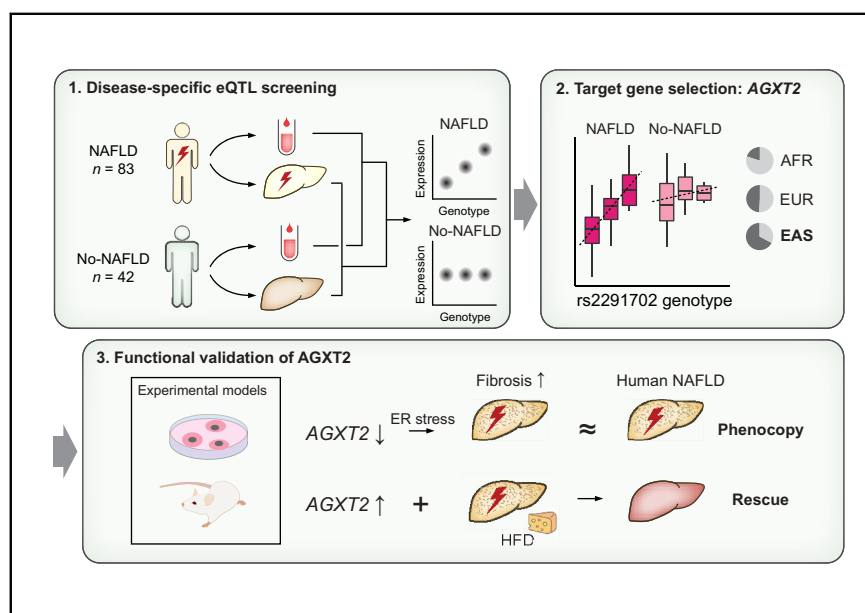


# Disease-specific eQTL screening reveals an anti-fibrotic effect of AGXT2 in non-alcoholic fatty liver disease

## Graphical abstract



## Authors

Taekyeong Yoo, Sae Kyung Joo, Hyo Jung Kim, ..., Jae Woo Kim, Won Kim, Murim Choi

## Correspondence

[drwon1@snu.ac.kr](mailto:drwon1@snu.ac.kr) (W. Kim),  
[murimchoi@snu.ac.kr](mailto:murimchoi@snu.ac.kr) (M. Choi).

## Lay summary

Elucidating causal genes for non-alcoholic fatty liver disease (NAFLD) has been challenging due to limited tissue availability and the polygenic nature of the disease. Using liver and blood samples from 125 Korean individuals (83 with NAFLD and 42 without NAFLD), we devised a new analytic method to identify causal genes. Among the candidates, we found that *AGXT2*-rs2291702 protects against liver fibrosis in a genotype-dependent manner with the potential for therapeutic interventions. Our approach enables the discovery of causal genes that act on the basis of genotype.

## Highlights

- Our 'response-eQTL' approach aimed to discover novel SNP-gene pairs that only function in NAFLD.
- NAFLD-specific repression of *AGXT2* was prominent in rs2291702:CC carriers.
- Lower *AGXT2* expression was associated with worse histological and metabolic features in rs2291702:CC carriers.
- The reduction of *AGXT2* mimicked human NAFLD features in mice, whereas overexpression rescued them.
- The reduced *AGXT2* caused increased cell death due to ER stress activation in HepG2 cells.

# Disease-specific eQTL screening reveals an anti-fibrotic effect of AGXT2 in non-alcoholic fatty liver disease

Taekyeong Yoo<sup>1</sup>, Sae Kyung Joo<sup>2</sup>, Hyo Jung Kim<sup>3</sup>, Hyun Young Kim<sup>4</sup>, Hyungtai Sim<sup>1</sup>, Jieun Lee<sup>2</sup>, Hee-Hoon Kim<sup>5</sup>, Sunhee Jung<sup>6</sup>, Youngha Lee<sup>1</sup>, Oveis Jamialahmadi<sup>7</sup>, Stefano Romeo<sup>7,8,9</sup>, Won-Il Jeong<sup>5</sup>, Geum-Sook Hwang<sup>6,10</sup>, Keon Wook Kang<sup>4</sup>, Jae Woo Kim<sup>3</sup>, Won Kim<sup>2,\*</sup>, Murim Choi<sup>1,\*</sup>, On behalf of the Innovative Target Exploration of NAFLD (ITEN) consortium

<sup>1</sup>Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea; <sup>2</sup>Department of Internal Medicine, Seoul National University College of Medicine, Seoul Metropolitan Government Boramae Medical Center, Seoul, Republic of Korea; <sup>3</sup>Department of Biochemistry and Molecular Biology, Brain Korea 21 PLUS Project for Medical Science, Yonsei University College of Medicine, Seoul, Republic of Korea; <sup>4</sup>College of Pharmacy and Research Institute of Pharmaceutical Sciences, Seoul National University, Seoul, Republic of Korea; <sup>5</sup>Laboratory of Liver Research, Graduate School of Medical Science and Engineering, KAIST, Daejeon, Republic of Korea; <sup>6</sup>Integrated Metabolomics Research Group, Western Seoul Center, Korea Basic Science Institute, Seoul, Republic of Korea; <sup>7</sup>Salhgremska Academy, Institute of Medicine, Department of Molecular and Clinical Medicine, University of Gothenburg, Sweden; <sup>8</sup>Salhgremska University Hospital, Cardiology Department, Sweden; <sup>9</sup>Department of Medical and Clinical Science, Clinical Nutrition Unit, University Magna Graecia, Catanzaro, Italy; <sup>10</sup>Department of Chemistry & Nanoscience, Ewha Womans University, Seoul, Republic of Korea

**Background & Aims:** Non-alcoholic fatty liver disease (NAFLD) poses an increasing clinical burden. Genome-wide association studies have revealed a limited contribution of genomic variants to the disease, requiring alternative but robust approaches to identify disease-associated variants and genes. We carried out a disease-specific expression quantitative trait loci (eQTL) screen to identify novel genetic factors that specifically act on NAFLD progression on the basis of genotype.

**Methods:** We recruited 125 Korean patients (83 with biopsy-proven NAFLD and 42 without NAFLD) and performed eQTL analyses using 21,272 transcripts and 3,234,941 genotyped and imputed single nucleotide polymorphisms. We then selected eQTLs that were detected only in the NAFLD group, but not in the control group (i.e., NAFLD-eQTLs). An additional cohort of 162 Korean individuals with NAFLD was used for replication. The function of the selected eQTL toward NAFLD development was validated using HepG2, primary hepatocytes and NAFLD mouse models.

**Results:** The NAFLD-specific eQTL screening yielded 242 loci. Among them, AGXT2, encoding alanine-glyoxylate aminotransferase 2, displayed decreased expression in patients with NAFLD homozygous for the non-reference allele of rs2291702, compared to no-NAFLD individuals with the same genotype ( $p = 4.79 \times 10^{-6}$ ). This change was replicated in an additional 162 individuals, yielding a combined  $p$  value of  $8.05 \times 10^{-8}$  from a total of 245 patients with NAFLD and 42 controls. Knockdown of AGXT2 induced palmitate-overloaded hepatocyte death by increasing endoplasmic reticulum stress, and exacerbated NAFLD

diet-induced liver fibrosis in mice, while overexpression of AGXT2 attenuated liver fibrosis and steatosis.

**Conclusions:** We identified a new molecular role for AGXT2 in NAFLD. Our overall approach will serve as an efficient tool for uncovering novel genetic factors that contribute to liver steatosis and fibrosis in patients with NAFLD.

**Lay summary:** Elucidating causal genes for non-alcoholic fatty liver disease (NAFLD) has been challenging due to limited tissue availability and the polygenic nature of the disease. Using liver and blood samples from 125 Korean individuals (83 with NAFLD and 42 without NAFLD), we devised a new analytic method to identify causal genes. Among the candidates, we found that AGXT2-rs2291702 protects against liver fibrosis in a genotype-dependent manner with the potential for therapeutic interventions. Our approach enables the discovery of causal genes that act on the basis of genotype.

© 2021 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved.

