medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

rs641738C>T near *MBOAT7* promotes steatosis, NASH, fibrosis and hepatocellular carcinoma in non-alcoholic fatty liver disease: a meta-analysis

Kevin Teo[1], Kushala W. M. Abeysekera[2], Leon Adams[3,4], Elmar Aigner[5], Rajarshi Banerjee[6], Priyadarshi Basu[7], Thomas Berg[8], Pallav Bhatnagar[9], Stephan Buch[10], Ali Canbay[11], Sonia Caprio[12], Ankita Chatterjee[7], Yii-Der Ida Chen[13], Abhijit Chowdhury[14], Christian Datz[15], Dana de Gracia Hahn[1], Johanna K. DiStefano[16], Jiawen Dong[1], Amedine Duret[1], EU-PNAFLD Investigators*, Connor Emdin[17], Madison Fairey[1], Glenn S Gerhard[18], GOLD Consortium*, Xiuqing Guo[13], Jochen Hampe[10], Matthew Hickman[2], Lena Heintz19, Christian Hudert[20], Harriet Hunter[1], Matt Kelly[5], Julia Kozlitina[21], Marcin Krawczyk[19,22], Frank Lammert[19], Claudia Langenberg[23], Joel Lavine[24], Hong Kai Lim[1], Panu Luukkonen[25,26,27], Phillip E. Melton[28,29], Trevor A. Mori[3], Constantinos A. Parisinos[30], Sreekumar G Pillai[9], Faiza Qayyum[31], Matthias C. Reichert[19], Stefano Romeo[32,33,34], Jerome Rotter[13], Yu Ri Im[1], Nicola Santoro[12], Clemens Schafmayer[35], Elizabeth K. Speliotes[36,37], Stefan Stender[31], Felix Stickel[38], Christopher D. Still[39], Pavel Strnad[40], Kent D. Taylor[13], Anne Tybjærg-Hansen[31], Giuseppina Rosaria Umano[12,41], Mrudula Utukuri[1], Luca Valenti[42,43], Nicholas J. Wareham[23], Julia Wattacheril[44], Hannele Yki-Järvinen[25,26], Jake P. Mann[23]

Affiliations

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

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

3. Medical School, Faculty of Medical and Health Sciences, University of Western Australia, Perth, Australia
4. Department of Hepatology, Sir Charles Gairdner Hospital, Perth, WA, Australia
5. Department of Internal Medicine I, Paracelsus Medical University, Salzburg, Austria
6. Perspectum Diagnostics Ltd., Oxford, UK.
7. National Institute of Biomedical Genomics, Kalyani, India
8. Devision of Hepatolgy, Clinic and Polyclinic for Gastroenterology, Hepatology, Infectiology and Pneumology, University Clinic Leipzig
9. Eli Lilly and Company, Indianapolis IN.
10. Medical Department 1, University Hospital Dresden, Technische Universität Dresden (TU Dresden), Dresden, Germany
11. Gastroenterology, Hepatology and Infectiology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany
12. Yale University, Department of Pediatrics, New Haven, CT
13. Translational Genomics and Population Sciences Lundquist Institute; Harbor-UCLA Medical Center
14. Institute of Post Graduate Medical Education And Research, Kolkata, India
15. Department of Internal Medicine, General Hospital Oberndorf, Teaching Hospital of the Paracelsus Medical University Salzburg, Oberndorf, Austria
16. Diabetes and Fibrotic Disease Unit Translational Genomics Research Institute (TGen)
17. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Boston, MA
18. Department of Medical Genetics and Molecular Biochemistry, Lewis Katz School of Medicine at Temple University, Philadelphia, PA
19. Department of Medicine II, Saarland University Medical Center, Saarland University, Homburg, Germany
20. Center for Chronically Sick Children, Charité - Universitätsmedizin Berlin, Berlin, Germany
21. Eugene McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, TX, USA
22. 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
23. MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, UK.
24. Department of Pediatrics, Columbia University, New York, USA
25. Minerva Foundation Institute for Medical Research, Helsinki, Finland.
26. Department of Medicine, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
27. Yale University School of Medicine, New Haven, Connecticut, US.
28. 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
29. School of Pharmacy and Biomedical Sciences, Faculty of Health Sciences, Curtin University, Perth, WA, Australia
30. Institute of Health Informatics, Faculty of Population Health Sciences, University College London, London, UK
31. Department of Clinical Biochemisty, Rigshospitalet Copenhagen University Hospital Copenhagen, Denmark
32. Department of Molecular and Clinical Medicine, University of Gothenburg, Sweden
33. Cardiology Department, Sahlgrenska University Hospital, Gothenburg, Sweden
34. Clinical Nutrition Unit, Department of Medical and Surgical Sciences, University Magna Graecia, Catanzaro, Italy
35. Department of Visceral and Thoracic Surgery, Kiel University, Kiel, Germany Thomas Berg: Department of Internal Medicine, University Hospital Leipzig, Leipzig, Germany
36. Division of Gastroenterology and Hepatology, Department of Medicine, University of Michigan Health System, Ann Arbor MI.
37. Department of Computational Medicine and Bioinformatics University of Michigan Medical School Ann Arbor MI.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

38. Department of Gastroenterology and Hepatology, University Hospital of Zurich, Switzerland
39. Geisinger Obesity Institute, Danville, PA 17822, United States
40. Medical Clinic III, University Hospital RWTH Aachen, Aachen, Germany
41. Department of the Woman, the Child, of General and Specialized Surgery, University of Campania Luigi Vanvitelli, Naples Italy
42. Department of Pathophysiology and Transplantation, Università degli Studi di Milano
43. Translational Medicine, Department of Transfusion Medicine and Hematology, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano, Milan, Italy
44. 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

*A list of contributors is included in the Acknowledgements

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

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Footnote page

Contact information: Dr Jake P. Mann, University of Cambridge, Department of Paediatrics, Cambridge, United Kingdom. jm2032@cam.ac.uk, T: +44 1223 763480, F: +44 1223 336996

Abbreviations (in order of appearance):

NAFLD, non-alcoholic fatty liver disease; NAFL; non-alcoholic fatty liver; NASH, non-alcoholic steatohepatitis; HDL, high-density lipoprotein; HOMA-IR, homeostatic model assessment of insulin resistance; BMI, body mass index; ALT, alanine aminotransferase; CI, confidence interval; GGT, gamma-glutamyl transferase; OR, odds ratio; *PNPLA3*, patatin-like phospholipase domain containing protein 3; *TM6SF2*, transmembrane 6 superfamily member 2; *TMC4*, transmembrane channel-like 4; *MBOAT7*, membrane bound O-acyltransferase domain containing 7.

