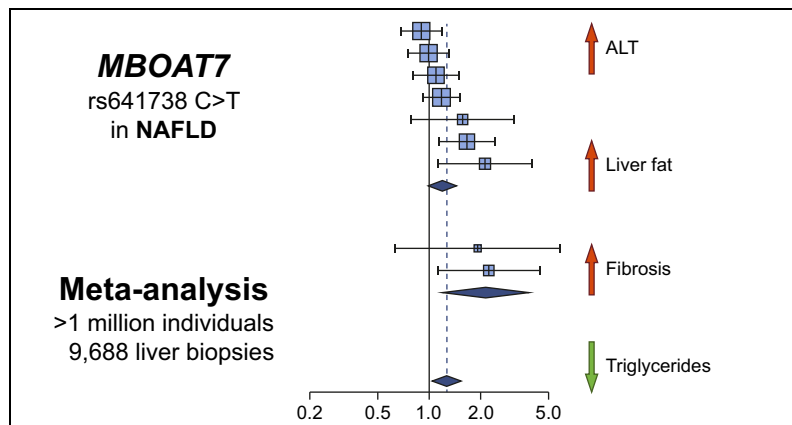


# rs641738C>T near *MBOAT7* is associated with liver fat, ALT and fibrosis in NAFLD: A meta-analysis

## Graphical abstract



## Authors

Kevin Teo, Kushala W.M. Abeysekera, Leon Adams, ..., Hannele Yki-Järvinen, Kendra A. Young, Jake P. Mann

## Correspondence

[jm2032@cam.ac.uk](mailto:jm2032@cam.ac.uk) (J.P. Mann).

## Lay summary

Fatty liver disease is a common condition where fat builds up in the liver, which can cause liver inflammation and scarring (including 'cirrhosis'). It is closely linked to obesity and diabetes, but some genes are also thought to be important. We did this study to see whether one specific change ('variant') in one gene ('*MBOAT7*') was linked to fatty liver disease. We took data from over 40 published studies and found that this variant near *MBOAT7* is linked to more severe fatty liver disease. This means that drugs designed to work on *MBOAT7* could be useful for treating fatty liver disease.

## Highlights

- Meta-analysis of 42 studies (>1 million participants) to assess the role of rs641738C>T near *MBOAT7* in NAFLD.
- rs641738C>T positively associated with liver fat, ALT, fibrosis and HCC.
- rs641738C>T negatively associated with serum triglycerides.
- Consistent associations found in studies of Caucasian populations only.



## rs641738C>T near *MBOAT7* is associated with liver fat, ALT and fibrosis in NAFLD: A meta-analysis

Kevin Teo<sup>1</sup>, Kushala W.M. Abeysekera<sup>2</sup>, Leon Adams<sup>3,4</sup>, Elmar Aigner<sup>5</sup>, Quentin M. Anstee<sup>6,7</sup>, Jesus M. Banales<sup>8</sup>, Rajarshi Banerjee<sup>9</sup>, Priyadarshi Basu<sup>10</sup>, Thomas Berg<sup>11</sup>, Pallav Bhatnagar<sup>12</sup>, Stephan Buch<sup>13</sup>, Ali Canbay<sup>14</sup>, Sonia Caprio<sup>15</sup>, Ankita Chatterjee<sup>10</sup>, Yii-Der Ida Chen<sup>16</sup>, Abhijit Chowdhury<sup>17</sup>, Ann K. Daly<sup>6</sup>, Christian Datz<sup>18</sup>, Dana de Gracia Hahn<sup>1</sup>, Johanna K. DiStefano<sup>19</sup>, Jiawen Dong<sup>1</sup>, Amedine Duret<sup>1</sup>, EU-PNAFLD Investigators<sup>†</sup>, Connor Emdin<sup>20</sup>, Madison Fairey<sup>1</sup>, Glenn S. Gerhard<sup>21</sup>, GOLD Consortium<sup>†</sup>, Xiuqing Guo<sup>16</sup>, Jochen Hampe<sup>13</sup>, Matthew Hickman<sup>2</sup>, Lena Heintz<sup>22</sup>, Christian Hudert<sup>23</sup>, Harriet Hunter<sup>1</sup>, Matt Kelly<sup>9</sup>, Julia Kozlitina<sup>24</sup>, Marcin Krawczyk<sup>22,25</sup>, Frank Lammert<sup>22</sup>, Claudia Langenberg<sup>26</sup>, Joel Lavine<sup>27</sup>, Lin Li<sup>28</sup>, Hong Kai Lim<sup>1</sup>, Rohit Loomba<sup>29</sup>, Panu K. Luukkonen<sup>30,31,32</sup>, Phillip E. Melton<sup>33,34,35</sup>, Trevor A. Mori<sup>3</sup>, Nicholette D. Palmer<sup>36</sup>, Constantinos A. Parisinos<sup>37</sup>, Sreekumar G. Pillai<sup>12</sup>, Faiza Qayyum<sup>38</sup>, Matthias C. Reichert<sup>22</sup>, Stefano Romeo<sup>39,40,41</sup>, Jerome I. Rotter<sup>16</sup>, Yu Ri Im<sup>1</sup>, Nicola Santoro<sup>15,42</sup>, Clemens Schafmayer<sup>43</sup>, Elizabeth K. Speliotes<sup>44,45</sup>, Stefan Stender<sup>38</sup>, Felix Stickel<sup>46</sup>, Christopher D. Still<sup>47</sup>, Pavel Strnad<sup>48</sup>, Kent D. Taylor<sup>16</sup>, Anne Tybjærg-Hansen<sup>38</sup>, Giuseppina Rosaria Umamo<sup>15,49</sup>, Mrudula Utukuri<sup>1</sup>, Luca Valenti<sup>50,51</sup>, Lynne E. Wagenknecht<sup>52</sup>, Nicholas J. Wareham<sup>26</sup>, Richard M. Watanabe<sup>53</sup>, Julia Wattacheril<sup>54</sup>, Hanieh Yaghoobkar<sup>55</sup>, Hannele Yki-Järvinen<sup>30,31</sup>, Kendra A. Young<sup>56</sup>, Jake P. Mann<sup>26,\*</sup>