## Introduction

Non-alcoholic fatty liver disease (NAFLD) is a growing burden that affects approximately a quarter of the world's population and contributes to liver-related morbidity.<sup>1,2</sup> Defined as a condition in which excess liver fat exists in the absence of secondary causes of lipid accumulation or clinically significant alcohol intake, NAFLD includes a spectrum of liver diseases ranging from non-alcoholic fatty liver (NAFL) to non-alcoholic steatohepatitis (NASH).<sup>1,3,4</sup> Due to unmet needs in the prediction and early detection of NAFLD, there have been continuous efforts at clarifying its pathomechanism to facilitate the identification of novel therapeutic targets and biomarkers. However, no pharmacotherapy has yet been approved for NAFLD.<sup>5,6</sup>

Genome-wide association studies (GWAS) have revealed loci that confer risk for NAFLD.<sup>1,4,7,8</sup> However, these signals demonstrate modest effect sizes and account for only a minor fraction of the overall heritability of NAFLD, which is estimated to range at

Keywords: NAFLD; NASH; eQTL; genetic variants; AGXT2.

Received 24 August 2020; received in revised form 22 March 2021; accepted 7 April 2021; available online xxx

\* Corresponding authors. Addresses: 103 Daehak-ro, Jongno-gu, Seoul, Republic of Korea. Tel: +82-2-740-8912. Fax: +82-2-3673-2167. (M. Choi), or Boramae-ro 5-gil, Dongjak-gu, Seoul, Republic of Korea. Tel: +82-2-870-2233. Fax: +82-2-831-2826. (W. Kim).

E-mail addresses: drwon1@snu.ac.kr (W. Kim), murimchoi@snu.ac.kr (M. Choi).

<https://doi.org/10.1016/j.jhep.2021.04.011>



ELSEVIER

between 22–50%.<sup>4,9</sup> For complex traits, the development of a polygenic risk score (PRS) has led to promising results in risk prediction.<sup>7</sup> However, PRS evaluation of NAFLD has not yet generated robust results.<sup>10,11</sup>

Mapping of expression quantitative trait loci (eQTLs) enables the identification of genetic variants that are associated with gene expression changes.<sup>12,13</sup> eQTL analysis has the advantage of providing interpretable molecular links between genetic variants and traits of interest.<sup>14,15</sup> These links also enable a substantial increase in statistical power, such that thousands of eQTLs can be detected even with a sample of just ~100 individuals.<sup>14,16</sup>

Alanine-glyoxylate aminotransferase 2 (AGXT2) is a mitochondrial aminotransferase, possessing multiple enzymatic activities on a wide array of substrates, including asymmetric dimethylarginine (ADMA) and 3-amino-isobutyrate (BAIB), and produces diverse metabolites including dimethylguanidino valeric acid (DMGV).<sup>17</sup> Knockout mouse and human-based studies have implicated AGXT2 in endothelial dysfunction, hypertension, and chronic heart failure, as ADMA is a potent inhibitor of nitric oxide synthase (NOS).<sup>18</sup> GWAS of BAIB and DMGV levels in urine and serum indicated a strong association with AGXT2 variants, highlighting its central role in regulating the levels of these molecules.<sup>19–21</sup> The gene is specifically expressed in the kidney and liver, and associated eQTLs have been identified in the liver tissues.<sup>22</sup> However, the molecular basis of its function in the liver remains elusive. Moreover, its pathophysiological role beyond the known enzymatic activity has not yet been clarified.

To circumvent existing technical limitations in understanding NAFLD, we performed eQTL mapping to identify genetic variants and their associated genes that confer susceptibility to NAFLD by collecting histologically confirmed liver tissue transcriptome and genotype data from 125 Korean individuals. We then developed a pipeline to select gene-eQTL pairs that are specifically active under the diseased state (hereafter referred to as “NAFLD-eQTLs”) and pinpointed a single nucleotide polymorphism (SNP) in the AGXT2 locus. With these efforts, we demonstrated that altered AGXT2 expression might modulate the progression of liver fibrosis via endoplasmic reticulum stress-mediated hepatocellular death. Overall, our results highlight a new approach for evaluating and selecting NAFLD-eQTLs, which can lead to the identification of novel therapeutic targets for NAFLD in an individual-specific manner.

## Materials and methods

### Patients

This study was approved by the institutional review board of Seoul Metropolitan Government Boramae Medical Center. We constructed a prospective cohort from the ongoing Boramae NAFLD registry (NCT 02206841) as previously described.<sup>23</sup> See the [supplementary methods](#) for eligibility and diagnostic criteria used in the study. Participants of both discovery (n = 125) and replication (n = 162) cohorts consisted of Korean individuals (Fig. S1), aged 19–80, who visited Seoul Metropolitan Government Boramae Medical Center. All participants were informed of the study protocol and provided written and signed consent. NAFLD activity scoring and fibrosis staging were performed following the Kleiner classification, and categorized into no-NAFLD, NAFL, and NASH (Table 1, Table S1 and S2).<sup>24</sup> In the subsequent analyses, we considered no-NAFLD as control and NAFL and NASH as the NAFLD group.

**Table 1. Summary statistics of the participants (n = 125).**

	No-NAFLD (n = 42)	NAFLD (n = 83)	p value
Age, years	56.7 (12.3)	53.3 (14.1)	0.19
Male, n (%)	25 (59.5)	38 (45.8)	0.15
BMI, kg/m <sup>2</sup>	24.0 (3.4)	28.1 (4.0)	1.21 × 10 <sup>-7</sup>
Fibrosis stage (0–4)	0.26 (0.5)	1.66 (1.0)	4.45 × 10 <sup>-19</sup>
Significant fibrosis (≥2) (%)	1 (2.4)	40 (48.2)	2.56 × 10 <sup>-7</sup>
NAFLD activity score	0.43 (0.6)	4.65 (1.3)	3.60 × 10 <sup>-49</sup>
SBP, mmHg	131.0 (16.5)	130.9 (18.5)	0.98
DBP, mmHg	79.4 (11.2)	79.9 (12.7)	0.84
HDL-cholesterol	46.4 (10.4)	44.7 (13.0)	0.45
Triglycerides, mg/dl	123.7 (54.8)	171.0 (88.0)	3.72 × 10 <sup>-4</sup>
AST, IU/L	29.9 (18.7)	67.5 (72.1)	1.91 × 10 <sup>-5</sup>
ALT, IU/L	30.3 (24.8)	80.5 (76.0)	2.51 × 10 <sup>-7</sup>
GGT, IU/L	49.4 (53.8)	66.7 (49.0)	8.72 × 10 <sup>-2</sup>
Albumin, g/dl	4.1 (0.3)	4.1 (0.3)	0.44
Platelet, ×10 <sup>9</sup> /L	228.7 (50.7)	224.4 (64.0)	0.68
HOMA-IR	2.7 (1.1)	6.2 (4.8)	9.11 × 10 <sup>-9</sup>
Adipo-IR	5.8 (4.5)	13.8 (11.4)	1.77 × 10 <sup>-7</sup>
Diabetes, n (%)	9 (21.4)	41 (49.4)	2.57 × 10 <sup>-3</sup>
Hypertension, n (%)	16 (38.1)	31 (37.3)	0.94

Values are given as mean (SD). p values are from independent t tests and  $\chi^2$  tests comparing between no-NAFLD and NAFLD groups.

Adipo-IR, adipose insulin resistance index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; DBP, diastolic blood pressure; GGT, gamma glutamyl-transferase; HOMA-IR, homeostasis model assessment of insulin resistance; NAFLD, non-alcoholic fatty liver disease.

### Transcriptome and genome data processing

Total RNA isolated from the liver was used for RNA sequencing on a HiSeq2500 platform. Reads were mapped and quantified with the human genome (hg19/GRCh37) based on GENCODE v19. Differentially expressed genes (DEGs) were called using the DESeq2 packages<sup>25</sup> with correction for sample batches. Genes within certain criteria were used in eQTL analysis (see [supplementary methods](#)).

DNA was acquired from blood and genotyped using an Illumina Infinium OmniExpress-24 kit or Omni2.5-8 kit. Genotype data were processed with exclusion criteria, matched with RNA-seq data and imputed using 1000 Genome Project Phase 3 haplotypes. After a further filtering process, genotyped and imputed calls were used in eQTL analysis (see [supplementary methods](#)).

### Cis-eQTL analysis

Genotype and gene expression data (21,272 genes and 3,234,941 genotyped and imputed SNPs) available from the 125 individuals were integrated for eQTL mapping using MatrixEQTL (version 2.2),<sup>26</sup> accounting for age and sex (or BMI and homeostatic model assessment of insulin resistance [HOMA-IR]). MatrixEQTL performed a linear regression on the transformed residuals with the corresponding imputed genotypes under the additive model. An eQTL within 1 Mb of a gene transcription start site (TSS) was considered a cis-eQTL. False discovery rate (FDR) was used to adjust for multiple testing.

