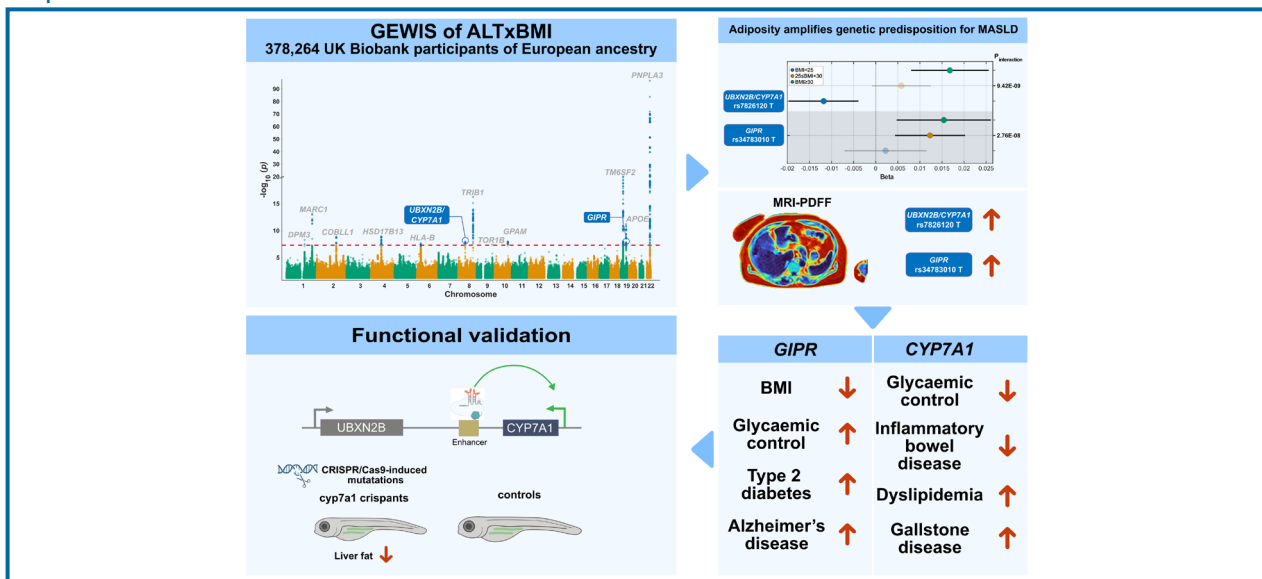


Genome-wide interaction study with body mass index identifies *CYP7A1* and *GIPR* as genetic modulators of metabolic dysfunction-associated steatotic liver disease

Oveis Jamialahmadi^{1,*}, Endrina Mujica^{2,*}, Lowri Morris^{3,*}, Rosellina Margherita Mancina^{1,4,5}, Ester Ciociola¹, Sami F. Qadri^{6,7}, Samantha Maurotti⁸, Francesco Malvestiti⁹, Ruifang Li-Gao¹⁰, Luisa Ronzoni¹¹, Federica Tavaglione¹², Hannah Maude³, Amin Allalou^{13,14}, Anastasia Emmanouilidou², Umberto Vespasiani-Gentilucci^{15,16}, Frits Richard Rosendaal¹⁰, Hannele Yki-Järvinen^{6,7}, Inês Cebola^{3,*}, Luca Valenti^{9,11,*}, Marcel den Hoed^{2,*}, and Stefano Romeo^{1,17,18,19,20,*}

¹Department of Molecular and Clinical Medicine, University of Gothenburg, Gothenburg; ²The Beijer Laboratory and Department of Immunology, Genetics and Pathology, Uppsala University and SciLifeLab, Uppsala, Sweden; ³Section of Genetics and Genomics, Department of Metabolism, Digestion, and Reproduction, Faculty of Medicine, Imperial College London, London, UK; ⁴Research Unit of Clinical Medicine and Hepatology, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma; ⁵Department of Life Science, Health, and Health Professions, Link Campus University, Rome, Italy; ⁶Department of Medicine, University of Helsinki and Helsinki University Hospital; ⁷Minerva Foundation Institute for Medical Research, Helsinki, Finland; ⁸Department of Experimental and Clinical Medicine, Magna Graecia University, Catanzaro; ⁹Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy; ¹⁰Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, Netherlands; ¹¹Precision Medicine – Biological Resource Center and Department of Transfusion Medicine, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy; ¹²MASLD Research Center, Division of Gastroenterology and Hepatology, University of California at San Diego, La Jolla, CA, USA; ¹³Division of Visual Information and Interaction, Department of Information Technology, Uppsala University; ¹⁴DanioReadout, Immunology Genetics and Pathology, Uppsala University, Uppsala, Sweden; ¹⁵Department of Medicine and Surgery, Università Campus Bio-Medico di Roma; ¹⁶Clinical Medicine and Hepatology Unit, Fondazione Policlinico Universitario Campus Bio-Medico, Rome; ¹⁷Clinical Nutrition Unit, Department of Medical and Surgical Sciences, Magna Graecia University, Catanzaro, Italy; ¹⁸Department of Cardiology, Sahlgrenska University Hospital, Gothenburg; ¹⁹Department of Medicine (H7), Karolinska Institute, Huddinge; ²⁰Department of Endocrinology, Karolinska University Hospital, Huddinge, Sweden

Graphical Abstract



Study Highlights

- We identified 13 BMI-ALT interaction loci in UK Biobank participants, including a novel locus, *GIPR*.
- *GIPR* rs34783010-T associates with lower BMI, higher HbA1c, and higher liver lipid content.
- The *UBXN2B/CYP7A1* locus associates with liver fat in UK Biobank and four European cohorts.
- Epigenomic, genetic, and clinical evidence implicates *CYP7A1* as the causal gene.
- CRISPR/Cas9 perturbation of zebrafish *CYP7A1* orthologues reduces liver fat in metabolically challenged larvae.
- Adiposity modifies the genetic susceptibility to MASLD; weight loss effects on liver fat may depend on *GIPR* and *CYP7A1* genotypes.
- Our data support *GIPR* agonists as therapy for type 2 diabetes and MASLD.

Background/Aims: Metabolic dysfunction-associated steatotic liver disease (MASLD) may progress to liver inflammation, fibrosis, cirrhosis and hepatocellular carcinoma. So far, genome-wide association studies explain a small fraction of MASLD heritability.

Methods: We sought to identify novel genetic determinants of MASLD by exploring interactions between genetic variants and body mass index (BMI). First, we examined genome-wide interactions with BMI for circulating alanine aminotransferase (ALT) levels using UK Biobank data. For identified loci, we next examined associations with hepatic proton density fat fraction (PDFF) in 35,146 independent UK Biobank participants. Associations with PDFF were replicated in four independent European cohorts, followed by a phenome-wide association study. Finally, we used human liver epigenomic maps and CRISPR/Cas9 experiments *in vitro* and *in vivo* to functionally characterize the *CYP7A1* locus.

Results: Thirteen loci interact with BMI for ALT ($P < 5E-8$), including eight well-known genetic modulators of MASLD. Two loci—*UBXN2B/CYP7A1* and *GIPR*—are additionally associated with PDFF. For the intronic rs34783010 in *GIPR*, the minor T allele is associated with lower BMI and higher HbA1c and liver triglyceride content in humans. The *UBXN2B/CYP7A1* locus is associated with PDFF in four additional European cohorts. Epigenomic data and *in vitro* experiments in human liver cells prioritise rs10504255 and *CYP7A1* as the functional effectors in this locus. Perturbation of *CYP7A1* orthologues using CRISPR/Cas9 results in less liver fat in 10-day-old, metabolically challenged zebrafish larvae.

Conclusions: A genome-wide single nucleotide polymorphism \times BMI design fuelled identification of two MASLD genes: *CYP7A1* and *GIPR*. (*Clin Mol Hepatol* 2025;31:1252-1268)

Keywords: Gene-environment interaction; Genome-wide association study; MASLD; *CYP7A1*; Gastric inhibitory polypeptide receptor

Corresponding author : Stefano Romeo

Department of Medicine, Huddinge Karolinska Institute, Stockholm, Sweden
Tel: +46(0)313426735, E-mail: stefano.romeo@ki.se
<https://orcid.org/0000-0001-9168-4898>

*Denotes equal contributions.

Editor: Murim Choi, Seoul National University College of Medicine, Korea

Received : Feb. 26, 2025 / **Revised :** May 28, 2025 / **Accepted :** May 29, 2025

Abbreviations:

ALT, alanine aminotransferase; ATAC-seq, assay for transposase-accessible chromatin using sequencing; BMI, body mass index; ChIP-seq, chromatin immunoprecipitation sequencing; CRE, cis regulatory element; CRISPRa, clustered regularly interspaced short palindromic repeats activation; GEWIS, genome-wide-environment interaction study; GWAS, genome-wide association study; HbA1C, glycosylated haemoglobin; HDL, high-density lipoprotein; HWE, Hardy-Weinberg equilibrium; ICD-10, International Classification of Diseases 10th edition; LD, linkage disequilibrium; LDL, low-density lipoprotein; MAF, minor allele frequency; MASH, metabolic-dysfunction associated steatohepatitis; MASLD, metabolic-dysfunction associated steatotic liver disease; MRI, magnetic resonance imaging; PDFF, proton density fat fraction; PIP, posterior inclusion probability; SLD, steatotic liver disease; SNP, single nucleotide polymorphism; TF, transcription factor; TPM, transcript per million; UK, United Kingdom

INTRODUCTION

Following the obesity and type 2 diabetes epidemics, the prevalence of metabolic dysfunction-associated steatotic liver disease (MASLD), previously known as non-alcoholic fatty liver disease, is 32% and rising worldwide. Despite its high prevalence, only one FDA-approved drug has shown efficacy in treating fibrotic steatohepatitis (metabolic-dysfunction associated steatohepatitis, MASH) or preventing its progression.¹ Weight loss is typically prescribed to reduce liver fat content and progression of liver disease.² However, not all patients with SLD are overweight or obese.