<sup>1</sup>School of Clinical Medicine, University of Cambridge, Cambridge, UK; <sup>2</sup>MRC Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, UK; <sup>3</sup>Medical School, Faculty of Health and Medical Sciences, University of Western Australia, Perth, WA, Australia; <sup>4</sup>Department of Hepatology, Sir Charles Gairdner Hospital, Perth, WA, Australia; <sup>5</sup>First Department of Medicine, Paracelsus Medical University Salzburg, Austria; <sup>6</sup>Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK; <sup>7</sup>Newcastle NIHR Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK; <sup>8</sup>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; <sup>9</sup>Perspectum Ltd, Oxford, UK; <sup>10</sup>National Institute of Biomedical Genomics, Kalyani, India; <sup>11</sup>Division of Hepatology, Department of Medicine II, Leipzig University Medical Center, Leipzig, Germany; <sup>12</sup>Eli Lilly and Company, Indianapolis, IN, USA; <sup>13</sup>Medical Department 1, University Hospital Dresden, Technische Universität Dresden (TU Dresden), Dresden, Germany; <sup>14</sup>Gastroenterology, Hepatology and Infectiology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany; <sup>15</sup>Yale University, Department of Pediatrics, New Haven, CT, USA; <sup>16</sup>The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA, USA; <sup>17</sup>Institute of Post Graduate Medical Education and Research, Kolkata, India; <sup>18</sup>Department of Internal Medicine, General Hospital Oberndorf, Teaching Hospital of the Paracelsus Medical University Salzburg, Oberndorf, Austria; <sup>19</sup>Diabetes and Fibrotic Disease Unit Translational Genomics Research Institute (TGen), Phoenix, AZ, USA; <sup>20</sup>Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Boston, MA, USA; <sup>21</sup>Department of Medical Genetics and Molecular Biochemistry, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA; <sup>22</sup>Department of Medicine II, Saarland University Medical Center, Saarland University, Homburg, Germany; <sup>23</sup>Department of Pediatric Gastroenterology, Charité - Universitätsmedizin Berlin, Berlin, Germany; <sup>24</sup>Eugene McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, TX, USA; <sup>25</sup>Laboratory of Metabolic Liver Diseases, Department of General, Transplant and Liver Surgery, Centre for Preclinical Research, Medical University of Warsaw, Warsaw, Poland; <sup>26</sup>MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, Cambridge, UK; <sup>27</sup>Department of Pediatrics, Columbia University, New York, NY, USA; <sup>28</sup>BioStat Solutions LLC, Frederick, MD, USA; <sup>29</sup>NAFLD Research Center, Division of Gastroenterology and Epidemiology, University of California at San Diego, La Jolla, CA, USA; <sup>30</sup>Minerva Foundation Institute for Medical Research, Helsinki, Finland; <sup>31</sup>Department of Medicine, University of Helsinki and Helsinki University Hospital, Helsinki, Finland; <sup>32</sup>Yale University School of Medicine, New Haven, CT, USA; <sup>33</sup>School of Global Population Health, Faculty of Health and Medical Sciences, The University of Western Australia, Perth, WA, Australia; <sup>34</sup>School of Pharmacy and Biomedical Sciences, Faculty of Health Sciences, Curtin University, Perth, WA, Australia; <sup>35</sup>Menzies Institute for Medical Research, College of Health and Medicine, University of Tasmania, Hobart, Australia; <sup>36</sup>Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC, USA; <sup>37</sup>Institute of Health Informatics, Faculty of Population Health Sciences, University College London, London, UK; <sup>38</sup>Department

Keywords: *MBOAT7*; Fibrosis; NAFLD; Triglyceride; Diabetes; ALSPAC.  
Received 21 April 2020; received in revised form 29 July 2020; accepted 20 August 2020;  
available online 31 August 2020

\* Corresponding author. Address: University of Cambridge, Institute of Metabolic Science, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 0QQ, UK. Tel: +44 1223 336792; fax: +44 1223 330598.

E-mail address: [jm2032@cam.ac.uk](mailto:jm2032@cam.ac.uk) (J.P. Mann).

† A list of contributors is included in the Acknowledgements.

<https://doi.org/10.1016/j.jhep.2020.08.027>



of Clinical Biochemistry, Rigshospitalet Copenhagen University Hospital, Copenhagen, Denmark; <sup>39</sup>Department of Molecular and Clinical Medicine, University of Gothenburg, Gothenburg, Sweden; <sup>40</sup>Cardiology Department, Sahlgrenska University Hospital, Gothenburg, Sweden; <sup>41</sup>Clinical Nutrition Unit, Department of Medical and Surgical Sciences, University Magna Graecia, Catanzaro, Italy; <sup>42</sup>Department of Medicine and Health Sciences 'V. Tiberio' University of Molise, Campobasso, Italy; <sup>43</sup>Department of Visceral and Thoracic Surgery, Kiel University, Kiel, Germany; <sup>44</sup>Division of Gastroenterology and Hepatology, Department of Medicine, University of Michigan Health System Ann Arbor, MI, USA; <sup>45</sup>Department of Computational Medicine and Bioinformatics, University of Michigan Medical School Ann Arbor, MI, USA; <sup>46</sup>Department of Gastroenterology and Hepatology, University Hospital of Zurich, Zurich, Switzerland; <sup>47</sup>Geisinger Obesity Institute, Danville, PA, USA; <sup>48</sup>Medical Clinic III, University Hospital RWTH Aachen, Aachen, Germany; <sup>49</sup>Department of the Woman, the Child, of General and Specialized Surgery, University of Campania Luigi Vanvitelli, Naples, Italy; <sup>50</sup>Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy; <sup>51</sup>Translational Medicine, Department of Transfusion Medicine and Hematology, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano, Milan, Italy; <sup>52</sup>Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, US; <sup>53</sup>Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; <sup>54</sup>Department of Medicine, Center for Liver Disease and Transplantation, Columbia University College of Physicians and Surgeons, New York Presbyterian Hospital, New York, NY, USA; <sup>55</sup>Genetics of Complex Traits, College of Medicine and Health, University of Exeter, Exeter, UK; <sup>56</sup>Department of Epidemiology, Colorado School of Public Health, University of Colorado Denver, Aurora, CO, USA

**Background & Aims:** A common genetic variant near *MBOAT7* (rs641738C>T) has been previously associated with hepatic fat and advanced histology in 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 to characterise its role in the regulation of related metabolic phenotypes through a meta-analysis.

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

**Results:** Data from 1,066,175 participants (9,688 with liver biopsies) across 42 studies were 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],  $p_z = 4.8 \times 10^{-5}$ ) and diagnosis of NAFLD (odds ratio [OR] 1.17 [95% CI 1.05–1.3],  $p_z = 0.003$ ) in Caucasian adults. The variant was also positively associated with presence of advanced fibrosis (OR 1.22 [95% CI 1.03–1.45],  $p_z = 0.021$ ) 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 ( $p_z = 1.5 \times 10^{-4}$ ). rs641738C>T was not associated with fasting insulin and no effect was observed in children with NAFLD.

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

**Lay summary:** Fatty liver disease is a common condition where fat builds up in the liver, which can cause liver inflammation and scarring (including 'cirrhosis'). It is closely linked to obesity and diabetes, but some genes are also thought to be important. We did this study to see whether one specific change ('variant') in one gene ('*MBOAT7*') was linked to fatty liver disease. We took data from over 40 published studies and found that this variant near *MBOAT7* is linked to more severe fatty liver disease. This means that drugs designed to work on *MBOAT7* could be useful for treating fatty liver disease.