### Selecting NAFLD-specific eQTLs (NAFLD-eQTLs)

To choose NAFLD-eQTLs, we divided samples into no-NAFLD (n = 42) and NAFLD (n = 83) groups as described above and performed cis-eQTL calling separately. Among the eQTLs from the NAFLD group, calls that were also found in the no-NAFLD group, with FDR-adjusted p < 0.05 excluded. To define NAFLD-specific loci, we selected the most significant eSNP per eGene, then kept only eQTLs with absolute  $\beta$  coefficient > 0.5 in NAFLD

and absolute  $\beta$  fold change (FC) between NAFLD and no-NAFLD  $>5$ . Finally, NAFLD-specific eGenes that overlapped with GTEx liver eGenes were excluded. GTEx participants free of NAFLD signatures ( $n = 79$ ) were used. RNA-seq and genotyping data from GTEx release v7 (<https://gtexportal.org>) were downloaded from the Database of Genotypes and Phenotypes (dbGaP) under accession phs000424.v7.p2.<sup>27</sup>

### Agxt2 overexpression and knockdown in mouse model

Male C57BL/J mice were purchased from Japan SLC (Shizuoka, Japan). Six-week-old mice were fed for 3 weeks, either a normal chow-diet for a normal control model or a choline-deficient, L-amino acid-defined, high-fat diet (CDAHFD, Research Diets, New Brunswick, NJ) for a NAFLD model. Then, mice were injected with adenoviruses containing short-hairpin RNA (shRNA) against *Agxt2* or mock, or adenoviruses expressing *Agxt2*-FLAG or GFP for knockdown and overexpression experiments, respectively. All procedures were performed under the standard protocols approved by the Committee on Animal Investigations of Yonsei University.

## Results

### Analysis of eQTLs

For transcriptome and genome-wide array analyses, liver biopsy and blood samples were acquired from 125 Korean individuals with varying metabolic and histological status. Using standardized pathological scores, we divided all samples into no-NAFLD ( $n = 42$ ) and NAFLD ( $n = 83$ , including NAFL and NASH individuals) groups for eQTL analysis (Fig. 1A and Table 1 and S1). After quality assessment of the genotype and transcriptome data, 21,272 transcripts and 3,234,941 SNPs were subject to *cis*-eQTL (i.e., SNP-gene pair within 1 Mb of gene TSS) mapping using an additive linear model. We identified 3,882 genes with *cis*-eQTLs at an FDR  $\leq 5\%$  (eGenes, Table S3), and 242,691 significant SNP-gene pairs from the set of 125 samples ("Liver-eQTLs"). We compared our liver eGenes to those from the GTEx database and found 40.1% of our eGenes (1,558/3,882) overlapped with GTEx liver eGenes, which was the largest proportion of overlap across all 48 GTEx tissues tested (Fig. S2A). Among the eQTL SNPs (eSNPs), 23.0% acted on multiple genes. As sample size may serve as a confounding factor for eQTL detection sensitivity, we compared our result to GTEx sets in terms of the correlation between sample numbers and eGene numbers. Our dataset fell within the range of GTEx correlations, suggesting that our data processing and eQTL calling were well-performed (Pearson's correlation  $R = 0.94$ ;  $p = 1.55 \times 10^{-24}$ ; Fig. S2B). In addition, we detected 2,577 significant *trans*-eQTLs at FDR  $\leq 5\%$ , of which 646 (25.1%) were also eQTLs for nearby genes. Among the *trans*-eQTLs, 30.0% predicted the expression of multiple non-local genes. We limited our subsequent analyses on *cis*-eQTLs as they have stronger and more direct implications on the regulation of target gene expression.

### Calling NAFLD-eQTLs

Under the hypothesis that *cis*-eQTLs may possess different activities under altered physiological status, such as NAFLD, we called eQTLs that are significantly associated in patients with NAFLD but not in the no-NAFLD group (i.e., NAFLD-eQTLs). Using the same sets of SNPs and genes as for the liver-eQTL calling as described above, 2,394 eGenes and 108,782 *cis*-eQTLs were detected specifically in the NAFLD group (Fig. 1B-1C). In comparison, 484 eGenes and 8,181 *cis*-eQTLs were detected as specific to the no-

NAFLD group. Like the liver-eQTLs, NAFLD-eQTLs clustered throughout the whole genome (Fig. S3). Both eQTL sets showed high enrichment of eSNPs in gene bodies and proximity to the gene starts and ends (Fig. S4A). Relative to intergenic regions, we also detected enrichments in untranslated regions, introns, and non-coding RNAs, which are putative functional regions (Fig. S4B). We additionally performed a genome-wide functional enrichment analysis of liver-eQTL and NAFLD-eQTL sets using GREGOR<sup>28</sup> and observed that eSNPs were significantly enriched in known transcription factor binding and histone modification sites (Fig. S4C and S4D). Enrichment in these regulatory regions provides further evidence that the eSNPs are functionally relevant to gene expression and regulation in the liver. The numbers of tissues that expressed our eGenes were comparable to those for GTEx liver eGenes (Fig. S4E). And in the liver, our eGenes displayed higher expression than non-eGenes ( $p < 2.2 \times 10^{-16}$ , Welch's *t*-test; Fig. S4F), suggesting a strong functional relevance. After additional filtering processes described in Fig. 1B and C, 242 NAFLD-eQTLs that alter expression in the NAFLD group but not in the no-NAFLD group, were selected for further analyses (Fig. 1B-1E and Table S4).

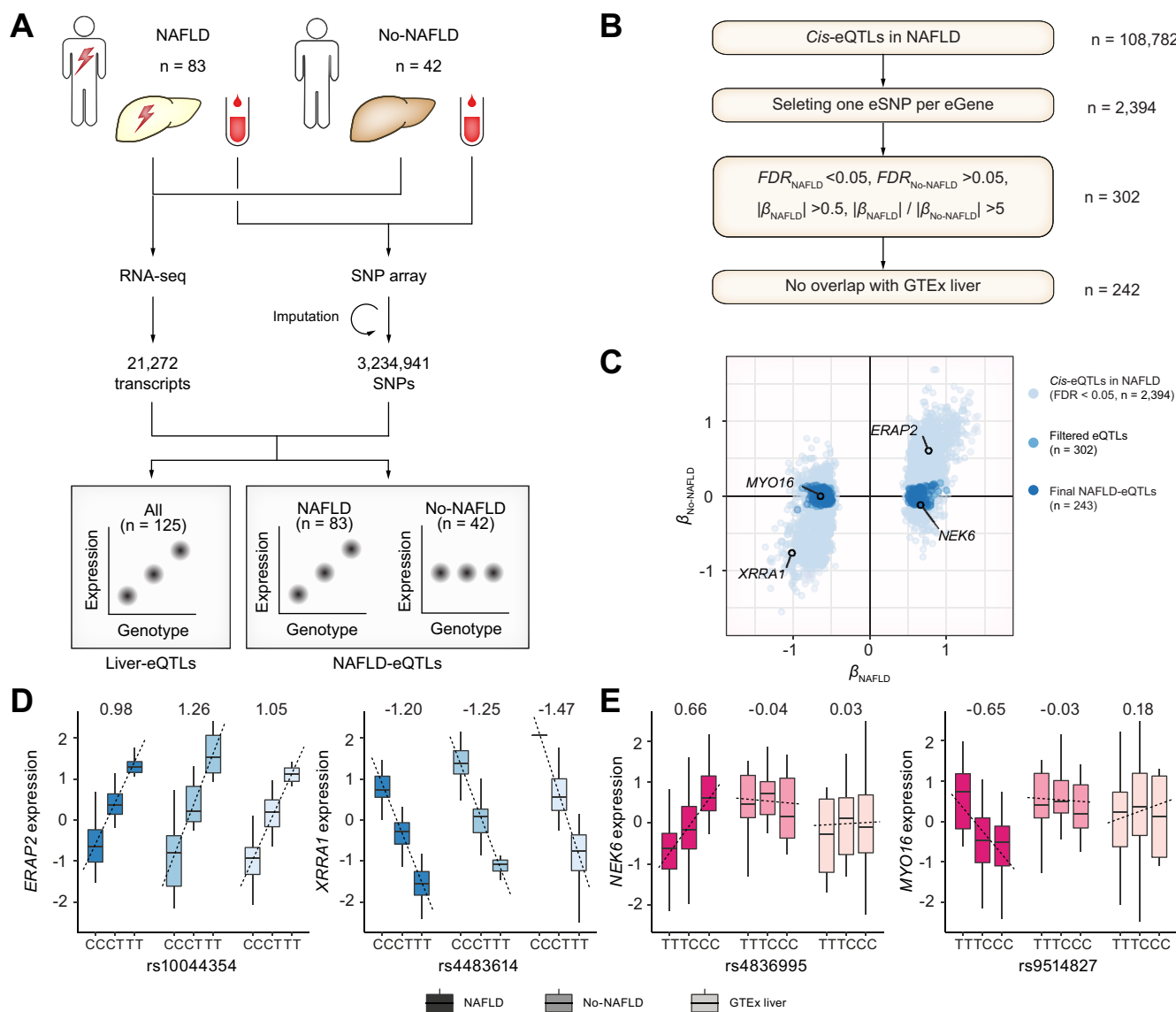
### AGXT2 expression is regulated by rs2291702 in NAFLD