Financial support: JPM is supported by a Wellcome Trust Fellowship (216329/Z/19/Z) and a Children's Liver Disease Foundation grant. NIH grants: R01HD028016 (SC), R01DK111038 (SC), R01DK114504 (NS), DK091601 (JKD), UL1TR001105 (JK). 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]. LV was supported by MyFirst Grant AIRC

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

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: KWMA & MH are supported by grants from MRC and Alcohol Research UK (MR/L022206/1). 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).

ABSTRACT

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. Therefore, we aimed to establish whether rs641738 is a risk factor for NAFLD through meta-analysis. Data from 131,096 participants (7,692 with liver biopsies and 45,419 with imaging) was included in the meta-analysis. The minor T-allele of rs641738C>T was associated with higher liver fat on CT/MRI using an additive genetic model (+0.05 standard deviations [95% CI: 0.01 – 0.09], p=0.025), and with an increased risk of

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

NAFLD (per-allele OR: 1.09 [95% CI: 1.01 - 1.17]), nonalcoholic steatohepatitis (OR: 1.11 [95% CI: 1.02 - 1.21]), advanced fibrosis (OR: 1.14 [95% CI: 1.05 - 1.23]), and hepatocellular carcinoma (OR: 1.43 [95% CI: 1.22 - 1.67]) in adults with NAFLD. Sub-group analysis did not demonstrate a difference in Caucasians and non-Caucasians. Rs641738C>T was not associated with markers of insulin resistance but was associated with higher risk of stroke in the UK Biobank. These data validate rs641738C>T near *MBOAT7* as a risk factor for the development, activity, and stage of NAFLD including hepatocellular carcinoma.

Abstract word count: 192

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

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

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). 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 locus for alcohol-related cirrhosis(3). It has since been implicated in the pathogenesis of NAFLD(4), hepatocellular carcinoma(5), as well as in fibrosis development in chronic hepatitis B and C.

This SNV 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. *MBOAT7* encodes lysophosphatidylinositol acyltransferase 1 (LPIAT1), which contributes to the regulation of free arachidonic acid in cells(6). Rs641738C>T is associated with lower hepatic expression of *MBOAT7* at both the mRNA(7) and protein level(4). Given its role in inflammatory lipid pathways, most mechanistic work relating to rs641738 has focused on *MBOAT7*.

In NAFLD, the rs641738 variant was first demonstrated to be associated with increased hepatic fat content and severity of fibrosis in individuals of European descent(4). Proton magnetic resonance spectroscopy data from

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

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(8,9). However, these results were not replicated in adults of other genetic ancestries(4,10–12) or in children(13). It is recognised that investigation of candidate genes in relatively small cohorts can generate false-positive findings(14).

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), however, rare likely pathogenic variants in *MBOAT7* are associated with HCC in NAFLD(16). In summary, the association between rs641738 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 diabetes and cardiovascular outcomes, have not been assessed.

Here, we conducted a large meta-analysis of published and unpublished data to determine if rs641738 influences the development or stage of NAFLD and associated cardio-metabolic phenotypes.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

PATIENTS AND METHODS

Data sources and study selection

Two data sources were included in the meta-analysis: published studies (and abstracts) and unpublished GWAS (or targeted genotyping) data.

Published studies were sourced through Medline and Embase using the search terms "(*MBOAT7* or membrane-bound-o-acyltransferase) or (rs641738 or rs626283) or (*TMC4*)". 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 1st October 2019. Reference lists of relevant publications were also reviewed. Titles and abstracts were screened for eligibility independently by two authors, with inclusion/exclusion criteria applied to potentially eligible full texts.

A search was conducted for all GWAS in NAFLD, NASH, and steatosis. Authors of all potentially relevant GWAS were contacted to request extraction of data regarding rs641738C>T. To assess for cardiometabolic phenotype associations, Phenoscanner(17) and GeneATLAS(18) were searched for summary statistics from published GWAS.

HuGENet guidelines were followed throughout. This study was prospectively registered on PROSPERO Database of Systemic Reviews (CRD42018105507) Available from:

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

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

Inclusion and exclusion criteria

Cohort and case-control studies related to NAFLD were included if genotyping of rs641738C>T (or other SNVs in linkage disequilibrium [LD, R^2>0.8]) was conducted and data on one of the outcomes of interest were reported. Review articles, *in vitro* studies, and investigations involving animal, fish, and invertebrates were excluded. Studies which investigated liver disease of other etiologies were also excluded.

Data collection

For each study, data on participant demographics (sex, age, ethnicity) were collated. 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 (from CT, MRS, MRI, PDFF), serum lipid profile, fasting insulin, and alanine aminotransferase levels were collected as continuous variables. Hepatic fat content was also assessed using semi-quantitative scoring in the Fenland cohort, as previously described(19), and using CAP. Histology data were extracted according to the NASH Clinical Research Network scoring system and, where not otherwise diagnosed by a pathologist's assessment, NASH was defined using the Fatty Liver Inhibition of Progression (FLIP) algorithm. The above data were collected for each genotype separately (CC, CT, and TT).

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Phenoscanner was used to assess for disease associations with rs641738C>T at P<0.01. Phenotype associations were filtered for those related to cardio-metabolic and liver disease.

The authors of 21 published studies were contacted for additional data, all of whom replied. Several studies reported outcomes from overlapping cohorts: ref (4) and ref (20); ref (9) and ref (21). In these instances, data only from the larger of the overlapping cohorts were included in analyses.

Cohorts with genome-wide data

The authors of 9 potentially relevant GWAS (and cohort studies with genome-wide data) were contacted, of which 8 replied and data were included from 6. These cohorts have been described elsewhere(1,22–24). Densely imputed genotyping data were available for rs641738C>T in all with >0.98 call rate. Unpublished data from Wilman et al.(25) was extracted from the UK BioBank under Application ID 9914 ('Determining the Outcomes of People with Liver Disease').