MASLD encompasses a spectrum of conditions defined by liver triglyceride accumulation greater than 5%. MASLD is a complex trait deriving from the interaction between genetic and environmental factors. Among the environmental factors, excess adiposity, quantified by elevated body mass index (BMI) is the major MASLD risk factor. When adipose tissue fails to expand, ectopic triglycerides accumulate in the liver, which can progress to inflammation, fibrosis and hepatocellular carcinoma.³

Genetic predisposition also contributes to MASLD susceptibility. Genome-wide association and candidate gene studies have identified common genetic variants associated with MASLD.⁴⁻⁶ However, few loci have been identified compared with other cardiometabolic diseases, and a large fraction of MASLD heritability remains unexplained. This in part reflects the invasive nature of MASLD's diagnosis, and concomitant small number of cases in genome-wide association studies (GWAS) for MASLD. Alternative approaches are required to improve our understanding of MASLD pathophysiology. We and others have already shown the merit of using proxy measures of MASLD,^{6,7} but interactions between genetic and environmental factors should be explored further.

Environmental factors may modify the effect of a genetic factor on a trait. In 2017, Stender et al.⁸ provided the first robust, formal evidence of such a gene-environment interaction for MASLD. By using a candidate gene approach, they showed a robust interaction between excess body mass and common genetic variants in *PNPLA3*, *TM6SF2*, and *GCKR* for liver fat content.⁸ This interaction is consistent with the previous observation that excess adiposity amplifies the effect of the *PNPLA3* rs738409 variant in inducing liver damage, as measured by circulating levels of

alanine aminotransferase (ALT), a clinical marker of MASLD.⁹

In this study, we sought to exploit gene-environment interactions between genetic variants and BMI for ALT levels in UK Biobank participants. We identified 13 loci interacting with BMI for ALT. One of these—in *GIPR*—is associated with lower BMI, yet worse glycaemic control, higher odds of diabetes, and higher liver fat content. Another—between UBX domain protein 2B (*UBXN2B*) and cytochrome P450 family 7 subfamily A member 1 (*CYP7A1*)—is likely associated with liver fat through hepatocyte cholesterol and bile acid metabolism. Activation of the transcriptional enhancer harbouring the putative causal variant of this locus upregulates *CYP7A1* in a human hepatic cell line. Furthermore, using zebrafish larvae, we show that perturbation of *cyp7a1* results in a lower liver fat content under metabolically challenging conditions, providing functional evidence of its role in MASLD pathogenesis.

MATERIALS AND METHODS

This research complies with the principles outlined in the Declaration of Helsinki. The UK Biobank received ethical approval from the National Research Ethics Service Committee North West Multi-Centre Haydock (reference 16/NW/0274). Data used in this study were obtained under application number 37142. The NEO study was approved by the medical ethical committee of the Leiden University Medical Center. The Liver BIBLE study was approved by the ethical committee of the Fondazione IRCCS Ca' Granda (ID 1650, revision 23 June 2020). The MAFALDA study was approved by the Local Research Ethics Committee (no. 16/20). The Helsinki cohort protocols were approved by the ethics committee of the Hospital District of Helsinki and Uusimaa (Helsinki, Finland).

Materials and methods are available online.

RESULTS

Genome-wide interactions with BMI for alanine aminotransferase

We first examine genome-wide interactions with BMI for

circulating ALT levels, a commonly used clinical biomarker for MASLD. To this end, we include data from 378,264 individuals of European ancestry in UK Biobank in whom ALT levels, but not liver fat based on proton density fat fraction (PDFF), are available. We examined interactions for 9,356,431 genetic variants with minor allele frequency (MAF) >0.01 . Variants at 13 loci interact with BMI for ALT ($P<5E-8$, Fig. 1). Four of these were not previously reported to interact with BMI for ALT (*GIPR*, *HLA*, *COBLL1*, *DPM3*).¹⁰ Of the 13 loci, seven show significant associations with ALT in all three BMI groups (lean and normal weight, overweight, and obese individuals), with larger effect sizes in individuals with a higher BMI (Fig. 1). This suggests elevated adiposity exacerbates the genetic predisposition for MASLD at these loci. While earlier results in a subset ($n=319,882$) of UK Biobank participants suggest the *UBXN2B/CYP7A1* locus is only associated with ALT in obese individuals,¹⁰ our results show this locus shows significant but oppositely directed associations with ALT in overweight and obese individuals on the one hand, and lean and normal-weight individuals on the other hand. In line with this interaction, the lead single nucleotide polymorphism (SNP) between *UBXN2B* and *CYP7A1* (rs7826120) does not show a main effect on ALT ($P=0.11$). Ten of the other 12 loci—all except *HLA-B* and *GIPR*—also show main effects on ALT (Table 1).

Interactions with BMI and main effects on proton density fat fraction

Of the 13 loci interacting with BMI for ALT, 6 loci (*PNPLA3*, *GPAM*, *UBXN2B/CYP7A1*, *HLA*, *APOE*, *TRIB1*) show directionally consistent trends for interactions with BMI for liver fat content assessed using PDFF in 35,146 independent UK Biobank participants (Fig. 1). Consistent with the interaction for ALT, the association between PDFF and rs7826120 (*UBXN2B/CYP7A1*) is stronger in individuals with a higher BMI (Fig. 1). Data from all 35,146 individuals with PDFF-assessed liver fat content reveal previously unanticipated associations with PDFF for the *GIPR* and *UBXN2B/CYP7A1* loci (Fig. 2).

We next used data from 3,931 individuals of four independent European cohorts to replicate the genetic association of liver fat content with the *UBXN2B/CYP7A1* locus. Of these, the Netherlands Epidemiology of Obesity Study

(NEO) ($n=1,822$) and Helsinki ($n=497$) studies have hepatic fat measurements from proton magnetic resonance spectroscopy, while the MAFALDA ($n=468$) and Liver-BIBLE ($n=1,144$) studies have data on controlled attenuation parameter values. After meta-analysis, the rs7826120 T allele is also associated with higher liver fat content in these independent cohorts (Fig. 3, $P=8.8E-3$ for both fixed- and random-effects models).

Fine mapping of independent loci

After excluding the *HLA-B* locus on chromosome 6, we performed functionally informed fine-mapping of the 12 remaining loci using summary statistics of the genotype-BMI interactions for ALT. The size of the first 95% credible set ranged from 1 to 65 variants per locus, with one variant at the *TRIB1* (rs2954021) and *TOR1B* (rs7029757) loci at a posterior inclusion probability (PIP) >0.94 , suggesting these are the causal variants at these two loci. SuSiE identified two 95% credible sets for the *APOE* and *GIPR* loci, suggesting the presence of multiple causal variants (Supplementary Table 1).¹¹ We did not identify credible sets at the *DPM3* and *COBLL1* loci, possibly reflecting a high purity filter ($r^2<0.25$). The *COBLL1* locus has been functionally implicated in increased risk of type 2 diabetes while decreasing adiposity, a presentation reminiscent of lipodystrophy.^{12,13} Among previously identified genetic variants associated with MASLD, *TM6SF2* rs187429064 and *PNPLA3* rs738409—both missense variants—are likely causal. Of note, when comparing the joint main and interaction term (2-df) P -values at the *TM6SF2* locus, the lead variant is the well-known E167K-encoding variant rs58542926, with a P_{2DF} of $1.20E-110$, compared with $1.06E-106$ for rs73001065. Moreover, the lead variants in *APOE*, *MARC1*, *TRIB1*, and *TOR1B*—recently reported in association with ALT and MASLD^{7,14-19}—have the highest PIP at their respective locus. In the *DPM3* locus, rs9330264 shows a higher PIP than the lead variant. Notably, when repeating the genome-wide-environment interaction study (GEWIS) using all available ALT data, rs9330264 is the lead variant. This SNP is in moderate linkage disequilibrium (LD) ($D' 0.96$, $r^2 0.31$) with the previously identified ALT lead SNP rs12904 in the 3' UTR of *EFNA1*. While *DPM3* is ubiquitously expressed, *EFNA1* expression is highest in the liver.^{20,21} Taken together, among the 12 loci