Among the 242 NAFLD-eQTLs, we sought to identify loci that are biologically relevant and may contribute to NAFLD pathogenesis. We focused on the *AGXT2* locus as it is the second strongest signal after a long non-coding RNA (*RP11-469A15.2*) and it is exclusively expressed in the liver and kidney. One of the eSNPs, rs2291702, forms a significant *cis*-eQTL in the NAFLD group ( $p = 7.21 \times 10^{-9}$ ), but not in no-NAFLD ( $p = 0.38$ ) or GTEx liver ( $p = 0.31$ ) sets (Fig. 2A). This significance persisted after adjusting for BMI and HOMA-IR in addition to age and sex (Fig. S5). Furthermore, rs2291702 also showed a clear association ( $p = 3.32 \times 10^{-5}$ ) between its genotypes and *AGXT2* expression level in an additional Korean cohort ( $n = 162$ ; Fig. 2B and Table S2). Therefore, combining this independent cohort with our cohorts ( $n = 287$ ) yielded stronger evidence that the eSNP functions in the NAFLD status ( $p = 3.69 \times 10^{-12}$ ; Fig. S6A and S6B). As a result, the difference in *AGXT2* expression between NAFLD and no-NAFLD ( $P = 3.47 \times 10^{-6}$ ) was largely attributed to CC carriers ( $p = 4.79 \times 10^{-6}$ ) and not to non-CC carriers ( $p = 0.16$ ; Fig. 2C). Immunohistochemistry on human liver biopsy samples confirmed that the protein expression pattern of *AGXT2* is consistent with its gene expression pattern, which is dependent on the disease state and rs2291702 genotype (Fig. 2D and S7). The top 8 *AGXT2*-eSNPs lie in a linkage disequilibrium (LD) block that spans  $\sim 4.7$  kb in the 4–7th introns of the gene (Fig. 2E), and these SNPs show variable allele frequencies (AF) across different populations. In the 1000 Genomes database, the reference T allele of rs2291702 predominates over the alternative C allele in African populations (mean AF = 0.802), and has a frequency that is roughly half in European populations (mean AF = 0.511), whereas it is minor in East Asians and Koreans (AF = 0.327 and 0.330, respectively; Fig. 2F). This observation implies a potential population-specific role of these SNPs along with differential susceptibility to the *AGXT2*-dependent NAFLD pathway.

### Role of AGXT2 in NAFLD progression

*AGXT2* encodes a mitochondrial alanine-glyoxylate aminotransferase, which is responsible for systemic regulation of metabolites such as asymmetrical and symmetrical dimethylarginine



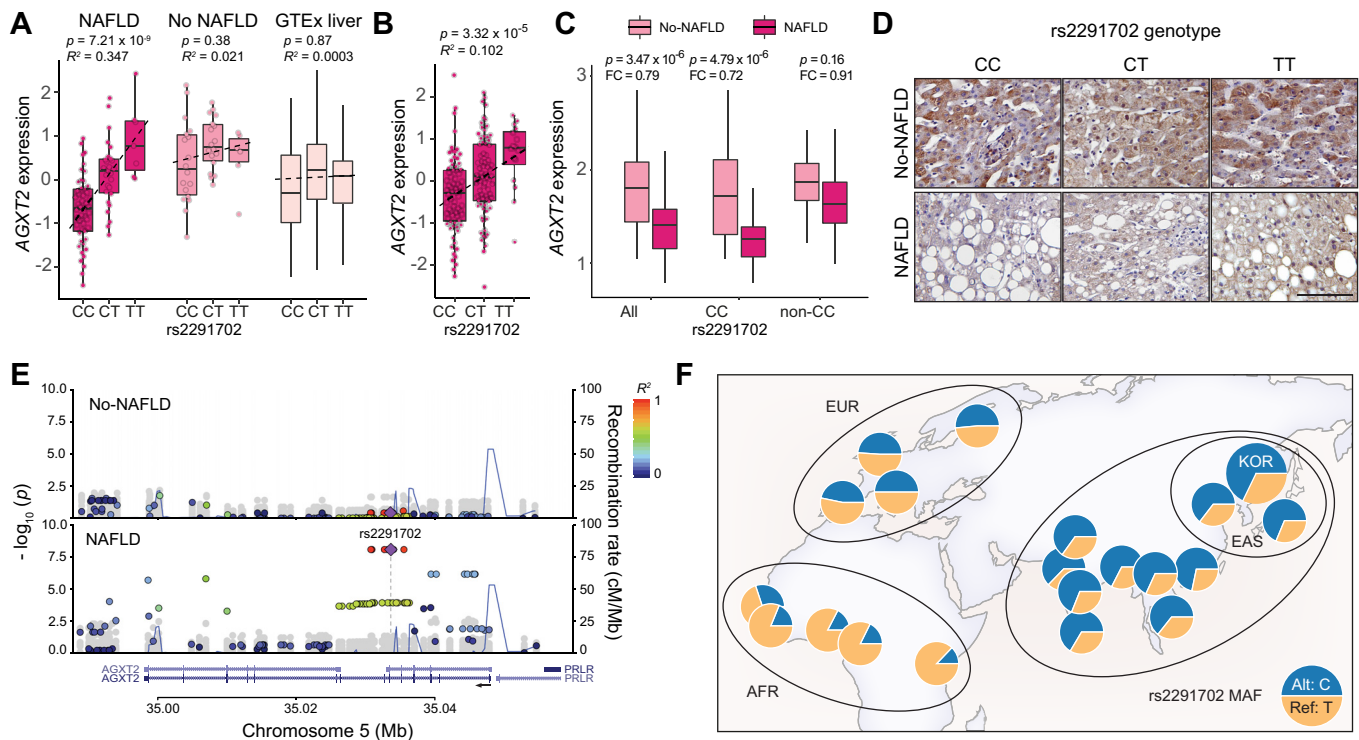


**Fig. 1. Screening NAFLD-eQTLs.** (A) Study scheme. (B) NAFLD-eQTL calling process. (C) Scatterplot of  $\beta$  values depicting the eQTL calls during the filtering process described in (B). Genes shown in D-E are indicated. (D-E) Examples of liver-QTLs that show significant associations in all physiological conditions (D), and of NAFLD-eQTLs that are significant only in the NAFLD group (E). The numbers above each genotype class denote  $\beta$  values. Darkness of bars depicts different groups (NAFLD, no-NAFLD, and GTEx liver). eQTL, expression quantitative trait loci; NAFLD, non-alcoholic fatty liver disease.

(ADMA, SDMA), and BAIB.<sup>19–21</sup> It is enriched in the liver and kidney (Fig. 3A), and within the liver, it is mainly expressed in hepatocytes, as evidenced by single-cell RNA-seq analysis of the human liver (Fig. 3B)<sup>29</sup> and a western blot in HepG2 cells (Fig. 3C). In comparison, AGXT2 is expressed at low levels in LX-2, a human hepatic stellate cell line (Fig. 3C). AGXT2 expression is significantly correlated with pathological and clinical features such as the degree of steatosis, ballooning, fibrosis, and lobular inflammation, and the levels of hyaluronic acid (HA), alanine aminotransferase (ALT), aspartate aminotransferase (AST), and HOMA-IR (Fig. 3D, 3E and S8). As expected, this effect was more evident in rs2291702:CC carriers, reflected by higher  $R^2$  values than those of non-CC carriers, whereas rs2291702:non-CC carriers displayed no significant correlations between the histological severity of NAFLD and AGXT2 expression (Fig. 3E, Table S5

and S6). Therefore, it is plausible to postulate that AGXT2-eSNPs function as causative variants in NAFLD, and the alteration of AGXT2 expression by the rs2291702 genotype modifies NAFLD pathogenesis.