© 2020 European Association for the Study of the Liver. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Introduction

Since the first genome-wide association study (GWAS) of liver fat,<sup>1</sup> >20 genetic single nucleotide variants (SNVs) have been associated with NAFLD.<sup>2</sup> These studies have deepened our understanding of the condition, its heritability and its relationship with cardiometabolic disease.

rs641738C>T near membrane-bound O-acyltransferase domain-containing 7 (*MBOAT7*) was initially identified as a genome-wide significant risk variant for alcohol-related cirrhosis [odds ratio (OR) = 1.35,  $p = 1.03 \times 10^{-9}$ ],<sup>3</sup> although this was not replicated in a more recent analysis.<sup>4</sup> It has since been implicated in the pathogenesis of NAFLD,<sup>5</sup> HCC,<sup>6</sup> as well as in fibrosis development in chronic HBV and HCV,<sup>7,8</sup> and primary sclerosing cholangitis.<sup>9</sup> However, unlike variants in patatin-like phospholipase domain containing protein 3 (*PNPLA3*), transmembrane 6 superfamily member 2 (*TM6SF2*), and 17 $\beta$ -hydroxysteroid dehydrogenase type 13 (*HSD17B13*), it was not identified to have genome-wide significance for liver fat or serum alanine aminotransferase (ALT).<sup>1,10,11</sup>

Rs641738 is located a few hundred base pairs downstream of the 3'-untranslated region of *MBOAT7*, which belongs to a family of genes that encode specific acyl donors and acceptors.<sup>12</sup> *MBOAT7* encodes lysophosphatidylinositol acyltransferase 1 (*LPIAT1*), which contributes to the regulation of free arachidonic acid in cells.<sup>13,14</sup> Rs641738C>T is associated with lower hepatic expression of *MBOAT7* at both the mRNA<sup>15</sup> and protein levels.<sup>5</sup> Given its role in inflammatory lipid pathways, most mechanistic work relating to rs641738 has focussed on *MBOAT7*.<sup>16</sup>

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.<sup>5</sup> Proton magnetic resonance spectroscopy data from 2,736 individuals showed a modest increase in hepatic fat in those with the TT-genotype (4.1%) compared with those with the CT- (3.6%) or CC-genotype (3.5%,  $p = 0.005$ ). Follow-up studies of European subjects corroborated the initial findings, and suggested a role in development of HCC.<sup>17,18</sup> However, these results were not replicated in adults of other ancestries<sup>5,19–21</sup> or in children.<sup>22</sup>

In addition, bi-allelic loss-of-function mutations in *MBOAT7* cause autosomal recessive mental retardation 57 (Online Mendelian Inheritance in Man #617188) and no liver phenotype has been reported in these patients to date.<sup>14,23</sup> However, rare likely pathogenic (coding) variants in *MBOAT7* are associated with HCC in NAFLD.<sup>24</sup>

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 *MBOAT7* were to be investigated as a drug target in NAFLD.

Here, we conducted a large meta-analysis to determine whether 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 that 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.

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 (see [Supplementary Methods](#)). There were no restrictions on either date or language. The search was completed on July 28, 2020. Reference lists of publications were also reviewed.

A separate search was conducted for all potentially relevant GWAS through GWAS Catalogue,<sup>25</sup> Phenoscanner,<sup>26</sup> Type 2 diabetes knowledge portal<sup>27</sup> and Cardiovascular disease knowledge portal<sup>28</sup> (see [Supplementary Methods](#)).

After removal of duplicates, titles and abstracts were screened for eligibility independently by 2 authors (investigators), with inclusion/exclusion criteria applied to potentially eligible full texts.

HuGenet guidelines<sup>29</sup> were followed throughout and MOOSE reporting guidelines<sup>30</sup> were used. This study was prospectively registered on the PROSPERO Database of Systematic Reviews (CRD42018105507; [www.crd.york.ac.uk/PROSPERO/display\\_record.php?ID=CRD42018105507](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<sup>31</sup>)/rs2576452C>T ( $R^2 = 0.92$  in Guzman *et al.*<sup>32</sup>), 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 that investigated liver disease of other aetiologies were also excluded. There was no restriction on ethnicity or ancestry. The types of study eligible for inclusion were case-control, cohort, GWAS, systematic reviews and meta-analyses. Preprint and abstract publications were not eligible for inclusion. Several studies reported on the same cohort (or patient sample) in more than 1 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 provided in [Table S1](#).

### Data collection

Details of the recruitment of controls and cases were obtained from each study and, where necessary, clarified by discussion

with the authors of the study. 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 (liver ultrasound, controlled attenuation parameter [CAP, with cut-off >248 dB/m], CT, MRI, or histological assessment) were used. Hepatic fat content was collected as a continuous variable from CT, magnetic resonance spectroscopy (MRS), MRI, and proton density fat fraction (PDFF). Non-invasive assessment of hepatic fat content was also assessed using semiquantitative scoring in the Fenland cohort, as previously described,<sup>33</sup> and using CAP.

Individual participant-level histology data were extracted according to the NASH Clinical Research Network scoring system<sup>34</sup> and, where not otherwise diagnosed by a pathologist's assessment, NASH was defined using the Fatty Liver Inhibition of Progression algorithm.<sup>35</sup> 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 mellitus (T2DM), and body mass index (BMI). Where possible, individual patient-level data were obtained.

The authors of 59 studies were contacted for additional data or clarification, of whom 49 replied. Data from 11 potentially relevant studies could not be included, which are listed in the [Supplementary Methods](#).

Additional details regarding cohorts with genome-wide data, the Avon Longitudinal Study of Parents and Children (ALSPAC)<sup>36–38</sup> data extracted from the UK Biobank (UKBB), quality assessment and statistical analysis are found in the [Supplementary Methods](#).

## Results

Database searches identified 1,167 articles ([Fig. S1](#)), of which 44 articles were included: 42 primary studies ([Tables S2–S4](#)), 1 systematic review, and 1 meta-analysis ([Table S5](#)).

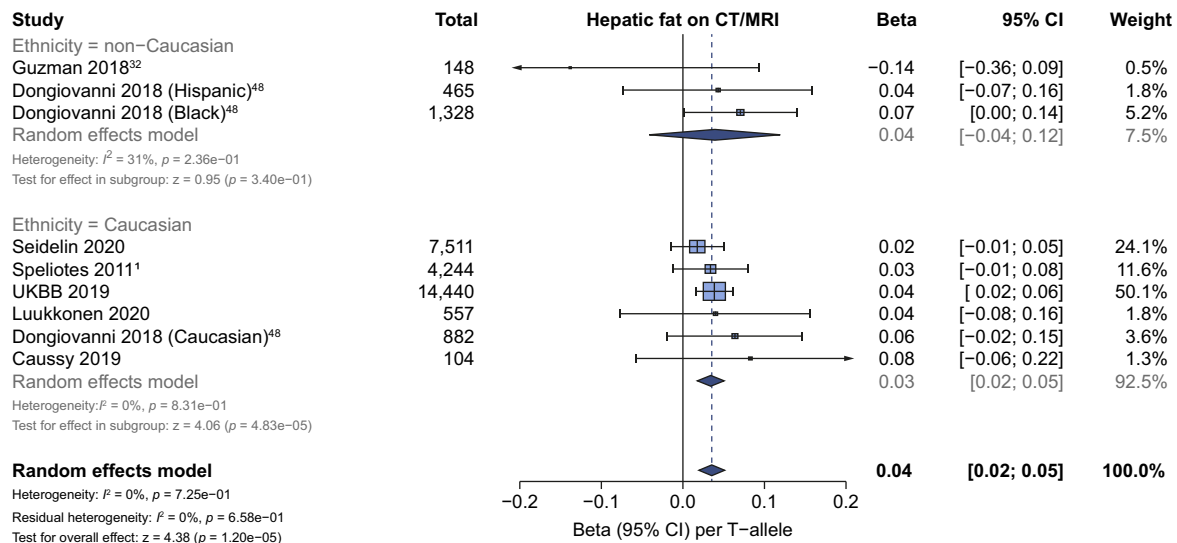
In total, 1,066,175 individuals (5,711 children) were included in the meta-analysis. Most studies were in adults (32/42, 76%) and in individuals from predominantly Caucasian populations (26/42, 62%). Of the 42 studies included, 14 (9,688 participants, including 584 children) reported data on liver histology.

