R, Chowdhury A, et al. Exome-wide association study with hepatic fat content in nonalcoholic fatty liver disease reveals significant association with 5 novel QTLs. Hepatol Int 2018;12:181.
[51] Chen VL, Chen Y, Du X, Handelman SK, Speliotes EK. Genetic variants that associate with cirrhosis have pleiotropic effects on human traits. Liver Int 2020;40:405–15.
[52] Di Costanzo A, Belardinilli F, Bailetti D, Sponziello M, D’Erasmo L, Polimeni L, et al. Evaluation of Polygenic Determinants of Non-Alcoholic Fatty Liver Disease (NAFLD) By a Candidate Genes Resequencing Strategy. Sci Rep 2018;8:1–10.
[53] Di Costanzo A, Pacifico L, Chiesa C, Perla FM, Ceci F, Angeloni A, et al. Genetic and metabolic predictors of hepatic fat content in a cohort of Italian

children with obesity. Pediatr Res 2019;85:671–7.
[54] Di Sessa A, Umano GR, Cirillo G, Del Prete A, Iacomino R, Marzuillo P, et al. The Membrane-bound O-Acyltransferase7 rs641738 Variant in Pediatric Nonalcoholic Fatty Liver Disease. J Pediatr Gastroenterol Nutr 2018;67:69–74.
[55] DiStefano JK, Kingsley C, Craig Wood G, Chu X, Argyropoulos G, Still CD, et al. Genome-wide analysis of hepatic lipid content in extreme obesity. Acta Diabetol 2014;52:373–82.
[56] Dongiovanni P, Stender S, Pietrelli A, Mancina RM, Cespiati A, Petta S, et al. Causal relationship of hepatic fat with liver damage and insulin resistance in nonalcoholic fatty liver. J Intern Med 2018;283:356–70.
[57] Gurdasani D, Carstensen T, Fatumo S, Chen G, Franklin CS, Prado-Martinez J, et al. Uganda Genome Resource Enables Insights into Population History and Genomic Discovery in Africa. Cell 2019;179:984–1002.e36.
[58] Kanai M, Akiyama M, Takahashi A, Matoba N, Momozawa Y, Ikeda M, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nat Genet 2018;50:390–400.
[59] Käräjämäki AJ, Hukkanen J, Kauma H, Kesäniemi YA, Ukkola O. Metabolic syndrome but not genetic polymorphisms known to induce NAFLD predicts increased total mortality in subjects with NAFLD (OPERA study). Scand J Clin Lab Invest 2019:1–8.
[60] Kawaguchi T, Shima T, Mizuno M, Mitsumoto Y, Umemura A, Kanbara Y, et al. Risk estimation model for nonalcoholic fatty liver disease in the Japanese using multiple genetic markers. PLoS One 2018;13:1–16.
[61] Koo BK, Joo SK, Kim D, Lee S, Bae JM, Park JH, et al. Development and Validation of a Scoring System, Based on Genetic and Clinical Factors, to Determine Risk of Steatohepatitis in Asian Patients with Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2020. https://doi.org/10.1016/j.cgh.2020.02.011.
[62] Krawczyk M, Jiménez-Agüero R, Alustiza JM, Emparanza JI, Perugorria MJ, Bujanda L, et al. PNPLA3 p.I148M variant is associated with greater reduction of liver fat content after bariatric surgery. Surg Obes Relat Dis 2016;12:1838–46.
[63] Krawczyk M, Bantel H, Rau M, Schattenberg JM, Grünhage F, Pathil A, et al. Could inherited predisposition drive non-obese fatty liver disease? Results from German tertiary referral centers. J Hum Genet 2018;63:621–6.
[64] Lin YC, Chang PF, Chang MH, Ni YH. Genetic determinants of hepatic steatosis and serum cytokeratin-18 fragment levels in Taiwanese children. Liver Int 2018;38:1300–7.
[65] Luukkonen PK, Tukiainen T, Juuti A, Sammalkorpi H, Haridas PAN, Niemelä O, et al. Hydroxysteroid 17-β dehydrogenase 13 variant increases phospholipids and protects against fibrosis in nonalcoholic fatty liver disease. JCI Insight 2020;5. https://doi.org/10.1172/jci.insight.132158.
[66] Mann JP, Pietzner M, Wittemans LB, De Lucia Rolfe E, Nicola D, Imamura F, et al. Insights into genetic variants associated with NASH-fibrosis from metabolite profiling. J Hepatol 2019;70:e383–624.
[67] Mann JP, Vreugdenhil A, Socha P, Jańczyk W, Baumann U, Rajwal S, et al. European paediatric non-alcoholic fatty liver disease registry (EU-PNAFLD): Design and rationale. Contemp Clin Trials 2018;75:67–71.
[68] Middelberg RPS, Ferreira MAR, Henders AK, Heath AC, Madden PAF, Montgomery GW, et al. Genetic variants in LPL, OASL and TOMM40/APOE-C1-C2-C4 genes are associated with multiple cardiovascular-related traits. BMC Med Genet 2011;12:123.
[69] Moon S, Kim YJ, Han S, Hwang MY, Shin DM, Park MY, et al. The Korea Biobank Array: Design and Identification of Coding Variants Associated with Blood Biochemical Traits. Sci Rep 2019;9:1382.
[70] Prins BP, Kuchenbaecker KB, Bao Y, Smart M, Zabaneh D, Fatemifar G, et al.