Despite the growing evidence that AGXT2 and its substrates may play an important role in the pathogenesis of cardiovascular and metabolic syndromes,<sup>17,18</sup> the exact role of AGXT2 in NAFLD is currently unclear. Therefore, we investigated whether altering expression of the gene will allow us to elucidate its function in regulating NAFLD progression. As the gene was downregulated in the NAFLD group (Fig. 2), we first tested whether the reduction of AGXT2 may contribute to NAFLD development. Agxt2 knockdown shRNA was injected into 4 nine-week-old mice on a normal diet. Seven days after the injection, histological analysis revealed that Agxt2-knockdown mouse livers featured an



**Fig. 2. AGXT2 expression is regulated by rs2291702 in a genotype-dependent manner.** (A) AGXT2 expression in NAFLD, no-NAFLD and GTEx liver samples. (B) AGXT2 expression in the replication cohort. (C) Difference in expression of AGXT2 by rs2291702 genotype. (D) Immunohistochemistry of liver biopsy samples corresponding to each genotype-phenotype category. Scale bar: 100  $\mu$ m. (E) Regional plots of the eQTL association between each SNP's genotype and AGXT2 expression level. AGXT2 locus displaying association from no-NAFLD (upper panel) and NAFLD (lower panel) groups. Color represents  $R^2$  values between rs2291702 and nearby SNPs. SNPs in grey denote signals calculated using genes other than AGXT2. Blue lines denote recombination rate collected from 1000 Genomes Project (Phase 3). The arrow shows direction of AGXT2 transcription. (F) World-wide allele frequency distribution of rs2291702. Data obtained from 1000 Genomes. eQTL, expression quantitative trait loci; NAFLD, non-alcoholic fatty liver disease; SNP, single nucleotide polymorphism.

increase in collagen deposition (Fig. 4A and S9A). We also observed an increase in serum AST/ALT levels and hepatic transcript levels of fibrogenesis (*Col1a1* and  $\alpha$ SMA), inflammation (*Tnf $\alpha$* , *Il-1 $\beta$*  and *Cd36*), and adipogenesis (*Adrp* and *Acaca*), reflecting the NAFLD status (Fig. 4B and S10A-S10D).

Next, we overexpressed *Agxt2* in 4 mice on CDAHFD, a widely-used NASH-fibrosis animal model.<sup>30</sup> Whereas the *Agxt2*- or *GFP*-injected mice were comparable in terms of fat accumulation (Fig. 4C), Masson's trichrome staining revealed the reduction of collagen deposition with the increased *Agxt2* dosage (Fig. 4C and S9B). The mice expressing *Agxt2* also displayed a decrease in serum AST/ALT levels and hepatic transcript levels of fibrogenesis, inflammation, and lipogenesis (Fig. 4D and S10E-S10H), indicating a protective role of AGXT2 in NAFLD progression. Moreover, *Agxt2*-overexpressed mice demonstrated a significant reduction in serum ADMA level, one of the main substrates of AGXT2 (Fig. 4D). Taken together, the knockdown and overexpression results suggest that *Agxt2* ameliorates fibrogenesis in the NAFLD mouse model, but show less pronounced effects on lipid accumulation and the degree of steatosis in the liver. In conclusion, AGXT2 protects from NAFLD progression by counterbalancing the fibrogenesis process.

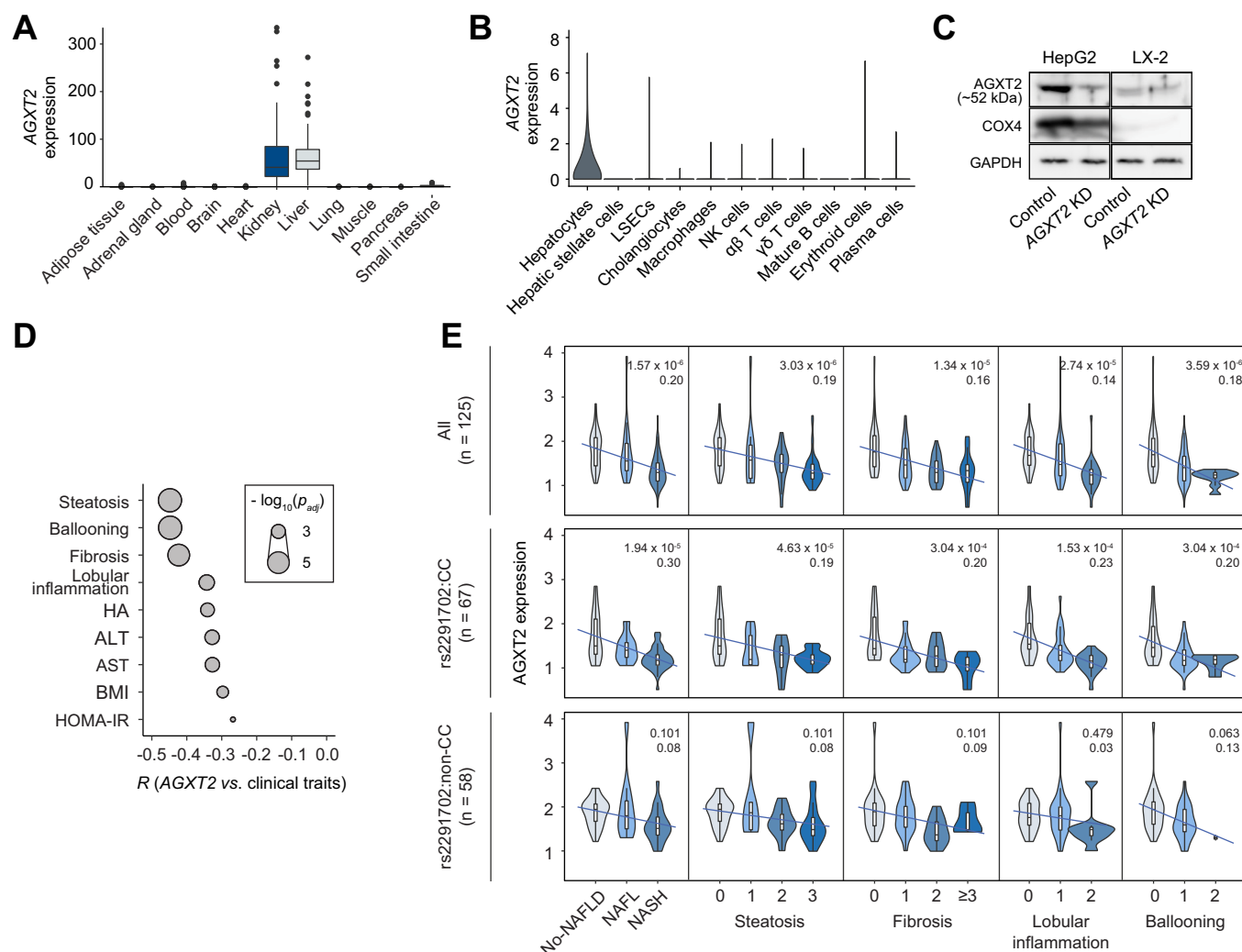
#### Alteration of the transcriptome in *Agxt2* knockdown mice

Next, we investigated whether the reduction of *Agxt2* can genetically mimic human NAFLD features. We obtained liver samples from *Agxt2* knockdown ( $n = 2$ ) and control ( $n = 2$ ) mice,

profiled transcriptional alteration by *Agxt2* depletion, and compared them with our human NAFLD transcriptomes (Fig. S11A). Of note, these mice were on a normal diet. We found a significant enrichment in the number of overlapped DEGs between human patients with NAFLD and *Agxt2* knockdown mice ( $p = 3.78 \times 10^{-3}$  for 163 concordantly regulated genes between human and mouse and  $p = 1.0$  for 41 discordant genes; Monte-Carlo simulation), implying that the *Agxt2* knockdown mouse model harbors similar physiological features to human NAFLD status. Next, we performed gene ontology analyses to elucidate the functional basis of these DEGs. Both DEG sets exhibited NAFLD-related terms in common, but enrichments in the concordant genes more prominently featured metabolic functions (e.g., amino acid and lipid metabolic processes) (Fig. S11B). This supports that mouse transcriptomic changes resulting from the reduction of *Agxt2* resemble those observed in patients with NAFLD and metabolic abnormalities.