Studies were generally of high quality, although, in 5 studies<sup>11,22,39–41</sup> (4 in adults and 1 in children), the control group was recruited from a different population or sample to the cases ([Table S3](#)).

One previous meta-analysis was included,<sup>42</sup> which used data from 5 case-control studies to assess the effect of rs641738C>T on the diagnosis of NAFLD. The meta-analysis included 2,560 cases and 8,738 controls and found no evidence of an association between this variant and a diagnosis of NAFLD ([Table S5](#)). One previous systematic review<sup>43</sup> found positive associations between rs641738C>T in adults of Caucasian, Hispanic, and Black descent, with limited data in children ([Table S5](#)).

### Liver fat, NAFLD and severe steatosis in adults

Seven studies (29,679 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 using an additive model of inheritance, with a per T-allele change of  $\beta$  0.034 (95% CI 0.018–0.051),  $p_z = 4.8 \times 10^{-5}$ ) standard deviations in inverse-normalised liver fat



**Fig. 1. The effect of rs641738C>T on liver fat.** Data from 29,679,916 individuals with CT, MRI or MRS liver fat. rs641738C>T was positively associated with liver fat in Caucasian populations (using an additive model of inheritance), where data represent SD change in normalised liver fat per T-allele. Meta-analysis was performed using random effects with DerSimonian-Laird method for estimation of tau<sup>2</sup>. Additional references are available in the [Supplementary Data](#). MRS, magnetic resonance spectroscopy; UKBB, UK BioBank.

(Fig. 1), whereas no consistent effect was observed in non-Caucasian populations. A similar trend was observed using a dominant model of inheritance in studies of Caucasian populations: mean difference in hepatic fat +0.18% (95% CI 0.2–0.34;  $p_z = 0.04$ ; [Table S6](#)).

Given the difference in sensitivity and specificity of modalities used to assess liver fat, a subanalysis by modality of imaging was performed. No significant differences were observed between studies using CT, MRI, or MRS for quantification of liver fat ([Fig. S2](#)).

A similar trend was observed using CAP and semiquantitative ultrasound to assess steatosis severity in 12,224 adults ( $\beta$  0.02 [95% CI -0.002–0.04],  $p_z = 0.08$ ; [Fig. S3](#)).

Data from a range of diverse modalities were used to assess the effect of this variant on the diagnosis of NAFLD, to reflect real-world diagnostic practice. rs641738C>T was associated with NAFLD as a trait (OR 1.15 [95% CI 1.05–1.26],  $p_z = 0.002$ ) using a recessive model of inheritance ([Fig. 2](#)) but not using additive or dominant models ([Table S7](#)). The effect was only observed in studies of Caucasian populations (OR 1.17 [95% CI 1.05–1.3],  $p_z = 0.003$ ). Subgroup analysis by modality of diagnosis found that the 95% CIs for all modalities overlapped, except for MRI-PDFF, which had only 1 study ([Fig. S4](#)). The association remained after excluding 4 studies in which there was a lack of similarity between cases and controls (OR 1.19 [95% CI 1.07–1.33],  $p_z = 0.0017$ ) using a recessive model of inheritance.

However, Egger's test suggested evidence of study distribution (publication) bias ( $p = 0.013$ ) and when using the Trim and Fill method to account for this bias, the positive association remained but was attenuated (OR 1.11 [95% CI 1.01–1.23],  $p_z = 0.037$ ; [Fig. S5](#)).

In patients with NAFLD, data from 8 studies (6,206 participants) showed that rs641738C>T was not significantly associated with the presence of severe steatosis (S1–S2 vs. S3) on

liver biopsy (OR 1.08 [95% CI 0.78–1.5],  $p_z = 0.64$ ; [Table 1](#) and [Fig. S6](#)).

#### Histological NASH in adults

Data from 9 studies (7,719 participants) found that rs641738C>T was not associated with the presence of NASH on biopsy in adults (OR 1.24 [95% CI 0.96–1.36],  $p_z = 0.128$ ; [Fig. S7](#)).

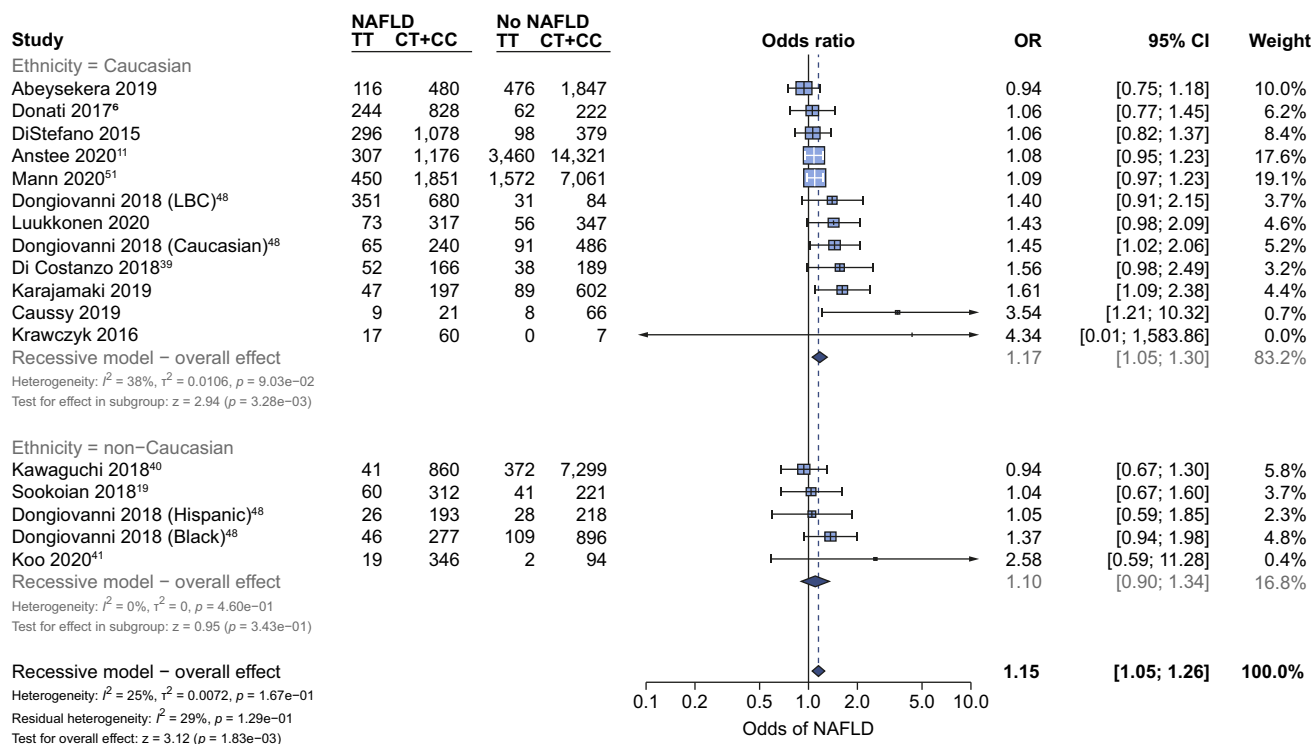
#### Fibrosis in adults

Liver biopsy data on the presence of advanced fibrosis were available from 8 studies (7,692 adults). Our primary outcome, presence of advanced fibrosis in adults (stage F0–F2 vs. stage F3–F4), showed a borderline positive association with rs641738C>T in Caucasian populations (OR 1.22 [95% CI 1.03–1.45],  $p_z = 0.021$ ; [Fig. 3](#)). In addition, 2 studies used International Statistical Classification of Diseases and Related Health Problems (ICD) codes in the UKBB cohort to identify individuals with NAFLD and advanced fibrosis or cirrhosis.<sup>44,45</sup> 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  $\beta$  1.22 (SE 0.06,  $p = 0.03$ ), using an additive genetic model.<sup>44</sup>

Data from 9 studies (8,389 participants) found that the presence of any fibrosis (F0 vs. F1–F4) was also borderline positively associated with rs641738C>T overall (OR 1.27 [95% CI 1.04–1.54],  $p_z = 0.018$ ) as well as in non-Caucasian populations as a subgroup ([Fig. S8](#)).