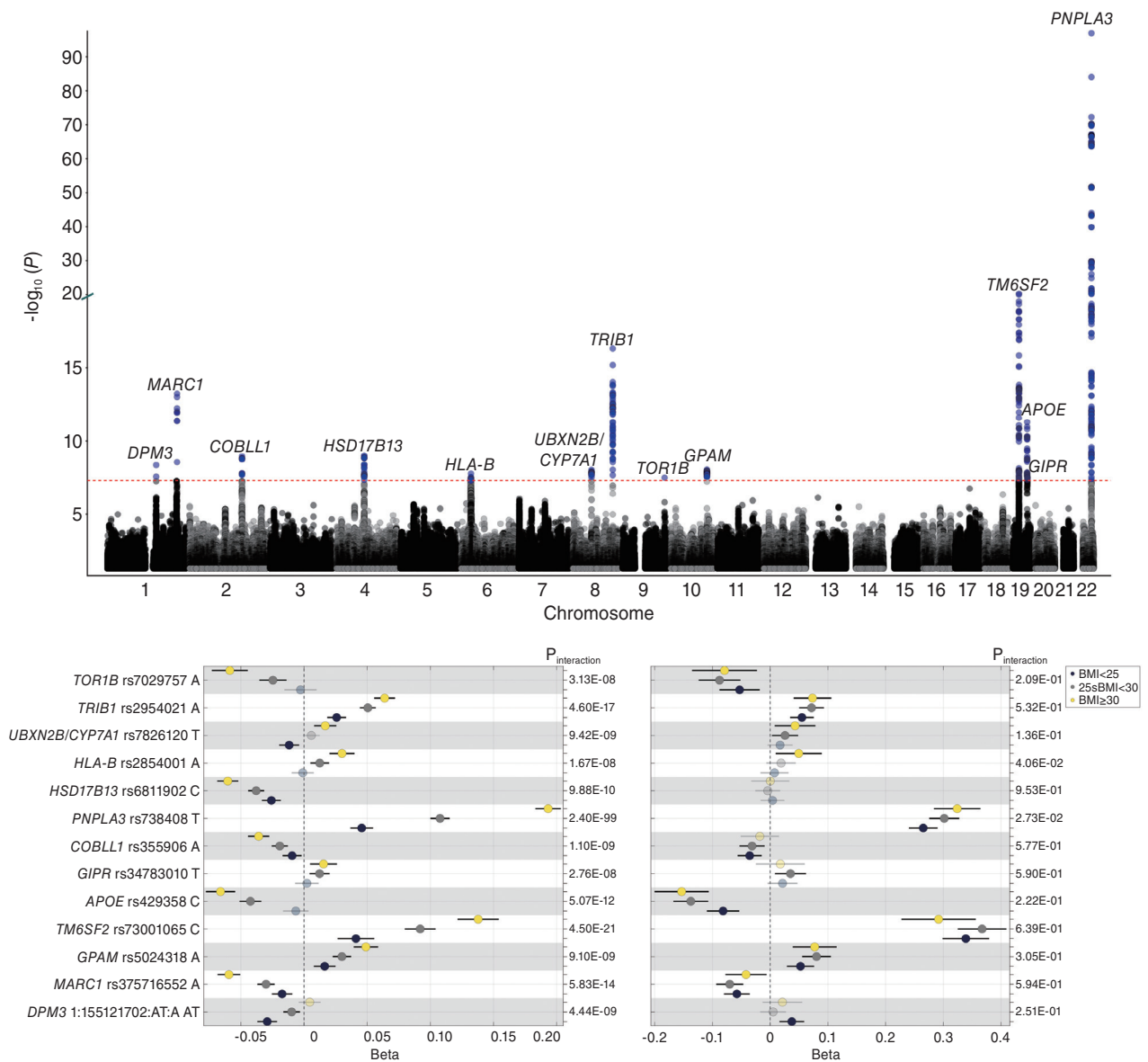


Figure 1. 13 loci interact with BMI for ALT. Top: Manhattan plot of genome-wide interaction analysis with BMI for ALT in European ancestry participants (UK Biobank). *P*-values were calculated by using a whole-genome regression model in REGENIE. Red dashed line represents the genome-wide significance level, 5×10^{-8} . Bottom: association of ALT (left) and PDFF (right) with 13 loci that interact with BMI for ALT stratified by BMI. Associations were examined using linear regression analysis adjusted for BMI, age, sex, age² and age² × sex, first 10 genomic principal components and array batch. For ALT, the association analyses have been performed after excluding individuals with PDFF data. The x-axis represents beta coefficients, with error bars showing 95% confidence intervals. Greyed-out circles represent associations with a nominal *P*-value > 0.05. The SNP-BMI interaction *P*-values are displayed as a secondary y-axis on the right-hand side of each panel. ALT, alanine aminotransferase; BMI, body mass index; PDFF, proton density fat fraction; SNP, single nucleotide polymorphism.

identified in the GEWIS, eight fine-mapped variants in *PNPLA3*, *TM6SF2*, *MARC1*, *GPAM*, *APOE*, *TRIB1*, *COBLL1*, *TOR1B* have previously been robustly associated with ALT and liver fat content.^{7,14-19}

To examine whether putative causal variants at the

UBXN2B/CYP7A1, *DPM3/EFNA1*, and *GIPR* loci influence the expression of nearby genes, we performed a Bayesian colocalization analysis between the novel GEWIS loci and expression quantitative trait locus (eQTL) summary statistics of liver and adipose tissue from the genotype-tissue

Table 1. GEWIS with BMI for ALT in the UK Biobank

Chr	Pos	Variant ID	Consequence	Beta _{GxE}	SE _{GxE}	A1 freq	A2	A1(effect)	P _{GxE}	P _{2DF}	P _{Cond}	P _{SNP}	Locus
1	155121702	1:155121702_AT_A	downstream_ gene_variant	0.003	4.8E-04	0.464	A	AT	4.44E-09	4.13E-16	1.42E-08	1.45E-11	DPM3/EFNA1*
1	220973761	rs375716552	intron_variant	-0.004	5.0E-04	0.313	AGC	A	5.83E-14	1.43E-55	7.22E-47	2.14E-07	MARC1
2	165642448	rs3355906	intron_variant	-0.003	4.7E-04	0.439	G	A	1.10E-09	6.06E-29	2.36E-23	7.36E-06	COBLL1
4	88213884	rs6811902	intergenic_ variant	-0.003	4.6E-04	0.437	T	C	9.88E-10	1.22E-79	9.15E-76	1.58E-03	HSD17B13
6	31323012	rs2854001	splice_ polypyrimidine_ tract_variant	0.003	5.4E-04	0.232	G	A	1.67E-08	4.54E-13	1.27E-07	1.81E-06	HLA-B
8	59371725	rs7826120	intergenic_ variant	0.003	4.9E-04	0.333	C	T	9.42E-09	3.22E-08	1.13E-01	4.47E-08	UBXN2B/ CYP7A1*
8	126482077	rs2954021	intron_variant	0.004	4.6E-04	0.494	G	A	4.60E-17	4.27E-112	2.17E-102	8.06E-07	TRIB1
9	132566666	rs7029757	non_coding_ transcript_ exon_variant	-0.004	7.8E-04	0.095	G	A	3.13E-08	2.59E-17	2.16E-12	1.15E-05	TOR1B
10	113933009	rs5024318	intron_variant	0.003	5.4E-04	0.247	T	A	9.10E-09	1.90E-37	1.95E-33	1.32E-04	GPAM
19	19460541	rs73001065	intron_variant	0.009	9.4E-04	0.071	G	C	4.50E-21	1.06E-106	1.48E-98	6.73E-10	TM6SF2
19	45411941	rs429358	missense_ variant	-0.004	6.3E-04	0.157	T	C	5.07E-12	3.43E-43	3.41E-35	9.99E-07	APOE
19	46180414	rs34783010	intron_variant	0.003	5.8E-04	0.193	G	T	2.76E-08	8.11E-10	3.50E-04	7.19E-07	GIPR*
22	44324730	rs738408	synonymous_ variant	0.012	5.7E-04	0.216	C	T	2.40E-99	2.22E-307	<4.94E-324	1.55E-47	PNPLA3

The GEWIS for ALT of genetic variants×BMI was performed using REGENIE, adjusting for age, sex, age², agexsex, age²×sex, first 10 genomic principal components and array batch. Beta and standard errors correspond to the interaction term (GxE), and P_{2DF} shows the two degree of freedom test of joint effect of interaction and main effects. P_{SNP} represents the main effect P-value in the GEWIS analysis. P_{cond} indicates the conditional P-value adjusted only for BMI and other covariates (corresponding to a typical GWAS). Loci * represent the novel loci identified here. The column Locus shows the nearest gene to the index variant (from a COJO analysis). ALT, alanine aminotransferase; BMI, body mass index; Chr, Chromosome; GEWIS, genome-wide-environment interaction study; GWAS, genome-wide association study; Pos, Position (GRCh37).