#### AGXT2 knockdown induces ER stress and cell death

Based on the protective effects of *Agxt2* against lipotoxicity of hepatocytes (Figs. 4 and S10), we sought to elucidate *in vitro* effects of AGXT2 reduction (see supplementary method). Knockdown of AGXT2 in the HepG2 cells (as demonstrated by western blot; Fig. 3C) sensitized them to cell death after palmitate treatment ( $p < 0.005$ , ANOVA with the Tukey's test; Fig. 5A) and increased cellular levels of ER stress (GRP-78 and CHOP) and apoptosis (cleaved Caspase-3 and PARP) markers, as

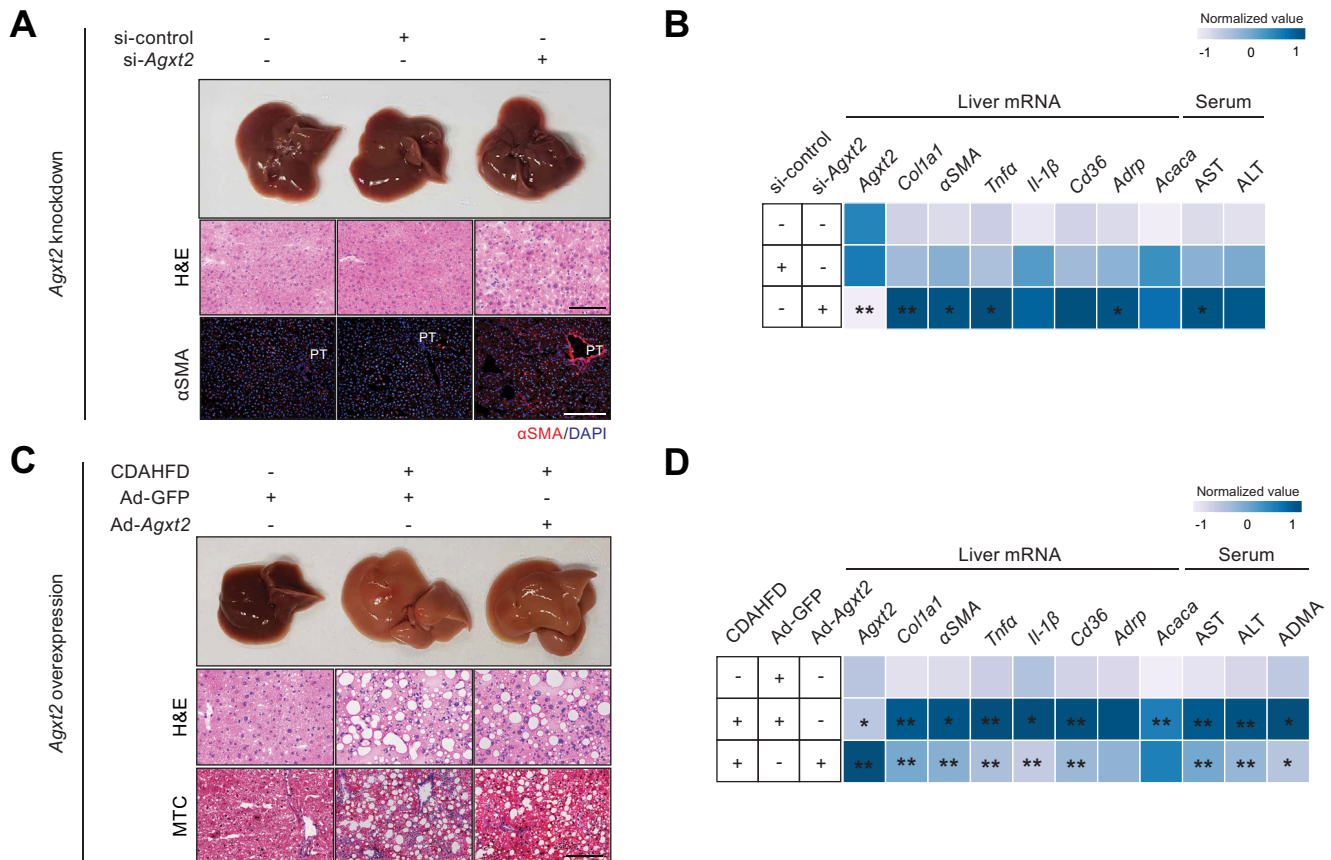


**Fig. 3. Hepatocyte-specific expression of AGXT2 is associated with liver steatosis and fibrosis.** (A) AGXT2 expression in GTEx tissues. (B) AGXT2 expression in human liver cell types analyzed at the single-cell level in the work of MacParland *et al.*<sup>29</sup> (C) AGXT2 expression in HepG2 and LX-2 cell lines. (D) Association of histological and metabolic parameters with AGXT2 expression. (E) Correlation between the histological severity of NAFLD and AGXT2 expression using all participants (top row), rs2291702:CC carriers (middle row) and rs2291702:non-CC carriers (bottom row). In each plot, the upper number indicates FDR-adjusted  $p$  value and the lower number shows  $R^2$  value in linear regression model. FDR, false discovery rate; LSEC, liver sinusoidal endothelial cell; NAFLD, non-alcoholic fatty liver disease.

detected by Western blot (Figs. 5B and S9C). Meanwhile, reduced expression of AGXT2 increased mitochondrial superoxide generation ( $p < 0.01$ , ANOVA with Tukey's test; Fig. 5C), which was enhanced by the addition of palmitate. We also observed reduced mitochondrial integrity and oxygen consumption rate by the reduction of AGXT2 (Fig. S12A-S12C). There were no differences in cell proliferation (Fig. S12D). Similar patterns were observed in murine primary hepatocytes, in which ER stress markers were increased by the addition of palmitate at 300  $\mu$ M (Figs. 5D and S9D). The expression was somewhat reduced at 500  $\mu$ M, reflecting our assumption that alternative pathways may be activated due to high PA level. As increased hepatic ER stress is known to cause cell death and fibrogenesis,<sup>31</sup> these results suggest that decreased AGXT2 exacerbates liver fibrosis by increasing ER stress-mediated hepatocyte death in NAFLD.

## Discussion

To accommodate the complex and polygenic nature of NAFLD, and to enhance the ability to identify genes underlying its pathomechanism, we present a disease-specific eQTL (*i.e.*, NAFLD-eQTL) mapping pipeline. Unlike conventional eQTL approaches, our pipeline enabled us to identify eQTLs and associated genes that are active in the NAFLD environment. We confirmed that altered transcriptional activity depends on rs2291702, which is located in AGXT2 and constitutes one of our top NAFLD-eQTLs. Using both mouse and cell models, we validated a NAFLD-preventive effect from AGXT2. Transcriptional changes of NAFLD-associated genes and pathways in *Agxt2*-knockdown mice on a normal diet indicate that suppression of *Agxt2* phenocopies human NAFLD with respect to gene expression. Lastly, at the cellular level, we present evidence that the reduction of AGXT2 causes ER stress and hepatic cell death.