Genome-wide analysis of health-related biomarkers in the UK Household Longitudinal Study reveals novel associations. Sci Rep 2017;7:11008.
[71] Reichert MC, Ripoll C, Casper M, Greinert R, Vandieken E, Grünhage F, et al. Common NOD2 Risk Variants as Major Susceptibility Factors for Bacterial Infections in Compensated Cirrhosis. Clin Transl Gastroenterol 2019;10:e00002.
[72] Seidelin A-S, Nordestgaard BG, Tybjærg-Hansen A, Stender S. Genetic variation at PPP1R3B increases hepatic CT attenuation and interacts with prandial status on plasma glucose. J Clin Endocrinol Metab 2020. https://doi.org/10.1210/clinem/dgaa151.
[73] Strnad P, Buch S, Hamesch K, Fischer J, Rosendahl J, Schmelz R, et al. Heterozygous carriage of the alpha1-antitrypsin Pi*Z variant increases the risk to develop liver cirrhosis. Gut 2018;Epub ahead:doi: 10.1136/gutjnl – 2018–316228.
[74] Tabassum R, Ramo JT, Ripatti P, Koskela JT, Kurki M, Karjalainen J, et al. Genetic architecture of human plasma lipidome and its link to cardiovascular disease. Nat Commun 2019;10:4329.
[75] Parisinos CA, Wilman HR, Thomas EL, Kelly M, Nicholls RC, McGonigle J, et al. Genome-wide and Mendelian randomisation studies of liver MRI yield insights into the pathogenesis of steatohepatitis. J Hepatol 2020. https://doi.org/10.1016/j.jhep.2020.03.032.
[76] Verma SS, Lucas AM, Lavage DR, Leader JB, Metpally R, Krishnamurthy S, et al. IDENTIFYING GENETIC ASSOCIATIONS WITH VARIABILITY IN METABOLIC HEALTH AND BLOOD COUNT LABORATORY VALUES: DIVING INTO THE QUANTITATIVE TRAITS BY LEVERAGING LONGITUDINAL DATA FROM AN EHR. Pac Symp Biocomput 2017;22:533–44.
[77] Viitasalo A, Eloranta A-M, Atalay M, Romeo S, Pihlajamaki J, Lakka TA. Association of MBOAT7 gene variant with plasma ALT levels in children: the PANIC study. Pediatr Res 2016;80:651–5.
[78] Wattacheril J, Lavine JE, Chalasani NP, Guo X, Kwon S, Schwimmer J, et al. Genome-Wide Associations Related to Hepatic Histology in Nonalcoholic Fatty Liver Disease in Hispanic Boys. J Pediatr 2017;190:100–7.e2.
[79] Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45:1274–83.
[80] Young KA, Palmer ND, Fingerlin TE, Langefeld CD, Norris JM, Wang N, et al. Genome-Wide Association Study Identifies Loci for Liver Enzyme Concentrations in Mexican Americans: The GUARDIAN Consortium. Obesity 2019;27:1331–7.
[81] Zusi C, Mantovani A, Olivieri F, Morandi A, Corradi M, Miraglia Del Giudice E, et al. Contribution of a genetic risk score to clinical prediction of hepatic steatosis in obese children and adolescents. Dig Liver Dis 2019;51:1586–92.
[82] Xia Y, Huang CX, Li GY, Chen KH, Han L, Tang L, et al. Meta-analysis of the association between MBOAT7 rs641738, TM6SF2 rs58542926 and nonalcoholic fatty liver disease susceptibility. Clin Res Hepatol Gastroenterol 2019:1–9.
[83] Emdin CA, Haas M, Khera AV, Aragam K, Chaffin M, Jian L, et al. A missense variant in Mitochondrial Amidoxime Reducing Component 1 gene and protection against liver disease. bioRxiv 2019:http://dx.doi.org/10.1101/594523.
[84] Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, et al. Mapping the genetic architecture of gene expression in human liver. PLoS Biol 2008;6:1020–32.
[85] Lee H-C, Inoue T, Sasaki J, Kubo T, Matsuda S, Nakasaki Y, et al. LPIAT1 regulates arachidonic acid content in phosphatidylinositol and is required for cortical lamination in mice. Mol Biol Cell 2012;23:4689–700.
[86] Shin S-Y, Fauman EB, Petersen A-K, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. Nat Genet

2014;46:543–50.
[87] Helsley RN, Varadharajan V, Brown AL, Gromovsky AD, Schugar RC, Ramachandiran I, et al. Obesity-linked suppression of membrane-bound O-acyltransferase 7 (MBOAT7) drives non-alcoholic fatty liver disease. Elife 2019;8. https://doi.org/10.7554/eLife.49882.
[88] Tanaka Y, Shimanaka Y, Caddeo A, Kubo T, Mao Y, Kubota T, et al. LPIAT1/MBOAT7 depletion increases triglyceride synthesis fueled by high phosphatidylinositol turnover. Gut 2020. https://doi.org/10.1136/gutjnl-2020-320646.
[89] Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet 2018;50:1505–13.
[90] Brouwers MCGJ, Simons N, Stehouwer CDA, Koek GH, Schaper NC, Isaacs A. Relationship Between Nonalcoholic Fatty Liver Disease Susceptibility Genes and Coronary Artery Disease. Hepatol Commun 2019;3:587–96.
[91] Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv 2019:531210.
[92] Mersha TB, Abebe T. Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. Hum Genomics 2015;9:1.
[93] Grimaudo S, Pipitone RM, Pennisi G, Celsa C, Cammà C, Di Marco V, et al. Association Between PNPLA3 rs738409 C>G Variant and Liver-Related Outcomes in Patients With Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2020;18:935–44.e3.