#### Development of HCC

Four cohorts (2,328 participants, 228 cases of NAFLD-HCC) reported on the 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],  $p_z = 0.003$ , [Fig. 4](#)).



**Fig. 2. rs641738C>T is associated with higher odds of diagnosis of NAFLD.** Data from 52,17333,263 adults (11,3019,713 cases and 40,87223,550 controls) with radiologically or histologically defined steatosis for presence vs. absence of NAFLD using a recessive model of inheritance (CC + CT vs. TT). Meta-analysis was performed using random effects with DerSimonian-Laird method for estimation of tau<sup>2</sup>. Additional references are available in the [Supplementary Data](#). LBC, Liver Biopsy Cohort; OR, odds ratio.

**Table 1. Summary of results in adults from meta-analyses for dichotomous outcomes.**

Outcome	Genetic model	Subanalysis	No. of studies	Heterogeneity		Effect summary	
				I <sup>2</sup>	p <sub>Q</sub>	OR (95% CI)	p <sub>Z</sub>
NAFLD diagnosis (control vs. NAFLD)	Recessive	Overall	17	0.25	0.17	1.15 (1.05–1.26)	0.0018
		Non-Caucasian	5	0	0.46	1.1 (0.9–1.34)	0.343
		Caucasian	12	0.38	0.09	1.17 (1.05–1.3)	0.0033
Severe steatosis (S1–S2 vs. S3)	Recessive	Overall	8	0.67	0	1.08 (0.78–1.5)	0.642
		Non-Caucasian	1	NA	NA	1.11 (0.39–3.16)	0.852
		Caucasian	7	0.72	0	1.08 (0.76–1.54)	0.676
NASH (NAFL vs. NASH)	Recessive	Overall	9	0.33	0.15	1.14 (0.96–1.36)	0.128
		Non-Caucasian	3	0	0.58	1.24 (0.81–1.9)	0.324
		Caucasian	6	0.53	0.06	1.14 (0.93–1.41)	0.213
Any fibrosis (F0 vs. F1–F4)	Recessive	Overall	9	0.52	0.03	1.27 (1.04–1.54)	0.0183
		Non-Caucasian	2	0	0.82	2.14 (1.2–3.84)	0.0105
		Caucasian	7	0.51	0.06	1.19 (0.99–1.45)	0.068
Advanced fibrosis (F0–F2 vs. F3–F4)	Recessive	Overall	8	0	0.65	1.2 (1.02–1.42)	0.027
		Non-Caucasian	2	0	0.64	0.96 (0.5–1.85)	0.911
		Caucasian	6	0	0.5	1.22 (1.03–1.45)	0.0206
HCC (NAFLD-HCC vs. NAFLD no-HCC)	Recessive	Overall	4	0	0.95	1.4 (0.99–1.98)	0.056

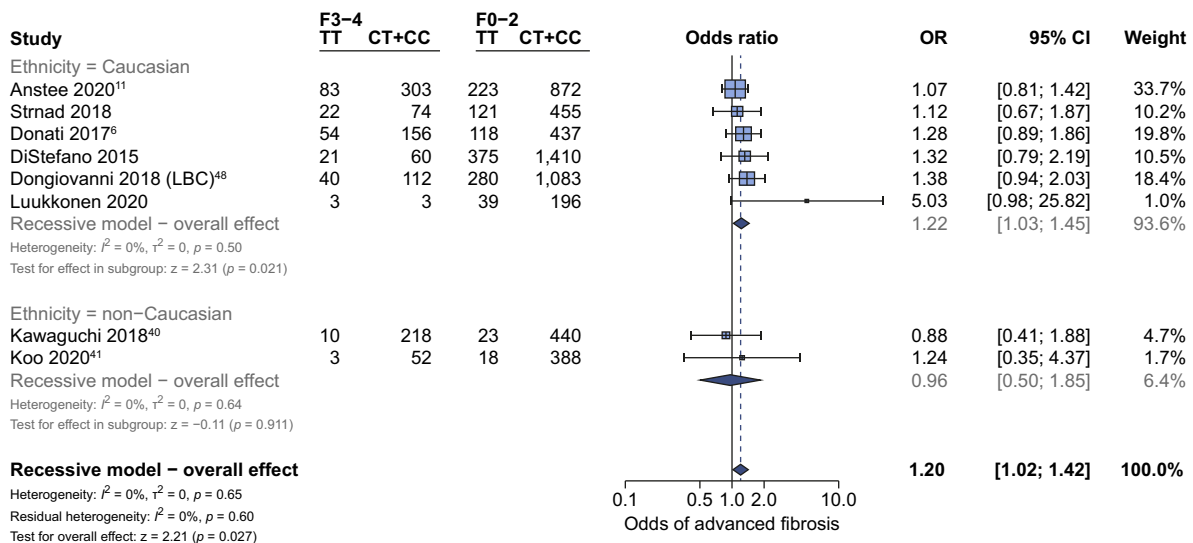
Meta-analyses were performed using random effects with subgroup analysis for Caucasian and non-Caucasian populations. Additive, recessive and dominant genetic models were tested for all outcomes. Results using a recessive model of inheritance (CC + CT vs. TT) are shown for all outcomes. Given the use of 3 genetic models, the critical p value for effect summary is p<sub>Z</sub> < 0.017. Full results (with all genetic models) are in [Table S5](#). Meta-analyses were performed using random effects with DerSimonian-Laird method for estimation of tau<sup>2</sup>. OR, odds ratio.

**Effect on alanine aminotransferase**

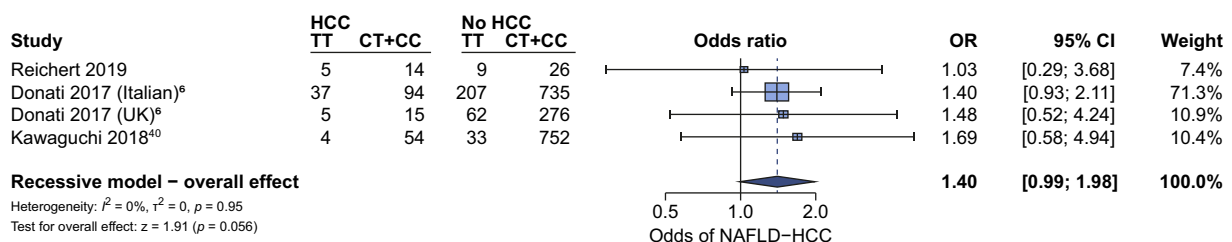
Data from GWAS using log-transformed ALT (609,794 participants) were available for meta-analysis to investigate the effect of rs641738C>T on ALT. The variant showed a positive association with ALT ( $\beta$  0.004 [95% CI 0.002–0.007], p<sub>Z</sub> = 0.002), which was observed in Caucasian populations but not in non-Caucasian populations on subanalysis ([Fig. 5](#); [Table S7](#)).