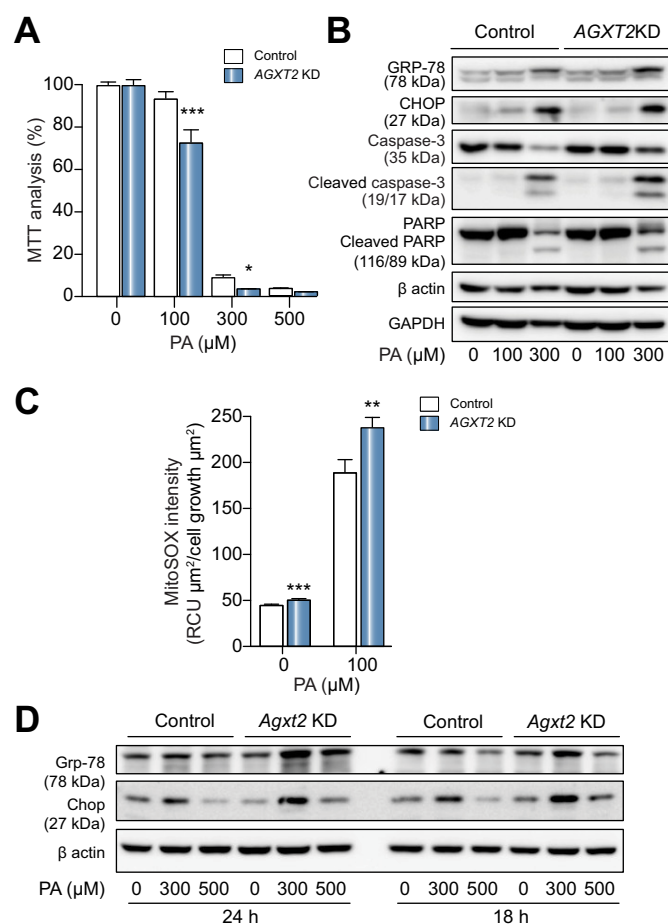
**Fig. 4. Agxt2 knockdown aggravates NAFLD phenotypes and its overexpression ameliorates them.** (A, C) Representative gross pictures, H&E, αSMA and MTC staining on sections of mouse livers. Scale bars: 200 μm. (B, D) Heatmaps representing normalized serum AST/ALT and ADMA levels and mRNA levels in Agxt2 knockdown and overexpression mice. \* $p < 0.05$ , \*\* $p < 0.01$ ; independent  $t$  test. ADMA, asymmetric dimethylarginine; ALT, alanine aminotransferase; AST, aspartate aminotransferase; NAFLD, non-alcoholic fatty liver disease; PT, portal triad.

A number of studies have attempted to map healthy liver-eQTLs, revealing eQTLs that are active in normal physiological status.<sup>22,32,33</sup> Notably, a recent meta-analysis from 1,183 individuals detected that approximately 75% of all genes are associated with *cis*-eQTLs, consistent with a GTEx study.<sup>27,34</sup> In contrast with previous eQTL approaches using liver tissues,<sup>32–36</sup> our approach enables: (i) identification of eQTLs with disease risk, (ii) ranking eQTLs by effect size in the diseased state relative to the normal, and (iii) pinpointing individuals that harbor specific genotypes that are more or less susceptible to disease. For example, our analysis demonstrated that rs2291702:CC carriers display AGXT2 downregulation and are more prone to the development and progression of NAFLD compared to others (*i.e.*, CT/TT carriers). This feature does not necessarily select SNPs that were also significant in GWAS, because we subdivided the cases and controls and this concomitantly reduced power. Nevertheless, this feature might be utilized in selecting proper target individual groups when developing and prescribing molecular targeted agents against NAFLD. A recent study demonstrated a similar pattern of the MBOAT7 variant.<sup>37</sup> They observed that rs641738, a SNP near MBOAT7 with significance from GWAS, confers the risk of liver fibrosis in a genotype-dependent manner.<sup>37</sup> This study and ours both highlight the functional implication of genetic variants that are active in a certain physiological condition. In addition, our approach illuminates the

power of using non-European cohorts to discover novel genetic players, as the susceptible allele predominates only in East Asian populations. It would be worthwhile to test whether the effect of genotype on AGXT2 expression is valid in other ethnic cohorts.

We found that reduced Agxt2 induces phenotypical and transcriptomic profiles similar to those of human NAFLD (Fig. 4), with livers of Agxt2 knockdown mice displaying an increased expression of fibrogenesis, inflammation, and adipogenesis markers (Fig. 4B). Conversely, increasing Agxt2 in high-fat diet-fed mice reversed the marker expression changes (Fig. 4D), suggesting a protective effect of the gene on NAFLD. At the cellular level, we observed that reducing AGXT2 induces an increase in reactive oxygen species level and ER stress-mediated cell death upon metabolic stress, which is a well-known cause of liver fibrosis. This observation raises a question – how does reduced AGXT2 contribute to ER stress and hepatic cell death? Investigations on the physiological role of AGXT2 have mainly focused on the cardiovascular system, as one of its main substrates, ADMA, is the most potent endogenous NOS inhibitor. Moreover, a knockout mouse model displayed cardiovascular phenotypes.<sup>18</sup> In light of such findings, we tested if altered ADMA due to AGXT2 dysregulation may affect nitric oxide levels in the liver. However, AGXT2 knockdown in HepG2 cells did not alter nitric oxide status, as assayed by 3NT western blotting (data not shown). Also, ADMA itself failed to worsen the ER stress induced





**Fig. 5. Downregulation of AGXT2 induces ER stress and hepatocyte cell death under palmitate treatment.** (A) Measurement of oxidative stress-mediated cell death by MTT analysis in HepG2 cells treated with PA for 48 hours ( $n = 4$ ). (B) Western blot of ER stress markers in HepG2 cells by AGXT2 knockdown and the addition of PA for 24 hours. (C) Measurement of mitochondrial ROS in HepG2 treated with PA ( $n = 3$ ) for 24 hours. (D) Western blot of ER stress markers in mouse primary hepatocytes by *Agxt2* knockdown (24 hours) and the addition of PA (additional 18 or 24 hours). Values represented as mean  $\pm$  SD, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.005$ ; ANOVA. ER, endoplasmic reticulum; PA, palmitic acid; ROS, reactive oxygen species.

by PA treatment, implying other or additional substrates may be engaged in this process. Another mechanism of AGXT2 pathogenesis is through altering amino acid metabolism, as AGXT2 utilizes various amino acid metabolites as substrates or products. Indeed, we observed that the levels of known substrates and amino acids that are closely linked with the AGXT2 enzymatic activity are altered by the reduction of AGXT2 (Fig. S13) in HepG2 cells. Although the causal relationship of amino acid dysregulation with NAFLD remains to be clarified, this observation is consistent with previous studies<sup>38,39</sup> and offers a plausible explanation for hepatic damage mediated by reduced AGXT2 levels. Lastly, the precise role of hepatic stellate cells in this process also remains as an open question, although we have observed similar changes of fibrosis and inflammation marker expression upon AGXT2 knockdown or overexpression in LX-2 cells (Fig. S14).

We observed that AGXT2-eSNPs are functional in NAFLD. Furthermore, we discovered that the SNPs possess transcriptional regulatory activity of varying degrees when placed on a

luciferase reporter, and the effect can influence across the LD block that they are located (Fig. S15). Nevertheless, precise molecular basis for AGXT2-eSNP's function needs to be further studied.

Despite the stringent filtering steps carried out in this study, it would be necessary to increase the sample size to achieve sufficient statistical power, and to systematically validate the function of each eQTL under both normal and affected status. Although our set of NAFLD-eQTLs possesses many previously associated NAFLD genes (Table S7), repeating our approach at the single-cell level would provide further insight into how cell type-specific NAFLD-eQTLs behave and confer the risk of NAFLD.

Herein, we presented a proof-of-concept method in which a disease-specific eQTL was selected and its function was validated toward the identification of therapeutic targets, specifically proposing AGXT2 as a novel druggable target against NAFLD. Given that NAFLD is a polygenic trait and a substantial portion of the NAFLD population is non-obese and probably considered to harbor substantial genetic susceptibility, it is expected that additional loci will be mapped with biological validation. Therefore, as we increase sample size to boost statistical power, it will become more feasible to find potent therapeutic targets and prospects for clinical interventions in an individual-specific manner. To date, therapeutic clinical trials have been unsuccessful largely because they failed to consider the genetic heterogeneity of patients with NAFLD. In this respect, our approach has substantial advantages for the NAFLD drug development process through selective enrollment of a high-risk NAFLD population carrying risk variants.

## Abbreviations

Adipo-IR, adipose insulin resistance index; ADMA, asymmetric dimethylarginine; AF, allele frequency; AGXT2, alanine-glyoxylate aminotransferase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BAIB, 3-amino-isobutyrate; CDAHFD, choline-deficient, L-amino acid-defined, high-fat diet; ChIP-seq, chromatin immunoprecipitation sequencing; DBP, diastolic blood pressure; DEG, differentially expressed gene; DMGV, dimethylguanidino valeric acid; eQTL, expression quantitative trait loci; ER, endoplasmic reticulum; FDR, false discovery rate; GGT, gamma-glutamyl transferase; GWAS, genome-wide association studies; HOMA-IR, homeostasis model assessment of insulin resistance; MBOAT7, membrane bound O-acyltransferase domain containing 7; NAFL, non-alcoholic fatty liver; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; NOS, nitric oxide synthase; MAF, minor allele frequency; PA, palmitic acid; PRS, polygenic risk score; SBP, systolic blood pressure; TSS, transcription start site.