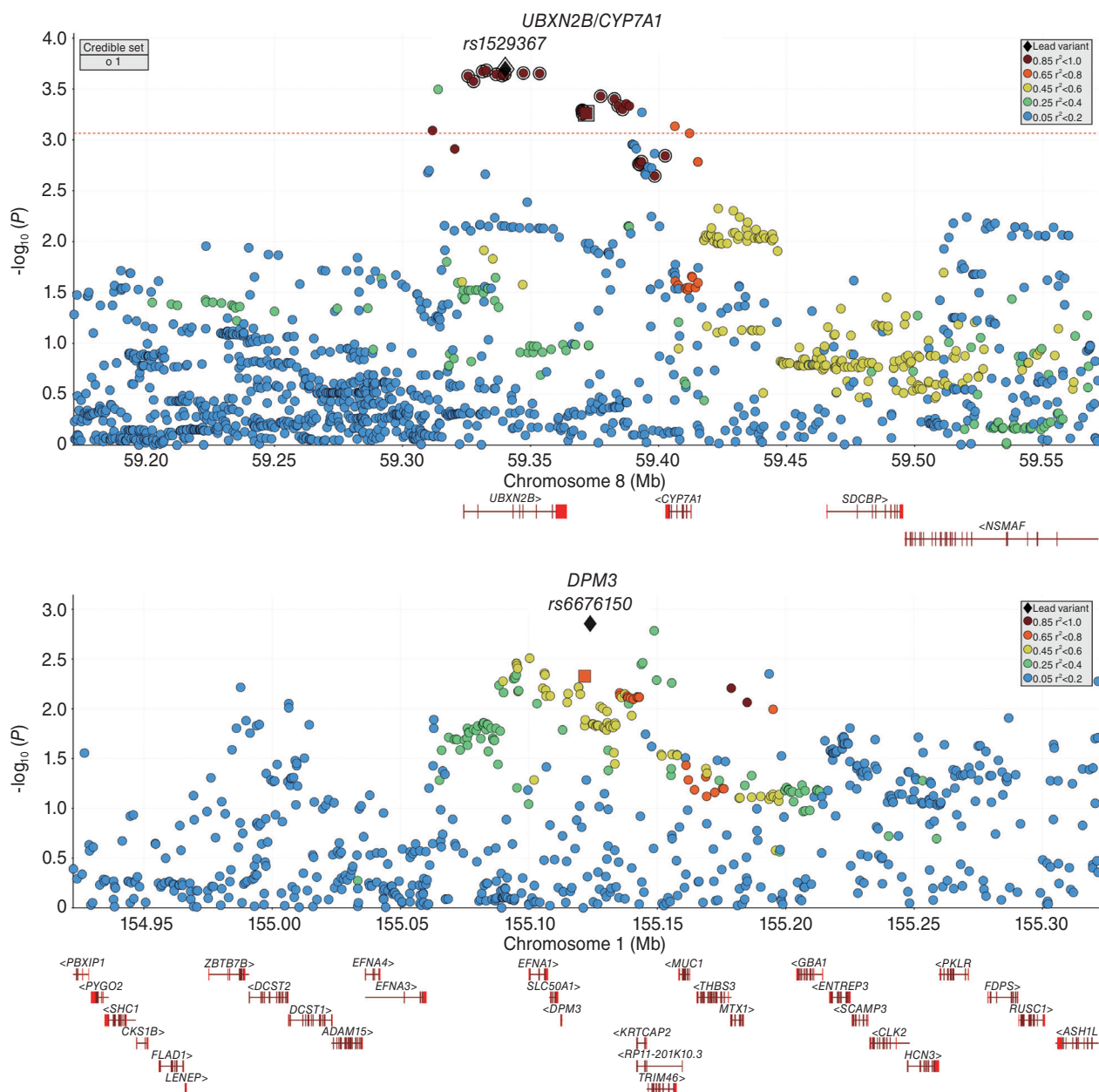


Figure 2. Regional plots for association of 3 new loci with PDFF. For each locus, a window of ± 200 kb around lead variants from GEWIS on ALT was considered. The lead variants from the GEWIS for ALT and GWAS on PDFF are marked with a square and black diamond, respectively. Credible set indicates putative causal variants from the GEWIS for ALT. Red dashed line represents the FDR threshold using the Benjamini-Hochberg method. Data from a total of 337,000 unrelated White-British participants from UK Biobank were used for in-sample LD structure. ALT, alanine aminotransferase; GEWIS, genome-wide-environment interaction study; GWAS, genome-wide association study; PDFF, proton density fat fraction; FDR, false discovery rate; LD, linkage disequilibrium.

expression (GTEx) project (v.8) uniformly processed by eQTL Catalogue (Supplementary Table 2).²²⁻²⁴ We do not observe evidence of shared causal variants between eQTL summary statistics and GEWIS signals at the *GIPR* locus. At the *UBXN2B/CYP7A1* locus, we observe colocalization

between interaction effects and eQTL signals in liver for *UBXN2B* (H4.PP=0.975), but not *CYP7A1*. In addition, we observe colocalization with the nearby *THBS3* (H4.PP=0.98) and *MUC1* (H4.PP=0.89) for the *DPM3/EFNA1* locus.

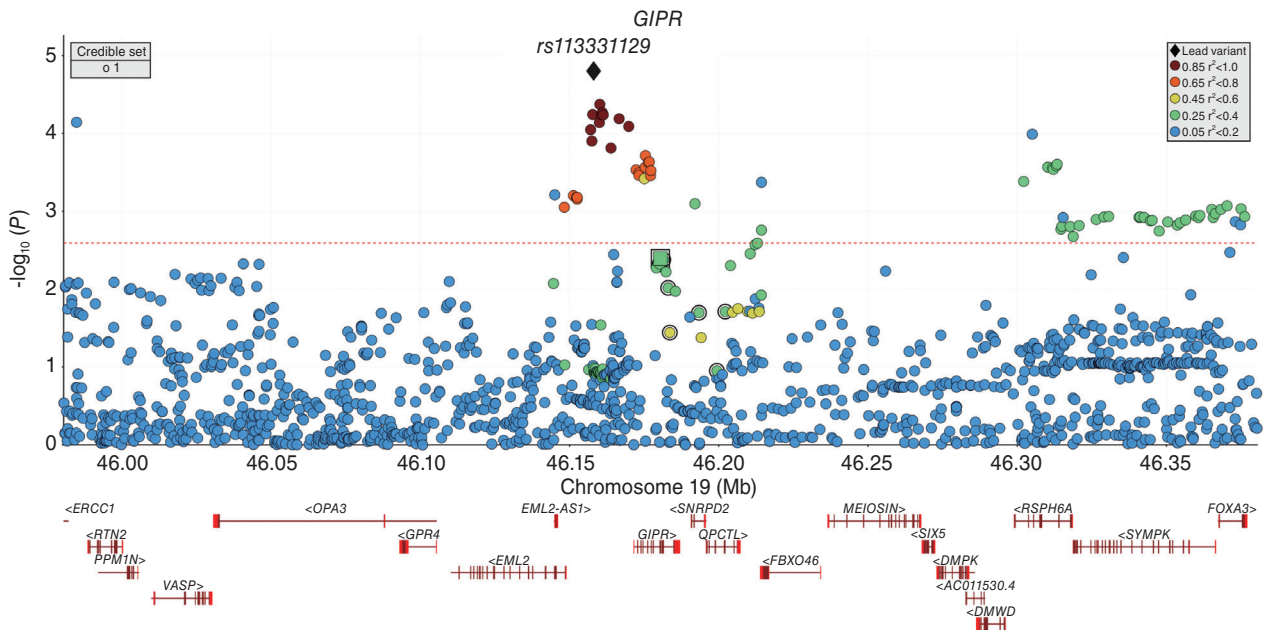


Figure 2. Continued.

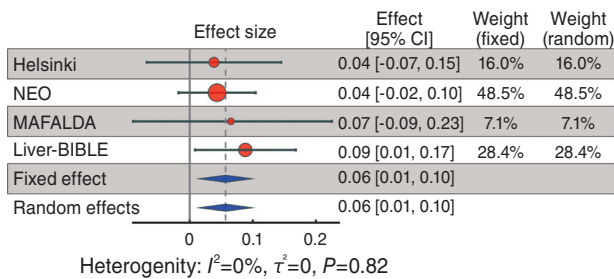


Figure 3. Forest plot for association of controlled attenuation parameter and magnetic resonance spectroscopy hepatic fat content with the *UB2N2B/CYP7A1* rs7826120 T allele in four European replication cohorts. The association was examined by a linear regression analysis under an additive genetic model adjusted for age, sex, and BMI. Pooled effect estimates were calculated using inverse-variance-weighted fixed- and random-effects meta-analysis. I^2 , τ^2 (between-study variance) and P -value for Cochran's Q heterogeneity test have been reported to assess the between-study heterogeneity. BMI, body mass index; CI, confidence interval; MAFALDA, Molecular Architecture of FAtty Liver Disease in patients with obesity undergoing bariatric surgery; NEO, the Netherlands Epidemiology of Obesity Study.

At the *UB2N2B/CYP7A1* locus, the lead variant of the PDFF GWAS is in complete LD with the GEWIS lead variant ($r^2=1$ in Europeans), and within the same credible set. In contrast, both lead variants at the *COBLL1* and *GIPR* loci are in weak LD ($r^2=0.16$ and 0.25 in Europeans), and belong to different credible sets, suggesting that the mech-

anisms for main and interaction effects may be different at these loci. The regional plots for other previously established MASLD-associated loci show a consistent pattern between putative causal variants of GEWIS and GWAS for PDFF (Supplementary Fig. 1).

GIPR locus

We show that the T allele of the intronic rs34783010 in *GIPR* is associated with higher PDFF. Tirzepatide, a dual GLP1R and GIPR agonist, was recently approved by the Food and Drug Administration and European Medical Agency to treat type 2 diabetes with obesity. Therefore, we performed a phenome-wide association analysis (214 traits, Supplementary Tables 3, 4) for rs34783010 in *GIPR*. In addition to being associated with higher PDFF, HbA1c and odds of diabetes, the *GIPR* rs34783010 T allele is also associated with lower BMI and with lower non-fasting glucose concentrations (Fig. 4). Carriers of the *GIPR* rs34783010 T allele also have higher odds of Alzheimer's disease ($P\text{-FDR}<0.05$, Fig. 4 and Supplementary Table 4). The *GIPR* rs34783010 T allele has been previously associated with lower levels of gastric inhibitory polypeptide (GIP) protein in plasma ($\beta=-0.079$, $P=5.0E-6$).²⁵ A recent phase two study showed that tirzepatide effectively re-

solves MASH and reduces liver fibrosis.²⁶ We do not observe associations of *GIPR* rs34783010 with chronic liver disease or liver cirrhosis (Supplementary Table 5).

***UBXN2B/CYP7A1* locus**

Each minor T allele at rs7826120 is associated with higher PDFF and higher odds of chronic liver disease, but not with cirrhosis (Supplementary Table 5). Further, the T allele is associated with higher circulating triglyceride and low-density lipoprotein (LDL) cholesterol levels and higher odds of dyslipidaemia, gallstone, and cardiovascular diseases²⁷ and lower odds of inflammatory bowel disease (Fig. 5A and Supplementary Table 6). Moreover, T allele carriers have higher cholesteryl ester in LDL and higher levels of very low-density lipoproteins (VLDL) particles (Fig. 5A and Supplementary Table 6). Using genome-wide summary statistics of the human plasma metabolome,²⁸ we show the rs7826120 T allele is associated with levels of nine metabolites ($P < 5E-8$): six annotated metabolites (five lower, one higher) and three unnamed compounds. Interestingly, all annotated metabolites are involved in primary and secondary bile acid metabolism (Supplementary Table 7). The trigger for these associations can be seen as an interaction with metabolic burden that requires further study to unravel the complex interplay between rs7826120/*CYP7A1* and bile acid synthesis, demand, composition and metabolism; dietary fat intake; cholesterol levels; inflammation; and oxidative stress.