Study quality assessment

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

Statistical Analysis:

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

For dichotomous outcomes, the effect statistic was calculated as an odds ratio between groups. Genetic association analyses were performed using an additive model to estimate the effect per T-allele as almost all included studies had used this model.

For analysis of effect on liver fat, 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 principal components of genetic ancestry (where available). In addition, data from the GOLD Consortium were adjusted for number of alcoholic drinks consumed.

Continuous quantitative liver fat data (from CT, MRI, MRS, or PDFF) and semi-quantitative data (ultrasound and CAP) were analyzed separately.

For other continuous variables, effect summary was calculated as a mean difference between CC and TT groups.

Meta-analysis was performed using random effects throughout.

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 with Caucasian (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, further sub-analysis by ethnicity may be affected by differences in linkage disequilibrium between genetic ancestries.

Leave-one-out sensitivity analysis was performed for all outcomes using additive model of inheritance and random effects.

Heterogeneity between groups was described using the Q statistic and I^2.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Bias was assessed using Egger's test and visually using funnel plots where more than 5 studies were included. P <0.025 (i.e. P<0.05/2) was considered statistically significant due to testing outcomes twice: in individuals of all ethnicities and Caucasians only. Analysis was performed using STATAv14 for Windows (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP), DistillerSR Forest Plot Generator from Evidence Partners (https://evidencepartners.com/resources/forest-plot-generator/), GraphPad Prism (v8.0 for Mac, GraphPad Software, La Jolla California, USA), and MetaGenyo(26).

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

RESULTS

Database search identified 405 abstracts, of which 18 studies were included. In addition, unpublished data were extracted from 12 cohorts (Table 1, Supplementary Fig. 1, and Supplementary Table 1).

Study Age group	Genetic ancestry (country)	Study design and sample size (N)	Female, n (%)	Features and patient characteristics	Liver biopsy (N)
Published					
Di Sessa, 2018; Paediatric (27)	Non-Finnish European (Italy)	Cases-only Hospital-based N=1002	466 (46.5%)	Children with hepatic steatosis measured by US	NA
Di Costanzo, 2018; Adult (28)	Non-Finnish European (Italy)	Case-control N=445	150 (33.7%)	Hepatic steatosis measured by US	NA
Dongiovanni, 2018b; Adult (29)	Mixed: Non-Finnish European and Finnish European	Cases-only N=1,388 (LBC)	728 (52.4%)	NAFLD diagnosed by LB (LBC)	1515
Lin, 2018; Paediatric (30)	East Asian (China)	Population-based N=831	257 (31.4%)	Hepatic steatosis measured by US	NA
Viitasalo, 2016; Paediatric (31)	Finnish European (Finland)	Population-based N=512	222 (47.5%)	Population cohort of children with measurement of ALT	NA
Koo, 2018; Adult (10)	East Asian (Korea)	Case-control Hospital-based N=525	264 (50.3%)	Adults with NAFLD diagnosed by LB, or US/MRI/CT	416
Published and unpublished data					
Hudert, 2018; Paediatric (13)	Non-Finnish European (Germany)	Case-control Hospital-based N=270	92 (34%)	Patients: children with NAFLD diagnosed by LB Controls: healthy population (adult) controls	70
Mann, 2018a; Adult (32)	Non-Finnish European (England)	Population cohort N=10,934	5,823 (53.2%)	Hepatic steatosis measured by US	NA

Mann, 2018b; Paediatric (33)	Non-Finnish European (Italy)	Hospital-based N=67	34 (50.7%)	Children with NAFLD diagnosed by LB	67
Umano, 2018; Paediatric (12)	Mixed: Non-Finnish European, African American, Hispanic (USA)	Cases-only Hospital-based N=860	509 (59.2%)	Hepatic steatosis measured by MRI	NA
Krawczyk, 2018; Adult (21)	Non-Finnish European (Germany)	Cases-only N=237	24 (38.1%)	Adults with NAFLD diagnosed by LB, or US/MRI/CT	63
Krawczyk, 2017; Adult (9)	Non-Finnish European (Germany)	Cases-only N=515	280 (54.4%)	Adults with NAFLD diagnosed by LB, or US/MRI/CT	320
Kawaguchi, 2018; Adult (34)	East Asian (Japan)	Case-control Mixed hospital- and population-based N=8,608	5111 (59.6%)	Patients: Adults with NAFLD diagnosed by LB Controls: healthy population controls	936
Dongiovanni, 2018; Adult (20)	Mixed: Non-Finnish European, African American, Hispanic (USA)	Population cohort: N= 4,570 (DHS), Cases-only Hospital-based: N=1,515 (LBC)	3,330 (54.7%)	Hepatic steatosis measured by H-MRS (DHS) or NAFLD diagnosed by LB (LBC)	1515
Mancina, 2016; Adult (4)	Mixed: Non-Finnish European,	Population cohort: N= 3,854 (DHS)	2754 (54.4%)	Hepatic steatosis measured by H-MRS (DHS) or NAFLD diagnosed by LB (LBC)	1149

	African American, Hispanic (USA)	Cases-only Hospital-based: N=1,149 (LBC)			
Luukkonen, 2016; Adult (8)	Finnish European (Finland)	Cases-control Hospital-based N=125	83 (66.4%)	Adults assessed for NAFLD by LB	125
Donati, 2017; Adult (5)	Non-Finnish European (Italy / UK)	Case-control Hospital-based N=765 (Italian) N=358 (UK NAFLD)	188 (24.6%) 143 (39.4%) 387 (34.5%)	Adults with NAFLD diagnosed by LB	1123
Sookoian, 2018; Adult (11)	Caucasian (Argentina)	Case-control Hospital-based N=634	360 (57.0%)	Patients: adults with NAFLD diagnosed by LB Controls: hepatic steatosis absent on US	372
Unpublished data					
Wilman, 2019; Adult (25)	Non-Finnish European (UK)	Population-based N=7,078	3,822 (54%)	GWAS of hepatic steatosis measured by MRI from the UK BioBank.	NA
DiStefano, 2015; Adult (22)	Non-Finnish European (USA)	Cases-only Hospital-based N=1,868	1,512 (80.9%)	GWAS of adults with NAFLD diagnosed by LB	1868
Adams, 2013; Paediatric (23)	Non-Finnish European (Australia)	Population-based N=928	444 (47.8%)	GWAS of adolescents with hepatic steatosis measured by US	NA
Lauridsen, 2018; Adult (35)	Non-Finnish European (Denmark)	Population-based N=7511	775 (53.9%)	Hepatic steatosis measured by CT, part of the Copenhagen General Population Study	NA