Additionally, in the UKBB cohort, rs641738C>T was associated with a small, but statistically significant (p = 2.0 × 10<sup>-8</sup>) increase in untransformed ALT: 0.18 IU/L higher ALT per T-allele in this variant ([Table S8](#)).

In the remaining cohort and case-control studies included in the meta-analysis (15,208 adults), rs641738C>T was not found to be significantly associated with a change in ALT; for example, the



**Fig. 3. The effect of rs641738C>T on the presence of advanced fibrosis in adult patients with NAFLD.** Data from 7,692,211 adults (1,214,828 cases and 6,478,383 controls) with biopsy-proven NAFLD comparing advanced fibrosis (F3–F4) vs. F0–F2, using a recessive model of inheritance (CC + CT vs. TT). Meta-analysis was performed using random effects with DerSimonian-Laird method for estimation of tau<sup>2</sup>. Additional references are available in the [Supplementary Data](#). LBC, Liver Biopsy Cohort; OR, odds ratio.



**Fig. 4. rs641738C>T is associated with higher odds of NAFLD-HCC.** Data from 2,328 adults with NAFLD assessing for the presence vs. absence of HCC, using a recessive model of inheritance (CC + CT vs. CT). Meta-analysis was performed using random effects with DerSimonian-Laird method for estimation of tau<sup>2</sup>.

mean difference using a recessive model (CC + CT vs. TT) was +0.32 IU/L [95% CI -0.06–0.7],  $p_z = 0.08$ ; [Table S9](#)) in Caucasian populations.

### 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 ( $\beta -0.01$  [95% CI -0.018, -0.006],  $p_z = 1.5 \times 10^{-4}$ ), which was observed in Caucasian populations but not in non-Caucasian populations on subanalysis ([Fig. S9](#)). Similar findings were obtained from a meta-analysis of cohort and case-control studies, particularly when using an additive model ( $\beta -0.03$  [95% CI -0.05, -0.01],  $p_z = 0.00091$ ; [Table S9](#)).

Data from GWAS (852,409 participants) found rs641738C>T to be positively associated with total cholesterol in Caucasian populations ( $\beta 0.007$  [95% CI 0.003–0.01],  $p_z = 2.1 \times 10^{-4}$ ), which was not observed in non-Caucasian populations ([Fig. S10](#)). A borderline positive association was also observed between rs641738C>T and HDL cholesterol ( $\beta 0.009$  [95% CI 0.001–0.02],  $p_z = 0.02$ ; [Table S7](#)). There was no effect on fasting insulin levels found in population-level GWAS ( $\beta 0.009$  [95% CI -0.03–0.04],  $p_z = 0.64$ ; [Table S7](#)). However, a negative association was observed using data from cohort and case-control studies with a dominant genetic model (mean difference -1.4 pmol/L [95% CI -2.1, -0.65],  $p_z = 0.004$ ; [Table S9](#)).

### Effect of rs641738C>T on paediatric NAFLD

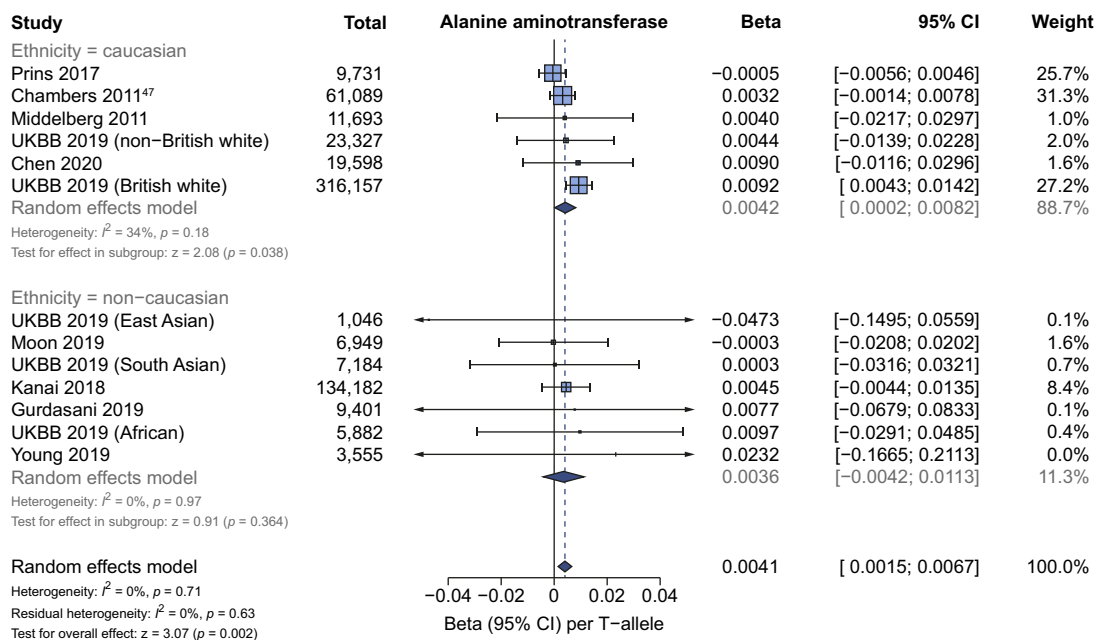
Data from 10 studies (5,711 children) were used in the meta-analysis. rs641738C>T was not significantly associated with the diagnosis of NAFLD, liver fat content, stage of liver histology, or serum biochemistry in children ([Table S10](#)).

### Meta-regression shows interaction between rs641738C>T and type 2 diabetes mellitus

Finally, we aimed to determine using meta-regression whether baseline participant characteristics influenced the association of rs641738C>T with histological outcomes. There was a negative association with the presence of T2DM and effect size for NASH vs. NAFL ( $\beta -1.8$  [SE 0.65],  $p = 0.006$ ; [Fig. S11A](#)). A similar negative trend with T2DM was observed for severe steatosis (S1–S2 vs. S3,  $\beta -2.6$  [SE 1.5],  $p = 0.08$ ) and presence of fibrosis (F0 vs. F1–4,  $\beta -1.5$  [SE 0.8],  $p = 0.06$ ; [Table S11](#)). In addition, the effect size for any fibrosis was greater in cohorts with an older mean age ( $\beta 0.05$  [SE 0.02],  $p = 0.014$ ; [Fig. S11D](#)).

### Discussion

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



**Fig. 5. rs641738C>T is positively associated with ALT in Caucasian populations in GWAS.** Meta-analysis using linear regression of GWAS summary statistics from 609,794 participants for the association between rs641738C>T and logarithmically transformed ALT. Meta-analysis was performed using random effects with DerSimonian-Laird method for estimation of tau<sup>2</sup>. Additional references are available in the Supplementary Data. GWAS, genome-wide association studies; UKBB, UK BioBank.