To gain further insights into the molecular mechanism underlying the association observed at this locus, we integrated the credible set with human liver epigenomic datasets. This revealed a single variant overlapping a liver cis-regulatory element (CRE), rs10504255 (Fig. 5B). The rs10504255 G allele—which co-segregates with the PDFF-increasing rs7826120-T allele—has also been associated with higher risk of intrahepatic cholestasis of pregnancy²⁹ and acute coronary syndrome.³⁰ None of the other variants in the credible set, including the index lead variant rs7826120, overlap a liver CRE. rs10504255 resides in a liver enhancer, showing accessible chromatin—detected by assay for transposase-accessible chromatin using sequencing (ATAC-seq) and enrichment in the active CRE histone mark H3K27ac across a series of samples (Supplementary Fig. 2A, Supplementary Table 8). In the liver, this enhancer

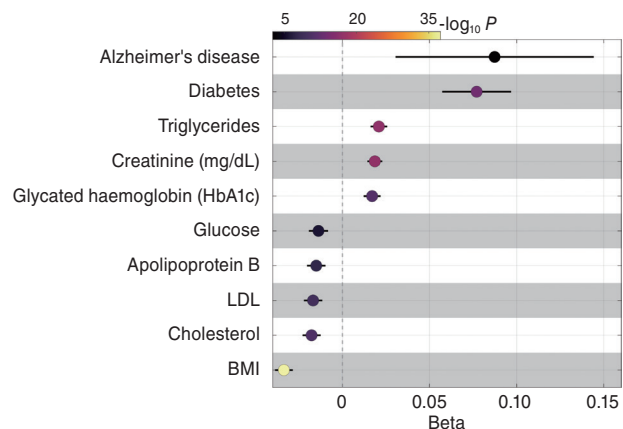


Figure 4. The association of a set of metabolic biomarkers and relevant diseases with the *GIPR* rs34783010 T allele in the UK Biobank. The association was examined by a linear or logistic regression analysis under an additive genetic model adjusted for BMI, age, sex, age \times sex, age² and age² \times sex, the first 10 genomic principal components, and array batch. For binary traits, log odds of effects are shown. The colour represents the $-\log_{10}$ -transformed P -values. BMI, body mass index; LDL, low-density lipoprotein.

is specifically and reproducibly accessible in adult and foetal hepatocytes, but not other cell lineages (Fig. 5B, Supplementary Fig. 2A).

To further interpret the functional impact of rs10504255, we carried out transcription factor (TF) motif analysis using MotifBreakR to compare the binding affinity profiles of the two predominant rs10504255 alleles. Motifs were filtered for strong allele effects and a minimal expression in liver and hepatocytes of at least one transcript per million (Supplementary Table 9). The presence of the minor G allele—associated with higher PDFF—increases the affinity for several TFs with repressive activity, including REST ($P=5.0E-4$) and FOXA1 ($P=6.3E-4$) (Fig. 5C, Supplementary Fig. 2B, Supplementary Table 9). On the other hand, the A allele is predicted to preferentially bind TFs with an established role in MASLD-associated phenotypes like aberrant lipid metabolism and inflammation, including HNF1A ($P=8.1E-4$) and PPARG ($P=7.0E-4$) (Fig. 5C, Supplementary Fig. 2B, Supplementary Table 9).³¹ Inspection of publicly available Chromatin ImmunoPrecipitation (ChIP)-seq datasets for human liver confirms the binding of several of the predicted TFs, including FOXA1 and PPARG. The G allele may therefore be associated with decreased activity of the enhancer leading to *CYP7A1* deficiency.

As mentioned, the index variant rs7826120 only colocalizes

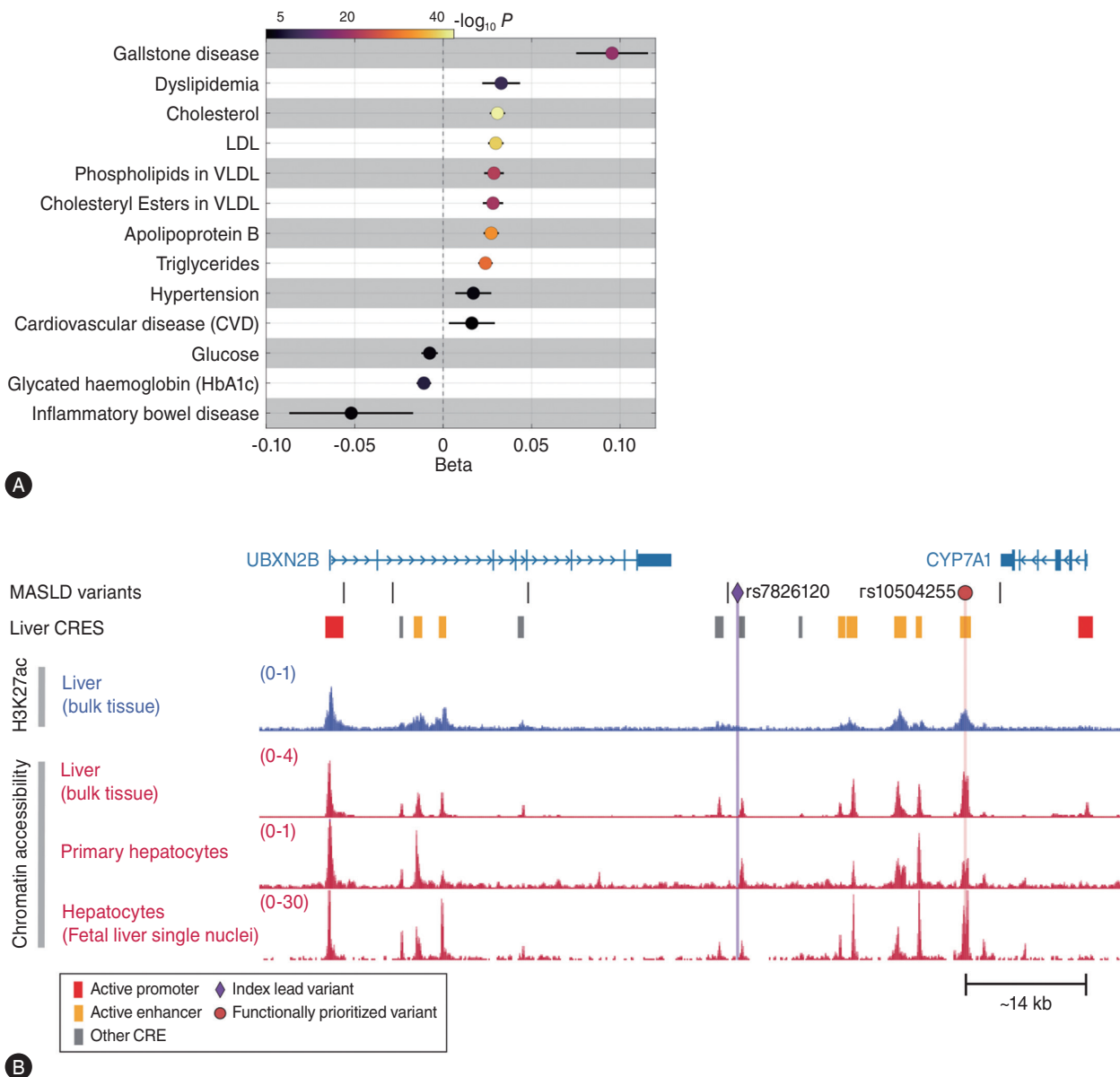
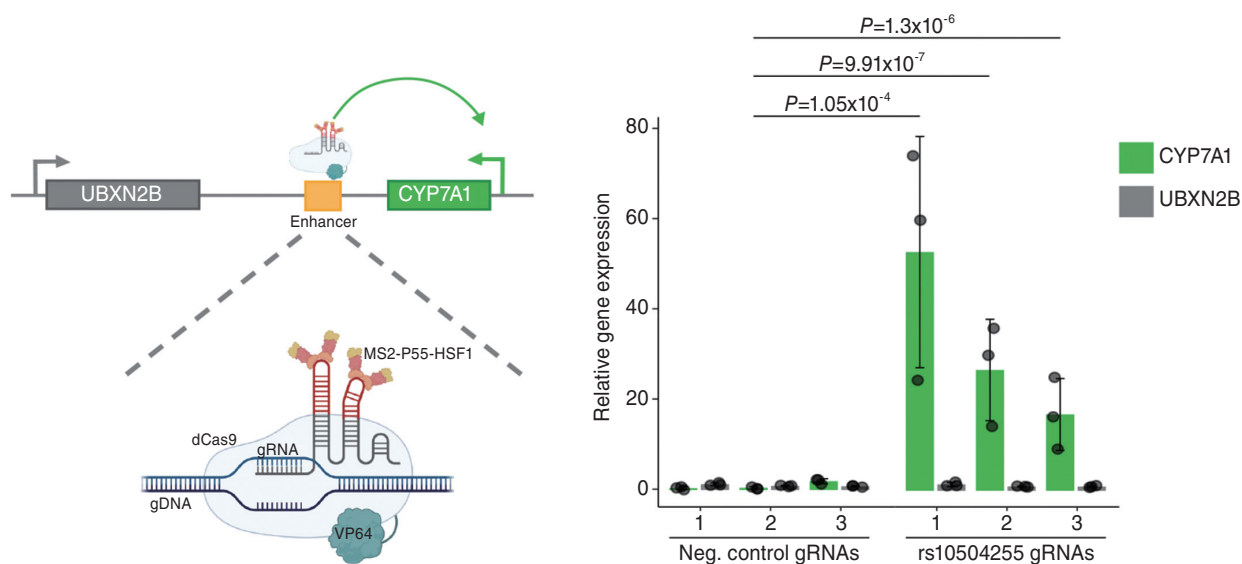


Figure 5. The association of relevant traits with the rs7826120 T allele (UK Biobank). (A) Associations were examined by additive linear or logistic regression analyses adjusted for BMI, age, sex, agexsex, age² and age²xsex, first 10 genomic principal components and array batch. For binary traits, log odds of effects are shown. The colour represents the $-\log_{10}$ -transformed P -values. (B) Overview of the metabolic-dysfunction associated steatotic liver disease (MASLD) locus *UBXN2B/CYP7A1*. Of all MASLD variants in the fine-mapped credible set at this locus, only rs10504255 (red circle) resides within an active liver *cis*-regulatory element (CRE). The lead variant (purple diamond) resides adjacent to an open chromatin region not enriched for the active CRE histone mark H3K27ac. Tracks show pooled normalised signal for H3K27ac chromatin immunoprecipitation sequencing (ChIP-seq) and chromatin accessibility detected by ATAC-seq across 4–6 independent samples. (C) MotifbreakR results showing position weight matrices (PWMs) of transcription factor (TF) motifs predicted to be affected by the variant rs10504255. PWMs matching the G allele, associated with higher proton density fat fraction, are shown at the top, and those matching the A allele are shown at the bottom. TFs with $P < 1E-3$ were prioritised as likely hits if they were expressed in both liver tissue and hepatocytes and are shown ranked top to bottom for each allele by significance. *Indicates TFs for which there is ChIP-seq evidence of binding at the enhancer element. (D) Schematic (left) of CRISPRa complex targeting the enhancer at the *UBXN2B/CYP7A1* locus. Bar plot (right) of relative expression detected by RT-qPCR of *CYP7A1* and *UBXN2B* upon CRISPRa targeting of the enhancer vs. negative control guides. Data are presented as the mean \pm standard deviation of biological replicates ($n=3$ independent transductions) and statistically analysed by Student's t -test.