Luukkonen, 2018; Adult (36)	Finnish European (Finland)	Cases-only Hospital-based N= 38	21 (55%)	Hepatic steatosis measured by MRS	NA
Speliotes 2011; Adult (1)	Mixed: Non-Finnish European, African American, Hispanic (USA, Iceland, Europe)	Population-based N=4,244	-	GWAS of hepatic steatosis measured by CT	NA
Strnad, Buch, & Hamesch, 2018; Adult (37,38)	Non-Finnish European (Germany, Austria, & Switzerland)	Case-control Hospital-based N=1184	573 (48.4%)	Adults with NAFLD diagnosed by LB	672
Emdin, 2019; Adult (39)	Non-Finnish European (UK)	Population-based N=77,464	42,144 (54%)	Adults with coded diagnosis of NAFLD and/or cirrhosis	NA
Reichert, 2019; Adult (40)	Non-Finnish European (Germany)	Hospital-based N= 54	24 (42.1%)	Adults with NAFLD cirrhosis diagnosed by LB, or US/MRI/CT	NA
Guzman, 2018; Adult (41)	Mixed: Hispanic and non-Hispanic (USA)	Case-control N=246 (GLDI study) Case-control N=158 (GLDJ study)	104 (42.3%) 57 (36.1%)	Adults with Type 2 Diabetes with hepatic steatosis measured by MRI	NA

Wattacheril, 2017; Paediatric (42)	Hispanic (USA)	Cases-only N=208	0 (all male)	GWAS of Hispanic boys with NAFLD diagnosed by LB	208
Chatterjee, 2019; Adult (43)	South Asian (India)	Hospital-based N=354	138 (38.9%)	GWAS of adults with NAFLD diagnosed by LB or US	132

Table 1. **Characteristic of studies included in the meta-analysis**. CT, computerized tomography; 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.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

<Table 1.>

In total, 131,096 individuals (4,174 children) were included in the meta-analysis. Most studies were in adults (23/30, 77%) and in individuals of European ancestry (20/30, 67%). Of the 30 included studies, 15 (totaling 7,692 unique participants, hereof 345 children) reported data on liver histology.

Liver fat, NAFLD, and severe steatosis in adults

Seven studies (21,924 participants) reported data on hepatic fat as a continuous variable assayed by CT or MR(1,20,25,35,41,43). In meta-analysis across these seven studies, rs641738 was associated with increased liver fat, with a per T-allele increase of 0.05 (95% CI 0.01 - 0.09) standard deviations in inverse normalized liver fat (Figure 1). There was significant heterogeneity between studies with I^2 = 67% and Tau^2=.002. This trend of association remained on sub-analysis including only cohorts with Caucasian (European) ethnicity (Supplementary Figure 2).

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

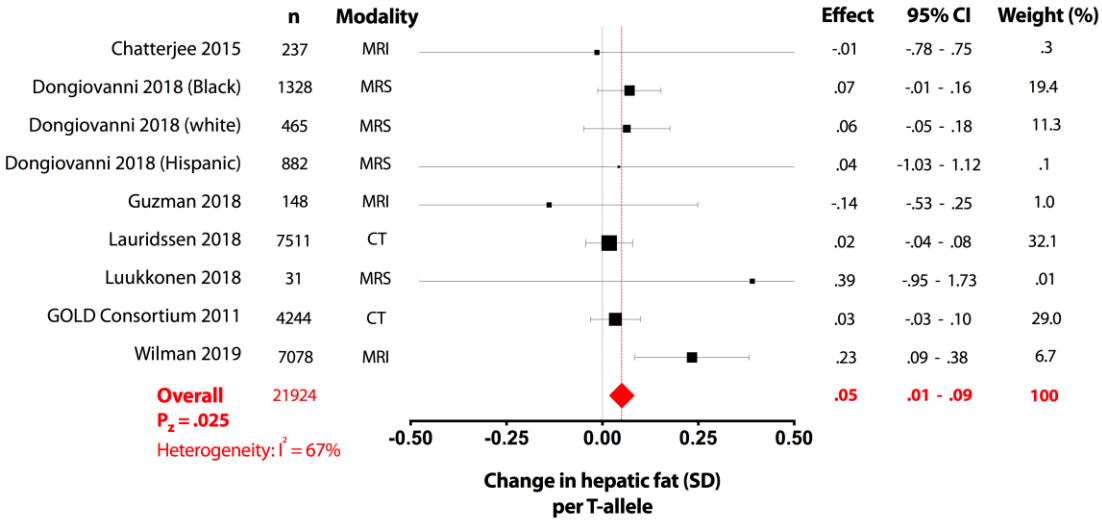

Figure 1. ***The effect of rs641738C>T on liver fat.*** *Data from 21,924 individuals with CT or MRI liver fat. T-allele was associated with a small increase in liver fat, where data represents standard deviation change in normalized liver fat per T-allele. CGPS, Copenhagen General Population Study; CI, confidence interval; ES, effect summary; GOLD, Genetics of Liver Disease; N, number of individuals included.*

The rs641738 variant C>T was also associated with NAFLD as a trait (OR 1.09 (95% CI 1.01, 1.17) using an additive model of inheritance (Figure 2A). Sensitivity analysis using the leave-one-out method did not demonstrate any individual study to affect the estimate (Supplementary Figure 3) and there was no evidence of study distribution bias on funnel plot (Supplementary Figure 4). The trend of a positive association was seen on sub-analysis in Caucasians (OR 1.12 (95% CI 0.997, 1.26), Supplementary Figure 5A).