A 2-stage GWAS initially identified rs641738C>T as a genome-wide significant locus for alcohol-related cirrhosis.<sup>3</sup> *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.<sup>5,17</sup> However, these associations have not been consistently replicated in the literature.<sup>19</sup> We conducted a meta-analysis to firmly establish the association of rs641738C>T with the presence and severity of NAFLD and associated metabolic traits.

**Main findings**

We found that rs641738C>T was associated with higher liver fat content, higher ALT, and with higher odds of NAFLD diagnosis, fibrosis, and HCC, particularly in Caucasian adults and in the homozygous ‘TT’ genotype. The effect sizes of rs641738C>T reported here are small compared with those of *PNPLA3* p.I148M and *TM6SF2* p.E167K, the 2 strongest steatogenic variants.<sup>46</sup> Also, the magnitude of change in ALT is small relative to that associated with variants in *PNPLA3*, *HSD17B13*, mitochondrial amidoxime reducing component 1 (*MTARC1*), and *TM6SF2*. This might account for the absence of this variant (or others near *MBOAT7*) from GWAS for NAFLD in the general population.<sup>1,10,11,45,47</sup> The effect size (and associated *p* value) was too small to be identified as significant genome-wide. The marginal positive effect on hepatic triglyceride content suggests that this variant acts through alterations in the composition as well as quantity of hepatic lipid.<sup>17</sup> This is consistent with pre-clinical data on lipotoxicity, where the composition of hepatic fats influences the development of NASH. By contrast, a recent Mendelian randomisation study using these variables as

instruments to assess causality of fatty liver in determining fibrosis showed that the effect of steatosis highly correlates with fibrosis in all the genetic variables, indicating that quantity of lipid rather than quality might be more important.<sup>48</sup> Functional studies are needed to understand the relationship between quality/quantity of fat and hepatotoxic/-protective mechanisms 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 the hepatic expression of *MBOAT7*. Results from the GTEx Consortium showed a strong negative association with T-allele,<sup>15</sup> which is supported by data from Schadt *et al.*<sup>49</sup> *MBOAT7* protein expression correlated with mRNA in liver biopsies from Mancina *et al.*,<sup>5</sup> but this finding was not replicated by Sookoian *et al.*<sup>19</sup> *MBOAT7* encodes LPIAT1, a 6-transmembrane domain protein involved in acyl-chain remodelling of membranes that influence intracellular membrane composition and circulating phosphatidylinositols.<sup>50</sup> Furthermore, recent metabolite profiling data implicated *MBOAT7* as the causal gene for this SNV.<sup>51</sup> Moreover, transmembrane channel-like 4 (*TMC4*) was found to have a low expression in the liver,<sup>5</sup> which 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.<sup>50</sup> Similarly, metabolite data from humans are strongly suggestive that rs641738C>T reduces *MBOAT7* function.<sup>52</sup> In addition, 2 independent groups found that loss of *MBOAT7* (but not *TMC4*) increases the severity of NAFLD in mice fed a high-fat diet.<sup>53,54</sup>

These analyses suggest that rs641738C>T impacts the severity of NAFLD through a recessive model of inheritance, although



some analyses using an additive genetic model were suggestive of a role (e.g. for liver fat and ALT). Other genetic variants are known to impact all-cause mortality in a recessive manner, notably variants that perturb homeostatic iron regulator protein (*HFE*).<sup>44</sup> Further mechanistic work is required to understand the extent to which the haplo-insufficient state affects hepatocyte function.

We found no evidence of an effect of rs641738C>T on insulin resistance (the key driver of hepatic steatosis) as determined by unaltered fasting insulin concentrations. GWAS meta-analyses of T2DM have implicated p.I148M in *PNPLA3* and p.E167K in *TM6SF2* as significant risk loci (albeit with very modest effect sizes compared with their effects on liver disease)<sup>55</sup> and Mendelian randomisation studies indicate a causal role in determining insulin resistance mediated by the degree of liver damage.<sup>48,56</sup> Similarly, these 2 variants are associated with reduced risk of coronary artery disease; although our analysis did find lower serum triglycerides to be associated with this variant, it has not been associated with lower rates of cardiovascular disease.<sup>57</sup> However, we did observe a negative association between effect size and prevalence of diabetes on meta-regression, suggesting that this variant has the greatest effect in less insulin-resistant individuals.

A strength of this meta-analysis is the large number of individuals with liver biopsy-derived phenotypic data as well as the 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 with the previous meta-analysis by Xia *et al.*<sup>42</sup>

### 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 with 0.53 in those of South Asian ancestry.<sup>58</sup> Moreover, the majority of studies included in this meta-analysis used self-reported ethnicity, rather than genetic ancestry.

Although 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 and it is likely that rs641738C>T is a proxy for the true causal variant. However, because of 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.

A limitation of using meta-analysis for a single variant is the lack of adjustment for population stratification. When further genome-wide data are available, a formal GWAS meta-analysis might be able to address this. We found significant differences between adult and paediatric histological analyses. Although 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, therefore, it is plausible that this is a true lack of association in children with NAFLD.

Data from multiple diagnostic or imaging modalities were combined in several analyses. Although we observed minimal heterogeneity between modalities, these techniques have differing accuracy for the diagnosis of steatosis, which has the

potential to affect results. The subgroup analysis of hepatic fat by modality suggested a marginally greater effect size in studies using MRS, which is regarded as a highly sensitive technique. There is potential that, through the inclusion of other modalities (e.g. CT), we have underestimated the effect size associated with this variant.

The magnitude of effect observed across all associations was small compared with other well-established variants. The clinical relevance of rs738409C>G in *PNPLA3* has been validated with hard end-points,<sup>59</sup> but large cohorts will be required to prospectively demonstrate the clinical risk associated with this variant near *MBOAT7*.

Although 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, limiting the power to assess for an association of this variant with non-cirrhotic HCC, as has been previously reported.<sup>6</sup> The HCC analysis was also unique in only demonstrating an effect in the dominant, rather than recessive, model of inheritance. Further work in this area might 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.

### Abbreviations

ALSPAC, Avon Longitudinal Study of Parents and Children; BMI, body mass index; CAP, controlled attenuation parameter; GWAS, genome-wide association study; *HFE*, homeostatic iron regulator protein; HSD17B13, 17 $\beta$ -hydroxysteroid dehydrogenase type 13; LPIAT1, lysophosphatidylinositol acyltransferase 1; *MBOAT7*, membrane bound O-acyltransferase domain containing 7; *MTARC1*, mitochondrial amidoxime reducing component 1; OR, odds ratio; MRS, magnetic resonance spectroscopy; PDFF, proton density fat fraction; *PNPLA3*, patatin-like phospholipase domain containing protein 3; SNV, single nucleotide variant; T2DM, type 2 diabetes mellitus; *TM6SF2*, transmembrane 6 superfamily member 2; *TMC4*, transmembrane channel-like 4; UKBB, UK BioBank.