C



D

Figure 5. Continued.

with an eQTL signal for *UBXN2B* (Supplementary Table 7). The same was observed for rs10504255 (GTEx: Liver - *UBXN2B* $P=2.4E-7$, Liver - *CYP7A1* $P=0.14$). Importantly, the degree of concordance between GWAS and *cis*-eQTLs is inherently limited.³² Therefore, we used CRISPR-mediated activation (CRISPRa) to identify the most likely target of the liver

enhancer hosting rs10504255 (Supplementary Table 10). CRISPRa leverages CRISPR/Cas9 technology to achieve nuclease-free targeting of specific sites in the genome, recruiting transcriptional activators to induce activation of target CREs or genes. We found that activation of the enhancer in Hep3B cells upregulates *CYP7A1* ($P=9.91E-7$ for

gRNA 2) and does not affect *UBXN2B* expression ($P=0.66$ across all 3 gRNAs, Fig. 5D, Supplementary Table 11). These results indicate that *CYP7A1* is the likely target gene of the enhancer harbouring rs10504255.

CYP7A1 encodes cytochrome P450 family 7 subfamily A member 1, which catalyses the first step of hepatic cholesterol catabolism and bile acid synthesis. In line with the associations described here and the known biological function of *CYP7A1*, three individuals homozygous for loss-of-function mutations in *CYP7A1* have been described to have 70% higher cholesterol levels in liver without fibrosis or inflammation, reduced bile acid secretion, elevated circulating LDL cholesterol levels that are resistant to statins, hypertriglyceridemia, and premature gallstone disease. One of three individuals had premature coronary and peripheral artery disease.³³ Together, these observations strongly point to *CYP7A1* as the causal gene in the locus. We hypothesize that the opposite direction of rs7826120's association with ALT in overweight and obese versus lean and normal-weight individuals reflects an interaction between *CYP7A1*-driven bile acid synthesis and the metabolic and/or inflammatory state of the liver. In lean individuals, increased *CYP7A1* activity may be neutral or beneficial, promoting cholesterol clearance and metabolic flexibility, supporting liver health, and resulting in lower ALT. In obesity on the other hand, it may exacerbate liver stress, inflammation, or injury and elevate ALT due to altered bile acid handling and increased hepatic vulnerability.

Functional experiments in zebrafish larvae

To further explore the role of *CYP7A1* in MASLD, we used CRISPR/Cas9-based gene editing, automated fluorescence imaging, and deep learning-based image analysis in zebrafish larvae. Both zebrafish orthologues of *CYP7A1* (*cyp7a1* and *CR354540.1*) were targeted simultaneously (Supplementary Table 12) and larvae with/without such mutations were fed on one of four diets from day 5 to 10 (feeding 3x more yes/no x4% extra dietary cholesterol yes/no). At day 10, we imaged 100 larvae across the four dietary conditions, with 23 to 28 larvae per condition. Of these, 57 larvae carried fragment length-confirmed CRISPR/Cas9-induced mutations in both *CYP7A1* orthologues, and 43 were sibling controls free from such mutations. The proportion of mutated larvae was similar across all four di-

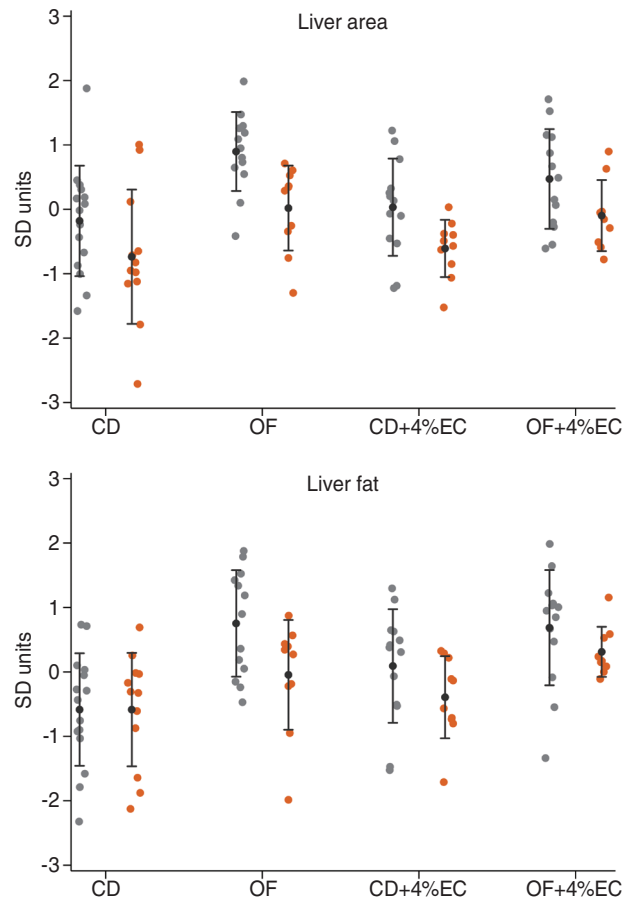


Figure 6. The effect of CRISPR/Cas9-induced mutations in both orthologues of *CYP7A1* on liver area and liver fat area in 10-day-old zebrafish larvae stratified by dietary condition. 4%EC, diet enriched with 4% extra cholesterol; CD, control diet; OF, overfeeding (3x more); SD, standard deviation. Orange, larvae carrying CRISPR/Cas9-induced mutations in both *CYP7A1* orthologues; grey, sibling controls free from such mutations.

etary conditions, ranging from 39% to 46%.

On average, larvae with CRISPR/Cas9-induced mutations in *cyp7a1* and *CR354540.1* are 5% shorter and have a 16% smaller liver with 19% less liver fat than their sibling controls. A linear regression analysis shows that when adjusting for time of day at imaging and whether or not larvae had been metabolically challenged, larvae with CRISPR/Cas9-induced mutations in *cyp7a1* and *CR354540.1* are 0.54 ± 0.18 standard deviation (SD) units shorter ($P=3.6E-3$), with a 0.63 ± 0.16 SD unit smaller liver ($P=1.8E-4$) than their sibling controls (Supplementary Table 13). For liver fat area, we observe a trend for an interaction of mutations in *cyp7a1* and *CR354540.1* with the presence/absence of a metabolic challenge (-0.62 ± 0.38 SD units, $P=0.10$,

Supplementary Table 14) that is not observed for liver or body size ($P > 0.2$). In metabolically challenged larvae, those with mutations in *cyp7a1* and *CR354540.1* have 0.55 ± 0.20 SD units less liver fat than their sibling controls ($P = 8.1E-3$). This effect is not observed in unchallenged larvae (0.15 ± 0.32 SD units, $P = 0.66$, Fig. 6, Supplementary Table 14). Given *CYP7A1*'s role as a critical regulatory enzyme of bile acid biosynthesis and cholesterol homeostasis in humans³³ and zebrafish,³⁴ it seems likely that mutations in the gene directly affect hepatic lipid accumulation, and that effects on liver and body size are secondary.

DISCUSSION

Here we performed a GEWIS with BMI for circulating ALT levels. Thirteen *loci* show evidence of such interactions. By excluding individuals in whom PDFF measurements are available from the primary GEWIS, we were able to subsequently examine the association of hits with liver fat content in a set of independent individuals. Two previously unanticipated *loci*, *GIPR*, and *CYP7A1*, show such associations with liver triglyceride content, complemented by the previously described *COBLL1* locus.⁶ Additionally, while a BMI-SNP interaction for ALT has been reported earlier for the *CYP7A1* locus,¹⁰ our larger sample size sheds new light on the nature of this interaction and expands it by showing that the locus is also associated with PDFF. *In vitro* and *in vivo* experiments pinpoint *CYP7A1* and the conversion of cholesterol into bile acid as relevant for MASLD.

An important finding of this study is the identification of a genetic variant at the *GIPR* locus, encoding the receptor of the incretin GIP. The rs34783010 T allele is associated with lower BMI, and with worse glycaemic control, higher odds of diabetes, and higher liver fat content, reminiscent of lipodystrophy. Murine models show inconsistent results after manipulation of *GIPR*. Mice treated with a long acting GIPR agonist have lower food intake, resulting in lower BMI and improved insulin sensitivity. This is consistent with the notion of GIPR agonism being favourable against diabetes.³⁵ Counterintuitively, *Gipr* knockout mice have a lower body mass, reminiscing receptor agonism.³⁵ However, a different mouse strain, susceptible to cardiovascular disease, does not show change in body mass and glucose metabolism when treated with a GIPR agonist.³⁶ Taken together, it is

difficult to reconcile the human genetics and *in vivo* experiments in mice.

Circulating levels of GIP were lower in carriers of the *GIPR* rs34783010 T allele. This suggests an effect on the function of the receptor or on hormone clearance from the circulation, even if the specific impact of the variant on *GIPR* expression and activity in specific tissues—including the gut, adipose tissue and the central nervous system—may be different. Hence, our results are consistent with an overall positive effect of the *GIPR* rs34783010 G allele, protecting against diabetes and MASLD despite higher BMI. This resembles the effect of pioglitazone, a peroxisome proliferator-activated receptor gamma agonist.