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

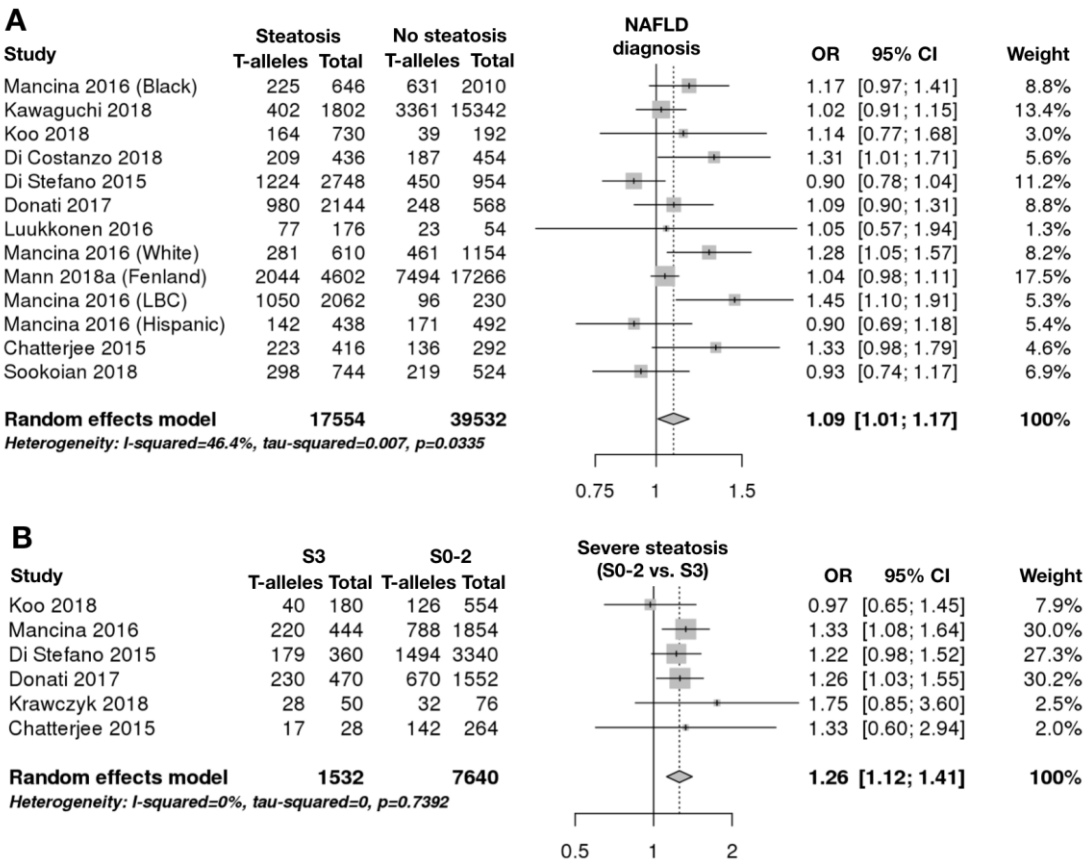

Figure 2. ***rs641738C>T is associated with higher odds of diagnosis of NAFLD and histological severity of steatosis****. Data from 28,543 adults with radiologically defined steatosis for presence versus absence of NAFLD (2A), and from 4,572 adults with liver biopsy data for presence of severe steatosis (S0-2 versus S3, 2B) using an additive model of inheritance.*

In patients with NAFLD, rs641738C>T was associated with the presence of severe steatosis (S0-2 vs. S3) on liver biopsy (OR 1.26 [95% CI 1.12, 1.41], Figure 2B). This association remained on sub-analysis in Caucasian individuals (OR 1.28 [95% CI 1.14, 1.45], Supplementary Figure 5B). Similar results were observed using CAP and semi-quantitative ultrasound to assess

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

steatosis severity (β .03 (95% CI .001, .06) standard deviations of inverse normalized liver fat score per T-allele, Supplementary Figure 6).

Histological NASH in adults

Data from 9 studies (6,155 participants) showed that rs641738C>T was positively associated with the presence of NASH on biopsy in adults (OR 1.11 (95% 1.02, 1.21, Figure 3). A similar magnitude of effect was observed on sub-analysis in Caucasian individuals (OR 1.13 (95% 1.01, 1.27, Supplementary Figure 7).

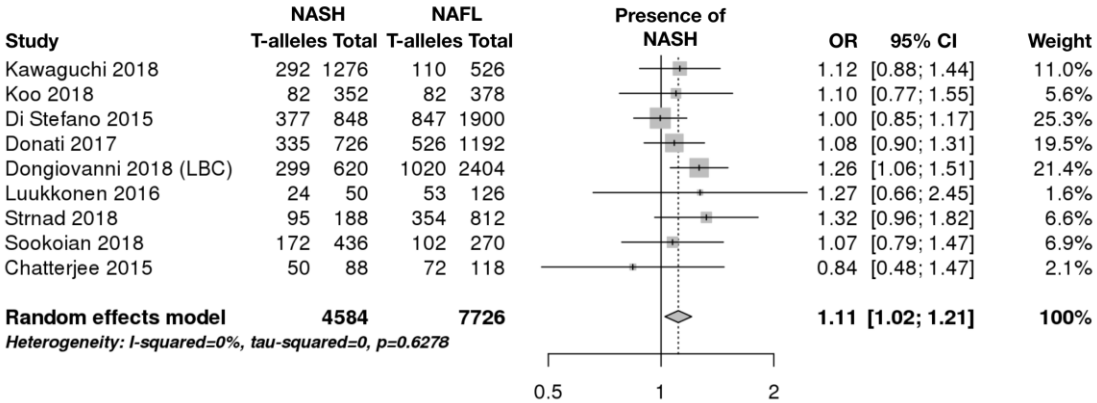

Figure 3. ***rs641738C>T is associated with higher odds of NASH on biopsy****. Data from 6,155 adults with NASH defined according to the FLIP algorithm for NAFL versus NASH, using an additive model of inheritance. NAFL, non-alcoholic fatty liver.*

Fibrosis in adults

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Data from 8 studies (82,857 adults, 6,787 with liver biopsy data) were included in meta-analysis of fibrosis. Our primary outcome, presence of advanced fibrosis in adults (stage F0-2 versus stage F3-4), was positively associated with T-allele (OR 1.14 (95% 1.05, 1.23), Figure 4A) in adults. Sensitivity analysis, including omission of coded cirrhosis data from Emdin *et al.*(39), did not alter the effect summary (Supplementary figure 8). Presence of any fibrosis (stage 0 versus stage 1-4) was also positively associated with rs641738C>T (OR 1.14 (95% 1.01, 1.28), Figure 4B). On sub-analysis of Caucasian individuals, rs641738C>T was associated with advanced fibrosis (OR 1.16 (95% 1.06, 1.26)) but not with any fibrosis (OR 1.15 (95% 0.99, 1.34)) despite a positive trend (Supplementary figure 9).