### Financial support

J.P.M. 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 (S.C.), R01DK111038 (S.C.), R01DK114504 (N.S.), DK091601 (J.K.D.) and UL1TR001105 (J.K.). 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 F.S. 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 ([www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf](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 K.W.M.A. and M.H. L.V. 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. German Federal Ministry of Education and Research (BMBF LiSyM 031L0051 to F.L.). P.L. 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 UU12015/5). R.B. and M.K. are employees of and shareholders in Perspectum Diagnostics Ltd. C.A.P. is funded by a Wellcome Trust Clinical PhD Programme (206274/Z/17/Z). J.M.B. is supported by the Spanish Carlos III Health Institute (ISCIII; PI15/01132, PI18/01075, CIBERehd, and Miguel Servet Program CON14/00129) co-financed by Fondo Europeo de Desarrollo Regional (FEDER) and La Caixa Scientific Foundation (HR17-00601). Research from the GUARDIAN Study was supported DK085175 and DK118062, and phenotyping in IRASFS was supported by HL060944, HL061019, HL060919 and HL060894 IRASFS from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). H.Y. is funded by a Diabetes UK RD Lawrence fellowship (17/0005594). Q.M.A. and A.D. supported by the EPoS (Elucidating Pathways of Steatohepatitis) consortium funded by the Horizon 2020 Framework Program of the European Union under Grant Agreement 634413. Q.M.A., A.D., L.V. and A.G. are members of the LITMUS (Liver Investigation: Testing Biomarker Utility in Steatohepatitis) consortium funded by the European Union Innovative Medicines Initiative 2 (IMI2) Joint Undertaking under grant agreement 777377. Q.M.A. is a Newcastle NIHR Biomedical Research Center Investigator.

### Conflict of interest

C.E. reports receiving personal fees from Navitor Pharma and Novartis. The other authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying [ICMJE disclosure](#) forms for further details.

### Authors' contributions

Study concept and design: J.P.M.; acquisition of data: all; analysis and interpretation of data: K.T., J.K., S.S., L.V., J.P.M.; drafting of the manuscript: K.T., J.P.M.; critical revision of the manuscript for important intellectual content: all authors; statistical analysis: K.T., J.K., S.S., L.V., J.P.M.; obtained funding: J.P.M., S.C., N.S., J.K.D.,

J.K., P.L., J.M.B., C.A.P., H.Y., K.W.M.A., L.A., Q.M.A., A.D.K., T.B., G.S.G., C.H., J.H., J.L., P.E.M., T.A.M., N.D.P., S.R., J.I.R., E.K.S., S.S., A.T.H., L.E.W., L.V., H.Y.J., K.A.Y.; study supervision: J.P.M.

### 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 was 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, and the late Valerio Nobili. The authors are also grateful to the members of the GOLD consortium: Gudny Eiriksdottir, Melissa E. Garcia, Vilmundur Gudnason, Tamara B. Harris, Lauren J. Kim, Lenore J. Launer, Michael A. Nalls, Albert V. Smith, Jeanne M. Clark, Ruben Hernaez, W. H. Linda Kao, Braxton D. Mitchell, Alan R. Shuldiner, Laura M. Yerges-Armstrong, Ingrid B. Borecki, J. Jeffrey Carr, Mary F. Feitosa, Jun Wu, Johannah L. Butler, Caroline S. Fox, Joel N. Hirschhorn, Udo Hoffmann, Shih-Jen Hwang, Joseph M. Massaro, Christopher J. O'Donnell, Cameron D. Palmer, Dushyant V. Sahani, Elizabeth K. Speliotes. We would also like to thank Naga Chalasani for his helpful comments. This research made use of the UK Biobank resource under project number 9914.

### Data availability

Raw data used in analyses is available in the [supplementary tables](#) online. Code used in analyses is available. Further details are available from the corresponding author upon request.

### Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.jhep.2020.08.027>.

### References

*Author names in bold designate shared co-first authorship*

- [1] **Speliotes EK, Yerges-armstrong LM, Wu J, Hernaez R, Kim IJ, 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] 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;47:1443–1448.
- [4] Innes H, Buch S, Hutchinson S, Guha IN, Morling JR, Barnes E, et al. Genome-wide association study for alcohol-related cirrhosis identifies risk loci in MARC1 and HNRNPUL1. *Gastroenterology* 2020. <https://doi.org/10.1053/j.gastro.2020.06.014>. In press.
- [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;150:1219–1230.
- [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;65:1840–1850.
- [9] Freund C, Wahlers A, Begli NH, Leopold Y, Klötters-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>. In press.
- [10] 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;378:1096–1106.
- [11] 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;73:505–515.
- [12] 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;13:879–888.
- [13] 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.
- [14] Johansen A, Rosti RO, Musaeov 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.
- [15] The GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580–585.
- [16] 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.
- [17] Luukkonen PK, Zhou Y, Hyötyläinen T, Leivonen M, Arola J, Orholm-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;65:1263–1265.
- [18] 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;58:247–255.
- [19] Sookoian S, Flichman D, Garaycochea ME, Gazzi C, Martino JS, Castañó 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.
- [20] 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.
- [21] 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;113:376–383.
- [22] 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;39:540–556.
- [23] 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;42:381–388.
- [24] 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.
- [25] 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–D901.
- [26] 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;35:4851–4853.
- [27] Type 2 diabetes knowledge portal. [www.type2diabetesgenetics.org/](http://www.type2diabetesgenetics.org/). [Accessed 30 September 2020].
- [28] Cardiovascular Disease Knowledge Portal. [www.broadcvdi.org/](http://www.broadcvdi.org/). [Accessed 30 September 2020].
- [29] Little J, Higgins J, editors. The HuGENet™ HuGE Review Handbook. Atlanta: Centers for Disease Control and Prevention; 2006.
- [30] Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies. *JAMA* 2000;283:2008–2012.
- [31] Arnold M, Raffler J, Pfeufer A, Suhre K, Kastenmüller G. SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics* 2015;31:1334–1336.
- [32] 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.
- [33] De Lucia Rolfe E, Brage S, Sleight 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.
- [34] 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;41:1313–1321.
- [35] 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;60:565–575.
- [36] 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;42:111–127.
- [37] 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.
- [38] 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;42:377–381.
- [39] Di Costanzo A, Belardinelli F, Ballelli D, Sponziello M, D’Erasmio 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.
- [40] 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.
- [41] 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;18:2592–2599.
- [42] 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;43:1–9.
- [43] Ismaiel A, Dumitrascu DL. Genetic predisposition in metabolic-dysfunction-associated fatty liver disease and cardiovascular outcomes: systematic review. *Eur J Clin Invest* 2020:e13331.
- [44] Emdin CA, Haas ME, Khera AV, Aragam K, Chaffin M, Klarin D, et al. A missense variant in Mitochondrial Amidoxime Reducing Component 1 gene and protection against liver disease. *PLoS Genet* 2020;16:e1008629.

- [45] 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;40:405–415.
- [46] **Sookoian S, Pirola CJ, Valenti L, Davidson NO.** Genetic pathways in nonalcoholic fatty liver disease: insights from systems biology. *Hepatology* 2020;72:330–346.
- [47] 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;43:1131–1138.
- [48] **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.
- [49] **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.
- [50] 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;23:4689–4700.
- [51] **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. *Hum Mol Genet* 2020:ddaa162. <https://doi.org/10.1093/hmg/ddaa162>. In press.
- [52] **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;46:543–550.
- [53] 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.
- [54] **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>. In press.
- [55] 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;50:1505–1513.
- [56] **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;73:241–251.
- [57] 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;3:587–596.
- [58] Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020;581:434–443.
- [59] **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;18:935–944.