Interestingly, the *CYP7A1* locus would not have been discovered in a GWAS focusing on main effects on ALT. The credible set in this locus is intergenic, and hence it is not immediately obvious which gene encodes the effector transcript. However, epigenomic analyses and experiments, common variant associations of the lead SNP, and clinical presentation in three individuals homozygous for loss-of-function mutations all point to *CYP7A1*. This gene encodes an enzyme that catalyses the rate limiting step in bile acid synthesis, by converting cholesterol to 7- α -hydroxycholesterol. Fine mapping shows that a variant ~14 kb downstream of the *CYP7A1* transcription start site (rs10504255) has the highest posterior probability of being causal. Moreover, integration of a series of epigenomic datasets revealed that rs10504255 resides within a liver CRE, likely an enhancer element, given its location and enrichment in the classic enhancer mark H3K27ac. In humans, we observe associations of the rs7826120 T allele with higher PDFF—especially in individuals with BMI ≥ 30 kg/m²—higher cholesteryl ester content, lower levels of primary and secondary bile acid metabolites, and higher risk of gallstone and cardiovascular disease, suggesting a shift in cholesterol metabolism from bile acid synthesis to hepatic lipid accumulation and VLDL secretion. These associations are reminiscent of an underlying loss of *CYP7A1* function,³³ or reduced expression, concordant with the prediction that the rs10504255 G allele—co-segregating with the rs7826120 T allele—increases the binding affinity of the *bona fide* transcriptional repressor REST and of TFs that are context-dependent repressors, like BCL3, FOXA1 and RORA. In the liver, RORA controls lipid homeostasis via negative regulation of PPARG,³⁷ while FOXA1 acts as a re-

pressor through competition for binding targets with FOXA2.³⁸ Moreover, the rs10504255 G allele is predicted to disrupt the binding of PPARG and HNF1A, two transcriptional activators that are key regulators of hepatic lipid homeostasis, including the regulation of cholesterol and bile acid metabolism.³⁹⁻⁴¹ The G allele also disrupts binding of BATF, which has been shown to ameliorate hepatic steatosis in mice challenged with high-fat diet.⁴² Based on our results, namely the CRISPRa of CRE hosting rs10504255 and the *in silico* motif disruption analyses, we postulate that the rs10504255 G allele promotes the repression of a transcriptional enhancer, leading to decreased expression of *CYP7A1*.

At odds with the direction of associations in humans but consistent with results in C57BL/6J mice,⁴³ CRISPR/Cas9-induced mutations in *CYP7A1* orthologues result in less liver fat in metabolically challenged zebrafish larvae. Several factors may contribute to these differences between species. First, mutations present from conception inherently induce compensatory adaptations from early embryonic development onwards. In line with this, human probands have twice higher *CYP27A1* activity than controls.³³ In humans, this upregulation does not compensate for loss of *CYP7A1*.³³ Zebrafish, however, have four *CYP27A1* orthologues, so overcompensation in response to loss of functional *cyp7a1* by upregulation of *CYP27A1* orthologues—and possibly other *CYPs*—seems plausible. Secondly, differences in adiposity between adult humans and zebrafish larvae may play a role. Adipocyte differentiation starts around day 8 in zebrafish larvae, resulting in very small lipid depots in only 24% of overfed and 8% of control fed larvae by day 10.⁴⁴ Hepatic bile acids are key metabolic hormones with the ability to modulate glucose and fatty acid metabolism in adipose tissues and anti-obesogenic potential.⁴⁵ On the other hand, adipokines secreted by white adipose tissue, such as Retnla, regulate hepatic cholesterol metabolism, inducing *Cyp7a1*.⁴⁶ Hence, the sparsity of adipose tissue in 10-day-old zebrafish larvae may influence the net effect of perturbations affecting bile acid synthesis on hepatic steatosis. Finally, it remains challenging to satisfactorily appreciate the nuances between association and perturbation, across outcomes, and across species. What humans and zebrafish share is that *CYP7A1* influences MASLD-related traits under metabolically challenging conditions.

Our findings have clinically important implications. Indeed, the presence of an interaction between these loci and BMI for ALT levels suggests that weight loss may differentially affect liver fat content depending on genetic background. In line with this, carriers of the *PNPLA3* rs738409 G allele have a larger reduction of liver fat content after bariatric surgery compared with non-carriers.⁴⁷ In conclusion, by exploiting genotype-BMI interactions for ALT levels, we have identified two loci associated with liver fat content, namely in *GIPR* and near *CYP7A1*. Through downstream functional perturbation experiments in zebrafish, we demonstrate that *CYP7A1* plays a role in MASLD. Our data provide genetic validation that affecting the conversion of cholesterol into bile acids influences the burden of MASLD and support the use of *GIPR* agonism as a therapeutic strategy against diabetes and MASLD.

Data Availability

All individual-level phenotype/genotype data from UK Biobank are accessible upon approval of applications to the UK Biobank at <http://www.ukbiobank.ac.uk>. Summary statistics of this study are available upon request. REGENIE software can be found at <https://rgcgithub.github.io/regenie/>. Epigenomic datasets generated as part of this study are available in EGAD50000001751.

Authors' contribution

OJ, SR, and MdH conceived and designed the study. OJ conducted the main analyses in the UK Biobank. EM and MdH conducted and interpreted the functional experiments in zebrafish. SFQ, SM, FM, RLG, LR, FT, UVG, FRR, HYJ, and LV contributed to the replication in the independent cohorts. LM, HM, AA, AE, and IC conducted and interpreted the epigenomic functional studies. OJ, SR, MdH, LV, EC, and RMM contributed to curating data and results. SR, OJ, MdH, LV, and IC wrote the initial draft of the paper. All authors contributed to and approved the final version of the paper.

Acknowledgements

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

The authors thank all individuals who participated in the NEO study, and all general practitioners for inviting eligible

participants. The authors also thank the NEO study group, Pat van Beelen, Petra Noordijk and Ingeborg de Jonge for the coordination, lab and data management of the NEO study.

Epigenomic data analysis was performed using the Imperial College Research Computing Service (RCS, DOI:10.14469/hpc/2232). Computations in the zebrafish experiments were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) through the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under project SNIC 2022/22-235.

Handling and storage of zebrafish imaging data were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725.

S.R. was supported by the Swedish Cancerfonden (22 2270 Pj), the Swedish Research Council (Vetenskapsrådet (VR), 2023-02079), the Swedish state under the Agreement between the Swedish government and the county councils (the ALF agreement, ALFGBG-965360), the Swedish Heart Lung Foundation (20220334), the Wallenberg Academy Fellows from the Knut and Alice Wallenberg Foundation (KAW 2017.0203), the Novonordisk Distinguished Investigator Grant - Endocrinology and Metabolism (NNF23OC0082114), the Novonordisk Project grants in Endocrinology and Metabolism (NNF20OC0063883). MdH is supported by the Swedish Heart-Lung Foundation (20230518), the Swedish Research Council (2023-02556), and the NIH/NIDDK-funded Accelerating Medicines Partnership for Common Metabolic Disorders (5UM1DK105554-5000826-5500002718). LV was supported by grants from the Italian Ministry of Health (Ministero della Salute, RF-2016-02364358, PNRR-MAD-2022-12375656), the Ricerca Corrente Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Fondazione IRCCS Ca' Granda "Liver BIBLE" (no. PR-0391), The European Union, (H2020-ICT-2018-20/H2020-ICT-2020-2 under grant agreement No. 101016726 - REVEAL, HORIZON-MISS-2021-CANCER-02-03 programme "Genial" under grant agreement "101096312"), Italian ministry of Research (MUR) (PNRR - M4 - C2 "ASSET", PRIN 2022 "DEFENDER"). SFQ was supported by grants from the Orion Research Foundation, the Yrjö Jahansson Foundation

(20207313), the Maud Kuistila Memorial Foundation (2021-0301B), the Emil Aaltonen Foundation (210182), the Finnish Medical Foundation (5843), and the Biomedicum Helsinki Foundation (20230241). HYJ was supported by the Academy of Finland (309263), the Sigrid Jusélius Foundation, and the Novo Nordisk Foundation (NNF19OC0057503). RL is supported by JPI HDHL-DIYUFOOD. LM was supported by a NC3Rs PhD studentship (NC/Y500628/1). IC is recipient of a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (224662/Z/21/Z). This work was also supported by the NIH/NIDDK-funded Accelerating Medicines Partnership for Type 2 Diabetes (AMP T2D, RFP16 awarded to IC).

Conflicts of Interest

SR has served as a consultant for AstraZeneca, Ribocure, Foresite Labs and Sanofi in the last 3 years. SR has received research grants from AstraZeneca in the last 3 years. MdH co-founded a contract research organization (Veyviser A/B) that offers target identification and validation as a service. LV received speaking fees from GSK and Gilead, served as a consultant for Gilead, Pfizer, AstraZeneca, Novo Nordisk, and received research grants from Gilead. OJ is a part-time consultant to Ribocure AB. IC has received research grants from Gilead and the Novo Nordisk Foundation and her partner has stock options in Ochre Bio. The remaining authors declare no competing interests.

SUPPLEMENTARY MATERIAL

Supplementary material is available at Clinical and Molecular Hepatology website (<http://www.e-cmh.org>).