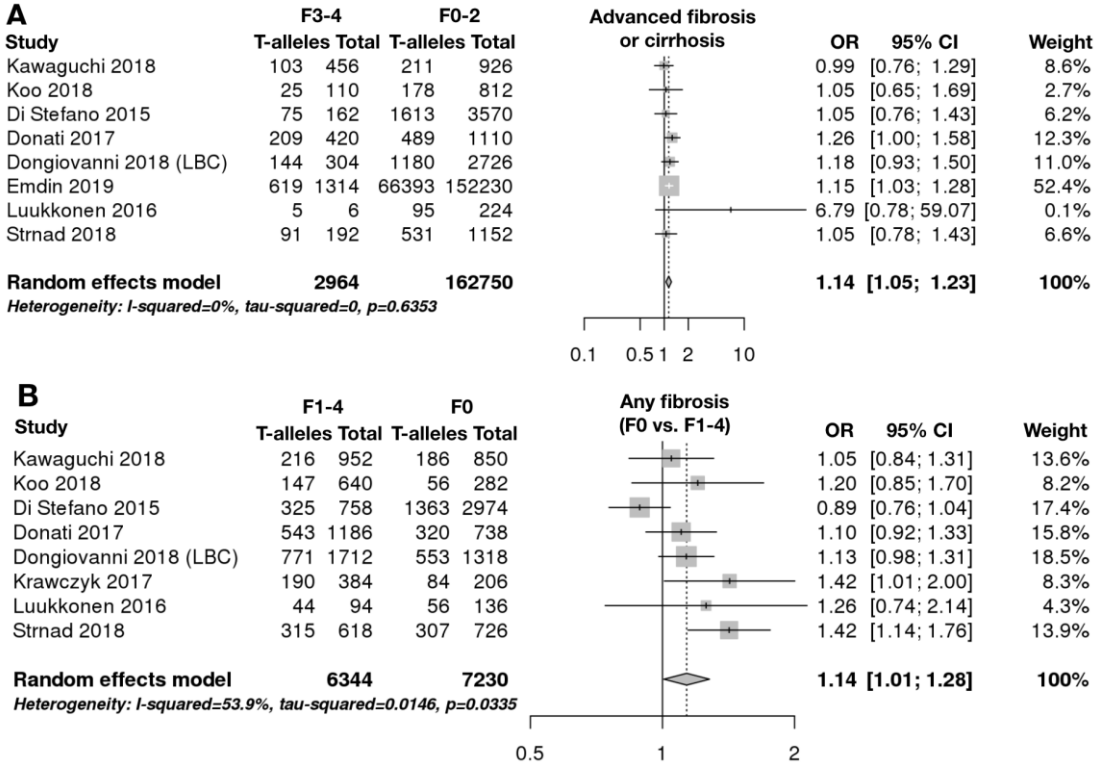

Figure 4. ***rs641738C>T is associated with increased fibrosis in NAFLD****. A, data from 6,787 adults with biopsy-proven NAFLD (plus coding data from*

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Emdin et al.) comparing advanced fibrosis (F3-4) versus F0-2, using an additive model of inheritance. B, data from 6,787 adults with biopsy-proven NAFLD comparing any fibrosis (F1-4) versus no fibrosis F0.

Development of hepatocellular carcinoma

Five cohorts (3,803 participants, 360 cases of NAFLD-HCC) reported on development of HCC in patients with NAFLD. Presence of T-allele was associated with increased odds of HCC in NAFLD (OR 1.43 (95% CI 1.22, 1.67, Figure 5).

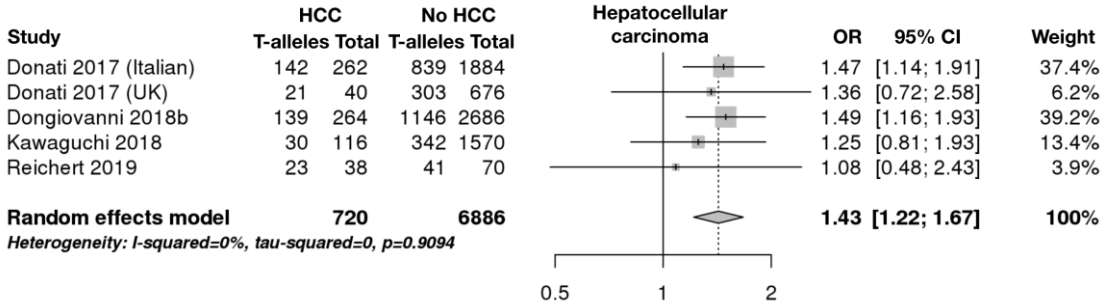

Figure 5. ***rs641738C>T is associated with higher odds of NAFLD-HCC****. Data from 3,803 adults with NAFLD assessing for the presence versus absence of HCC, using an additive model of inheritance.*

Effect on aminotransferases, lipids, and fasting insulin

Data from 12 studies (17,148 participants) was available for meta-analysis of serum biochemical parameters. T-allele was associated with lower triglycerides (mean difference CC versus TT genotype -3.7 mg/dL (95% CI -

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

7.2, -0.2)) but no other effect on aminotransferases, serum lipids, or fasting insulin (Table 2).

	Number of cohorts	P_Q	I^2	Random effects		
				Mean difference	[95% CI]	P_z
ALT, IU/L (n=17,102)	14	.44	.01	-.27	-.91, 0.37	.41
Triglycerides, mg/dL (n=17,148)	15	.19	.23	-3.71	-7.22, -.20	.04
Total cholesterol, mg/dL (n=16,822)	13	.009	.52	.45	-2.41, 3.31	.76
High-density lipoprotein, mg/dL (n=9,843)	10	.08	.39	-.26	-1.31, .80	.63
Low density lipoprotein, mg/dL (n=8,800)	7	.18	.30	2.06	-.30, 4.42	.09
Fasting insulin, mU/L (n=6,269)	4	.004	.71	-.66	-2.52, 1.20	.49

Table 2. **Meta-analysis for the effect of rs641738C>T on biochemical indices liver damage dyslipidemia, and insulin resistance**. Data represents the mean difference between CC and TT genotypes using random effects. N represents the sum of individuals with CC and TT genotypes included in each analysis.

Disease outcomes in adults

Using data from previous meta-analyses via Phenoscanner and UK BioBank data via GeneAtlas, rs641738C>T was weakly positively associated stroke (β 0.0007, p=0.004), Supplementary table 2). There was no evidence of an association with type 2 diabetes, coronary artery disease, or chronic kidney disease. It was also associated with higher alkaline phosphatase (β 0.005, $p=6.1x10^{-6}$).

Effect of rs641738C>T on paediatric NAFLD

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Data from seven studies (4,174 children) was used in the meta-analysis. rs641738C>T was not significantly associated with any disease outcome studied (Supplementary table 3 and Supplementary Figure 10). However there was a trend towards increasing hepatic fat fraction (0.19 SD (95% CI -0.05, 0.42)) and severity of steatosis (OR 1.21 (95% 0.89, 1.64)).

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

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 the importance of rs641738C>T near *MBOAT7* on the full spectrum of NAFLD in adults.

A two-stage GWAS initially identified rs641738C>T as a genome-wide significant locus for alcohol-related cirrhosis(3). *MBOAT7* was a potentially interesting target as an enzyme involved in (phosphor)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(4,8). However, these associations have not been consistently replicated in the literature(11). We conducted a meta-analysis to firmly establish the association of rs641738C>T with NAFLD.

Main findings

We found that the T-allele of rs641738C>T was associated with higher liver fat content, and with an increased risk of NASH, fibrosis, and HCC. 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(2). Also, unlike NASH-associated variants in *PNPLA3*, *HSD17B13*, *MARC1*, and *TM6SF2*, there was no association between this *MBOAT7* variant and alanine or aspartate aminotransferase. The marginal positive

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

effect of this variant on hepatic triglyceride content may suggest alterations in the composition of hepatic lipid, rather than quantity(8). 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(20). 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 rs641738 is associated with changes in hepatic expression of *MBOAT7*. Results from the GTEx Consortium show a strong negative association with T-allele(7), which is supported by data from Schadt et al.(44). *MBOAT7* protein expression correlated with mRNA in liver biopsies from Mancina *et al.*(4) but this finding was not replicated by Sookoian *et al.*(11). *MBOAT7* encodes LPIAT1, a 6 transmembrane domain protein involved in acyl-chain remodeling of membranes that influence intracellular membrane composition and circulating phosphatidylinositols(8). Further recent metabolite profiling data implicates *MBOAT7* as the causal gene for this SNV(32). Moreover, TMC4 was found with a low expression in the liver(4) that is consistent with no mechanistic data supporting its role in NAFLD.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

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(45). In addition, loss of *MBOAT7*, but not TMC4, increases the severity of NAFLD in mice fed a high-fat diet(46). It is not known whether these genetically modified animals have increased susceptibility to HCC.

This variant shows a particularly strong association with development of HCC. It is unclear whether this reflects the effect on NASH-fibrosis or if there is a specific causal relationship between *MBOAT7* and HCC.

We found no evidence of rs641738 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(47) (albeit with very modest effect size as compared to their effects on liver disease) and a Mendelian randomization study indicates a causal role in determining insulin resistance mediated by the degree of liver damage(20). Similarly, these two variants are associated with reduced risk of coronary artery disease whereas rs641738 has no effect. It does, however, appear to be weakly associated with higher prevalence of stroke in the UK BioBank(18). Our analysis also found lower serum triglycerides in those with TT-genotype versus CC-genotype, though this was not replicated in the Global Lipid Genetic Consortium data(48).

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

There is some evidence that genetic variants affect response to drug treatment (for *PNPLA3*(49)) but this is yet to be explored for *MBOAT7*. It will be equally interesting to assess whether somatic genotype of variants associated with HCC affects treatment response.

A strength of this analysis is the inclusion of data from individuals of multiple ethnicities (and genetic ancestries). We found no difference in the impact of the variant on liver disease among Caucasian and non-Caucasian individuals. Another strength is the large number of individuals with liver biopsy-derived phenotypic data.

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(50).

Studies measured hepatic fat using several different imaging modalities, which have varying sensitivity for quantification of liver fat. This may have accounted for some of the heterogeneity observed in these analyses. There was a trend towards more positive associations in population-based studies using more sensitive techniques (MRI or MRS). It is possible that weighting towards large CT-based studies could have underestimated the true effect size.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

We found significant differences between adult and paediatric histological analyses though a consistent trend was observed in the analysis of hepatic fat fraction. There are several potential reasons, including: sample size insufficient to demonstrate an effect, variations in imaging quantification of fat, too few clinical events (i.e. with fibrosis) to demonstrate an effect, different histology of paediatric NASH, or a true alternative effect of this variant on paediatric NAFLD.

Though there was minimal heterogeneity across included studies, the numbers of individuals with NAFLD and HCC were comparatively low. Further work in this area may improve the accuracy of effect estimates.

Conclusions

rs641738C>T near *MBOAT7* increases risk of NASH, fibrosis, and HCC in NAFLD with a small, positive effect on total liver fat and no impact on insulin resistance. 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

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

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). This study has been conducting 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.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

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. Eslam M, Valenti L, Romeo S. Genetics and epigenetics of NAFLD and NASH: Clinical impact. J. Hepatol. [Internet]. 2018;68:268–279. Available from: https://doi.org/10.1016/j.jhep.2017.09.003
3. 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. [Internet]. 2015;47:1443–1448. Available from: http://dx.doi.org/10.1038/ng.3417
4. 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;150:1219-1230e6.
5. 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. [Internet]. 2017;7:4492. Available from: http://www.nature.com/articles/s41598-017-04991-0
6. Gijón MA, Riekhof WR, Zarini S, Murphy RC, Voelker DR. Lysophospholipid acyltransferases and arachidonate recycling in human neutrophils. J. Biol. Chem. 2008;283:30235–30245.
7. Consortium TGte. The Genotype-Tissue Expression (GTEx) project.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Nat. Genet. 2013;45:580–585.