REFERENCES

1. Harrison SA, Bedossa P, Guy CD, Schattenberg JM, Loomba R, Taub R, et al. A Phase 3, randomized, controlled trial of resmetirof in NASH with liver fibrosis. *N Engl J Med* 2024;390:497-509.
2. Powell EE, Wong VW, Rinella M. Non-alcoholic fatty liver disease. *Lancet* 2021;397:2212-2224.
3. Bianco C, Jamialahmadi O, Pelusi S, Baselli G, Dongiovanni P, Zanon I, et al. Non-invasive stratification of hepatocellular

- carcinoma risk in non-alcoholic fatty liver using polygenic risk scores. *J Hepatol* 2021;74:775-782.
4. Du X, DeForest N, Majithia AR. Human Genetics to identify therapeutic targets for NAFLD: challenges and opportunities. *Front Endocrinol (Lausanne)* 2021;12:777075.
 5. Sveinbjornsson G, Ulfarsson MO, Thorolfsdottir RB, Jonsson BA, Einarsson E, Gunnlaugsson G, et al. Multiomics study of nonalcoholic fatty liver disease. *Nat Genet* 2022;54:1652-1663.
 6. Vujkovic M, Ramdas S, Lorenz KM, Guo X, Darlay R, Cordell HJ, et al. A multiancestry genome-wide association study of unexplained chronic ALT elevation as a proxy for nonalcoholic fatty liver disease with histological and radiological validation. *Nat Genet* 2022;54:761-771.
 7. Jamialahmadi O, Mancina RM, Ciociola E, Tavaglione F, Luukkonen PK, Baselli G, et al. Exome-wide association study on alanine aminotransferase identifies sequence variants in the GPAM and APOE associated with fatty liver disease. *Gastroenterology* 2021;160:1634-1646.e1637.
 8. Stender S, Kozlitina J, Nordestgaard BG, Tybjaerg-Hansen A, Hobbs HH, Cohen JC. Adiposity amplifies the genetic risk of fatty liver disease conferred by multiple loci. *Nat Genet* 2017;49:842-847.
 9. Romeo S, Sentinelli F, Dash S, Yeo GS, Savage DB, Leonetti F, et al. Morbid obesity exposes the association between PNPLA3 I148M (rs738409) and indices of hepatic injury in individuals of European descent. *Int J Obes (Lond)* 2010;34:190-194.
 10. Gao C, Marcketta A, Backman JD, O'Dushlaine C, Staples J, Ferreira MAR, et al. Genome-wide association analysis of serum alanine and aspartate aminotransferase, and the modifying effects of BMI in 388k European individuals. *Genet Epidemiol* 2021;45:664-681.
 11. Zou Y, Carbonetto P, Wang G, Stephens M. Fine-mapping from summary data with the "Sum of Single Effects" model. *PLoS Genet* 2022;18:e1010299.
 12. Glunk V, Laber S, Sinnott-Armstrong N, Sobreira DR, Strobel SM, Batista TM, et al. A non-coding variant linked to metabolic obesity with normal weight affects actin remodelling in subcutaneous adipocytes. *Nat Metab* 2023;5:861-879.
 13. Maude H, Cebola I. Zooming into process-specific risk. *Nat Metab* 2023;5:730-731.
 14. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2008;40:1461-1465.
 15. Kozlitina J, Smagris E, Stender S, Nordestgaard BG, Zhou HH, Tybjaerg-Hansen A, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to non-alcoholic fatty liver disease. *Nat Genet* 2014;46:352-356.
 16. Fairfield CJ, Drake TM, Pius R, Bretherick AD, Campbell A, Clark DW, et al. Genome-wide association study of NAFLD using electronic health records. *Hepatol Commun* 2022;6:297-308.
 17. Ghodsian N, Abner E, Emdin CA, Gobeil É, Taba N, Haas ME, et al. Electronic health record-based genome-wide meta-analysis provides insights on the genetic architecture of non-alcoholic fatty liver disease. *Cell Rep Med* 2021;2:100437.
 18. Liu Y, Bastly N, Witcher B, Bell JD, Sorokin EP, van Bruggen N, et al. Genetic architecture of 11 organ traits derived from abdominal MRI using deep learning. *Elife* 2021;10.
 19. 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.
 20. Chen VL, Du X, Chen Y, Kuppa A, Handelman SK, Vohnoutka RB, et al. Genome-wide association study of serum liver enzymes implicates diverse metabolic and liver pathology. *Nat Commun* 2021;12:816.
 21. Emdin CA, Haas M, Ajmera V, Simon TG, Homburger J, Neben C, et al. Association of genetic variation with cirrhosis: a multi-trait genome-wide association and gene-environment interaction study. *Gastroenterology* 2021;160:1620-1633.e1613.
 22. Kerimov N, Hayhurst JD, Peikova K, Manning JR, Walter P, Kolberg L, et al. A compendium of uniformly processed human gene expression and splicing quantitative trait loci. *Nat Genet* 2021;53:1290-1299.
 23. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 2020;369:1318-1330.
 24. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* 2014;10:e1004383.
 25. Pietzner M, Wheeler E, Carrasco-Zanini J, Cortes A, Koprulu M, Wörheide MA, et al. Mapping the proteo-genomic convergence of human diseases. *Science* 2021;374:eabj1541.
 26. Loomba R, Hartman ML, Lawitz EJ, Vuppalanchi R, Boursier J, Bugianesi E, et al. Tirzepatide for metabolic dysfunction-associated steatohepatitis with liver fibrosis. *N Engl J Med* 2024;391:299-310.

27. Qayyum F, Lauridsen BK, Frikke-Schmidt R, Kofoed KF, Nordestgaard BG, Tybjaerg-Hansen A. Genetic variants in CYP7A1 and risk of myocardial infarction and symptomatic gallstone disease. *Eur Heart J* 2018;39:2106-2116.
28. Surendran P, Stewart ID, Au Yeung VPW, Pietzner M, Raffler J, Wörheide MA, et al. Rare and common genetic determinants of metabolic individuality and their effects on human health. *Nat Med* 2022;28:2321-2332.
29. Dixon PH, Levine AP, Cebola I, Chan MMY, Amin AS, Aich A, et al. GWAS meta-analysis of intrahepatic cholestasis of pregnancy implicates multiple hepatic genes and regulatory elements. *Nat Commun* 2022;13:4840.
30. Vargas-Alarcón G, Pérez-Méndez Ó, Posadas-Sánchez R, González-Pacheco H, Luna-Luna M, Escobedo G, et al. Associations of the CYP7A1 gene polymorphisms located in the promoter and enhancer regions with the risk of acute coronary syndrome, plasma cholesterol, and the incidence of diabetes. *Biomedicines* 2024;12.
31. Gross B, Pawlak M, Lefebvre P, Staels B. PPARs in obesity-induced T2DM, dyslipidaemia and NAFLD. *Nat Rev Endocrinol* 2017;13:36-49.
32. Mostafavi H, Spence JP, Naqvi S, Pritchard JK. Systematic differences in discovery of genetic effects on gene expression and complex traits. *Nat Genet* 2023;55:1866-1875.
33. Pullinger CR, Eng C, Salen G, Shefer S, Batta AK, Erickson SK, et al. Human cholesterol 7 α -hydroxylase (CYP7A1) deficiency has a hypercholesterolemic phenotype. *J Clin Invest* 2002;110:109-117.
34. Enya S, Kawakami K, Suzuki Y, Kawaoka S. A novel zebrafish intestinal tumor model reveals a role for cyp7a1-dependent tumor-liver crosstalk in causing adverse effects on the host. *Dis Model Mech* 2018;11.
35. Liskiewicz A, Khalil A, Liskiewicz D, Novikoff A, Grandl G, Maity-Kumar G, et al. Glucose-dependent insulinotropic polypeptide regulates body weight and food intake via GABAergic neurons in mice. *Nat Metab* 2023;5:2075-2085.
36. Sachs S, Götz A, Finan B, Feuchtinger A, DiMarchi RD, Döring Y, et al. GIP receptor agonism improves dyslipidemia and atherosclerosis independently of body weight loss in preclinical mouse model for cardio-metabolic disease. *Cardiovasc Diabetol* 2023;22:217.
37. Kim K, Boo K, Yu YS, Oh SK, Kim H, Jeon Y, et al. ROR controls hepatic lipid homeostasis via negative regulation of PPAR transcriptional network. *Nat Commun* 2017;8:162.
38. Duncan SA, Navas MA, Dufort D, Rossant J, Stoffel M. Regulation of a transcription factor network required for differentiation and metabolism. *Science* 1998;281:692-695.
39. Li T, Chiang JY. Regulation of bile acid and cholesterol metabolism by PPARs. *PPAR Res* 2009;2009:501739.
40. Patitucci C, Couchy G, Bagattin A, Cañeque T, de Reyniès A, Scoazec JY, et al. Hepatocyte nuclear factor 1 suppresses steatosis-associated liver cancer by inhibiting PPAR transcription. *J Clin Invest* 2017;127:1873-1888.
41. Jung D, Kullak-Ublick GA. Hepatocyte nuclear factor 1 α : a key mediator of the effect of bile acids on gene expression. *Hepatology* 2003;37:622-631.
42. Zhang Z, Liao Q, Pan T, Yu L, Luo Z, Su S, et al. BATF relieves hepatic steatosis by inhibiting PD1 and promoting energy metabolism. *Elife* 2023;12.
43. Ferrell JM, Boehme S, Li F, Chiang JY. Cholesterol 7 α -hydroxylase-deficient mice are protected from high-fat/high-cholesterol diet-induced metabolic disorders. *J Lipid Res* 2016;57:1144-1154.
44. Mazzaferro E, Mujica E, Zhang H, Emmanouilidou A, Jenseit A, Evcimen B, et al. Functionally characterizing obesity-susceptibility genes using CRISPR/Cas9, in vivo imaging and deep learning. *Sci Rep* 2025;15:5408.
45. Thomas C, Pellicciari R, Pruzanski M, Auwerx J, Schoonjans K. Targeting bile-acid signalling for metabolic diseases. *Nat Rev Drug Discov* 2008;7:678-693.
46. Lee MR, Lim CJ, Lee YH, Park JG, Sonn SK, Lee MN, et al. The adipokine Retnla modulates cholesterol homeostasis in hyperlipidemic mice. *Nat Commun* 2014;5:4410.
47. Krawczyk M, Jiménez-Agüero R, Alustiza JM, Emparanza JI, Perugorria MJ, Bujanda L, et al. PNPLA3 p.I148M variant is associated with greater reduction of liver fat content after bariatric surgery. *Surg Obes Relat Dis* 2016;12:1838-1846.