8. 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. [Internet]. 2016;65:1263–1265. Available from: http://www.sciencedirect.com/science/article/pii/S0168827816304214

9. 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. [Internet]. 2017;58:247–255. Available from: http://www.jlr.org/lookup/doi/10.1194/jlr.P067454

10. 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;33:1277–1285.

11. 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.

12. 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. [Internet]. 2018;113:376–383. Available from:

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

http://dx.doi.org/10.1038/ajg.2018.1

13. Hudert CA, Selinski S, Rudolph B, Bläker H, Christoph, Loddenkemper, et al. Genetic determinants of steatosis and fibrosis progression in pediatric non-alcoholic fatty liver disease. Liver Int. 2018;In Press:doi: 10.1111/liv.14006.

14. Border R, Johnson EC, Evans LM, Smolen A, Berley N, Sullivan PF, et al. No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples. Am. J. Psychiatry. 2019;doi.org/10.1176/appi.ajp.2018.18070881.

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;99:912–916.

16. 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;9:1–10.

17. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: A database of human genotype-phenotype associations. Bioinformatics. 2016;32:3207–3209.

18. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. Nat. Genet. [Internet]. 2018;50:1593–1599. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30349118

19. De Lucia Rolfe E, Brage S, Sleigh A, Finucane F, Griffin SJ, Wareham NJ, et al. Validity of ultrasonography to assess hepatic steatosis

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

compared to magnetic resonance spectroscopy as a criterion method in older adults. PLoS One. 2018;13:87–99.

20. 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;283:356–370.

21. 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. [Internet]. 2018;63:621–626. Available from: http://dx.doi.org/10.1038/s10038-018-0420-4

22. 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;52:373–382.

23. 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.

24. Stender S, Smagris E, Lauridsen BK, Kofoed KF, Nordestgaard BG, Tybjærg-Hansen A, et al. Relationship between genetic variation at PPP1R3B and levels of liver glycogen and triglyceride. Hepatology. 2018;67:2182–2195.

25. Wilman H, Parisinos C, Kelly M, Neubauer S, Thomas L, Bell J, et al. Genome-wide association studies of abdominal MRI scans identifies loci associated with liver fat and liver iron in the UK Biobank. J. Hepatol. 2019;70:e135.

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

26. Martorell-Marugan J, Toro-Dominguez D, Alarcon-Riquelme ME, Carmona-Saez P. MetaGenyo: A web tool for meta-analysis of genetic association studies. BMC Bioinformatics. 2017;18:1–6.

27. 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.

28. 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.

29. Dongiovanni P, Meroni M, Mancina RM, Baselli G, Rametta R, Pelusi S, et al. Protein phosphatase 1 regulatory subunit 3B gene variation protects against hepatic fat accumulation and fibrosis in individuals at high risk of nonalcoholic fatty liver disease. Hepatol. Commun. 2018;2:666–675.

30. 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;38:1300–1307.

31. 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;80:651–655.

32. Mann JP, Pietzner M, Wittemans LB, Rolfe EDL, Kerrison N, Allison ME, et al. Metabolomic patterns associated with known genetic variants for hepatic steatosis and non-alcoholic steatohepatitisidentify

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

biomarkers that may be of utility in predicting adverse liver outcomes. J Hepatol. 2018;68:S331–S332.

33. 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.

34. 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.

35. Lauridsen BK, Stender S, Kristensen TS, Kofoed KF, Køber L, Nordestgaard BG, et al. Liver fat content, non-alcoholic fatty liver disease, and ischaemic heart disease: Mendelian randomization and meta-analysis of 279 013 individuals. Eur. Heart J. 2018;39:385–393.

36. Luukkonen PK, Sädevirta S, Zhou Y, Kayser B, Ali A, Ahonen L, et al. Saturated fat is more metabolically harmful for the human liver than unsaturated fat or simple sugars. Diabetes Care. 2018;41:1732–1739.

37. 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 [Internet]. 2018;Epub ahead:doi: 10.1136/gutjnl-2018-316228. Available from: http://gut.bmj.com/lookup/doi/10.1136/gutjnl-2018-316228

38. von Schönfels W, Beckmann JH, Ahrens M, Hendricks A, Röcken C, Szymczak S, et al. Histologic improvement of NAFLD in patients with obesity after bariatric surgery based on standardized NAS (NAFLD activity score). Surg. Obes. Relat. Dis. 2018;14:1607–1616.

39. Emdin CA, Haas M, Khera A V, Aragam K, Chaffin M, Jian L, et al. A

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

missense variant in Mitochondrial Amidoxime Reducing Component 1 gene and protection against liver disease. bioRxiv. 2019;http://dx.doi.org/10.1101/594523.

40. Reichert M, Ripoll C, Casper M, Horn P, Bruns T, Grünhage F, et al. Increased prevalence of low-frequency and rare NOD2 variants in patients with liver cirrhosis. J. Hepatol. 2019;70:e442.

41. 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 . Hepatol. Commun. 2018;2:561–570.

42. 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. [Internet]. 2017;190:100-107.e2. Available from: https://doi.org/10.1016/j.jpeds.2017.08.004

43. 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.

44. 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;6:1020–1032.

45. Lee H-C, Inoue T, Sasaki J, Kubo T, Matsuda S, Nakasaki Y, et al. LPIAT1 regulates arachidonic acid content in phosphatidylinositol and is

medRxiv preprint doi: https://doi.org/10.1101/19013623; this version posted December 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

required for cortical lamination in mice. Mol. Biol. Cell. 2012;23:4689–4700.

46. 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:e49882.

47. 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. [Internet]. 2018;50:1505–1513. Available from: http://www.nature.com/articles/s41588-018-0241-6

48. 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;45:1274–1285.

49. Scorletti E, West AL, Bhatia L, Hoile SP, McCormick KG, Burdge GC, et al. Treating liver fat and serum triglyceride levels in NAFLD, effects of PNPLA3 and TM6SF2 genotypes: Results from the WELCOME trial. J. Hepatol. [Internet]. 2015;63:1476–1483. Available from: http://dx.doi.org/10.1016/j.jhep.2015.07.036

50. 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 [Internet]. 2019;531210. Available from: https://www.biorxiv.org/content/10.1101/531210v1