## Financial support

This work was funded by the following grants from the National Research Foundation of Korea (2019R1A2C2010789, 2021M3A9E4021818, and 2021R1A2C2005820), Seoul National University College of Medicine (the Collaborative Research Program of SNU Boramae Medical Center and Basic Medical Science, 800-20210005), Korean Gastroenterology Fund for Future Development, the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare, Republic of Korea (HI21C0538), and Korea Basic Science Institute (C170200).

## Conflicts of interest

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

## Authors' contributions

W.K. and M.C. conceived and designed the study; S.K.J., J.L. and W.K. contributed to sample acquisition; T.Y., Y.L. and M.C. contributed to sample processing, quality control and data generation; T.Y., O.J., S.R., W.K. and M.C. contributed to the analysis and interpretation of the data; H.J.K., H.Y.K., H.S., H.-H.K., S.J., W.-I.J., G.-S.H., K.W.K. and J.W.K. contributed to the functional analyses and interpretation; T.Y., W.K. and M.C. drafted the manuscript; all authors read and approved the manuscript.

## Data availability statement

The data that support the findings of this study are available from the corresponding authors, upon reasonable request. The RNA-seq data from mouse livers are available from the Korean Nucleotide Archive (<https://kobc.re.kr/bps/kona>; Accession ID: PRJKA200013).

## Acknowledgments

We thank the participants of the study and Jana Kneissl for technical assistants. Silvia Sookoian, Tae Soo Kim, Yoon Jung Kim, Sangmoon Lee and Yongjin Yoo provided critical comments.

## Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jhep.2021.04.011>.

## References

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

- [1] Eslam M, George J. Genetic contributions to NAFLD: leveraging shared genetics to uncover systems biology. *Nat Rev Gastroenterol Hepatol* 2020;17:40–52. <https://doi.org/10.1038/s41575-019-0212-0>.
- [2] Estes C, Anstee QM, Arias-Loste MT, Bantel H, Bellentani S, Caballeria J, et al. Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016–2030. *J Hepatol* 2018;69:896–904. <https://doi.org/10.1016/j.jhep.2018.05.036>.
- [3] Segovia-Miranda F, Morales-Navarrete H, Kücken M, Moser V, Seifert S, Repnik U, et al. Three-dimensional spatially resolved geometrical and functional models of human liver tissue reveal new aspects of NAFLD progression. *Nat Med* 2019;25:1885–1893. <https://doi.org/10.1038/s41591-019-0660-7>.
- [4] Eslam M, Valenti L, Romeo S. Genetics and epigenetics of NAFLD and NASH: clinical impact. *J Hepatol* 2018;68:268–279. <https://doi.org/10.1016/j.jhep.2017.09.003>.
- [5] Farrell GC, Larter CZ. Nonalcoholic fatty liver disease: from steatosis to cirrhosis. *Hepatology* 2006;43:99–112. <https://doi.org/10.1002/hep.20973>.
- [6] Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. *Nat Med* 2018;24:908–922. <https://doi.org/10.1038/s41591-018-0104-9>.
- [7] Trépo E, Valenti L. Update on NAFLD genetics: from new variants to the clinic. *J Hepatol* 2020. <https://doi.org/10.1016/j.jhep.2020.02.020>.
- [8] 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. <https://doi.org/10.1016/j.jhep.2020.04.003>.
- [9] Sookoian S, Pirola CJ. Genetic predisposition in nonalcoholic fatty liver disease. *Clin Mol Hepatol* 2017;23:1–12. <https://doi.org/10.3350/cmh.2016.0109>.
- [10] 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. <https://doi.org/10.1111/joim.12719>.
- [11] 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. <https://doi.org/10.1371/journal.pone.0185490>.
- [12] Rockman MV, Kruglyak L. Genetics of global gene expression. *Nat Rev Genet* 2006;7:862–872. <https://doi.org/10.1038/nrg1964>.
- [13] Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. *Nat Rev Genet* 2015;16:197–212. <https://doi.org/10.1038/nrg3891>.
- [14] Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. *Philos Trans R Soc B Biol Sci* 2013;368. <https://doi.org/10.1098/rstb.2012.0362>.
- [15] Müller-Deile J, Jobst-Schwan T, Schiffer M. Moving beyond GWAS and eQTL analysis to validated hits in chronic kidney disease. *Cell Metab* 2019;29:9–10. <https://doi.org/10.1016/j.cmet.2018.12.009>.
- [16] Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, et al. Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature* 2010;464:773–777. <https://doi.org/10.1038/nature08903>.
- [17] Rodionov RN, Jarzebska N, Weiss N, Lentz SR. AGXT2: a promiscuous aminotransferase. *Trends Pharmacol Sci* 2014;35:575–582. <https://doi.org/10.1016/j.tips.2014.09.005>.
- [18] Caplin B, Wang Z, Slaviero A, Tomlinson J, Dowsett L, Delahaye M, et al. Alanine-glyoxylate aminotransferase-2 metabolizes endogenous methyl-arginines, regulates NO, and controls blood pressure. *Arterioscler Thromb Vasc Biol* 2012;32:2892–2900. <https://doi.org/10.1161/ATVBAHA.112.254078>.
- [19] Nicholson G, Rantalainen M, Li JV, Maher AD, Malmmodin D, Ahmadi KR, et al. A genome-wide metabolic QTL analysis in europeans implicates two Loci shaped by recent positive selection. *PLoS Genet* 2011;7. <https://doi.org/10.1371/journal.pgen.1002270>.
- [20] O'Sullivan JF, Morningstar JE, Yang Q, Zheng B, Gao Y, Jeanfavre S, et al. Dimethylguanidino valeric acid is a marker of liver fat and predicts diabetes. *J Clin Invest* 2017;127:4394–4402. <https://doi.org/10.1172/JCI95995>.
- [21] Rueedi R, Ledda M, Nicholls AW, Salek RM, Marques-Vidal P, Morya E, et al. Genome-wide association study of metabolic traits reveals novel gene-metabolite-disease links. *PLoS Genet* 2014;10. <https://doi.org/10.1371/journal.pgen.1004132>.
- [22] 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. <https://doi.org/10.1371/journal.pbio.0060107>.
- [23] 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. <https://doi.org/10.1111/jgh.14056>.
- [24] Brunt EM, Kleiner DE, Wilson LA, Belt P, Neuschwander-Tetri BA. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. *Hepatology* 2011;53:810–820. <https://doi.org/10.1002/hep.24127>.
- [25] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:1–21. <https://doi.org/10.1186/s13059-014-0550-8>.
- [26] Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 2012;28:1353–1358. <https://doi.org/10.1093/bioinformatics/bts163>.
- [27] Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, et al. Genetic effects on gene expression across human tissues. *Nature* 2017;550:204–213. <https://doi.org/10.1038/nature24277>.
- [28] Schmidt EM, Zhang J, Zhou W, Chen J, Mohlke KL, Chen YE, et al. GREGOR: evaluating global enrichment of trait-associated variants in epigenomic features using a systematic, data-driven approach. *Bioinformatics* 2015;31:2601–2606. <https://doi.org/10.1093/bioinformatics/btv201>.
- [29] MacParland SA, Liu JC, Ma XZ, Innes BT, Bartczak AM, Gage BK, et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat Commun* 2018;9:1–21. <https://doi.org/10.1038/s41467-018-06318-7>.
- [30] Matsumoto M, Hada N, Sakamaki Y, Uno A, Shiga T, Tanaka C, et al. An improved mouse model that rapidly develops fibrosis in non-alcoholic steatohepatitis. *Int J Exp Pathol* 2013;94:93–103. <https://doi.org/10.1111/iep.12008>.

- [31] Lebeaupin C, Vallée D, Hazari Y, Hetz C, Chevet E, Bailly-Maitre B. Endoplasmic reticulum stress signalling and the pathogenesis of non-alcoholic fatty liver disease. *J Hepatol* 2018;69:927–947. <https://doi.org/10.1016/j.jhep.2018.06.008>.
- [32] Innocenti F, Cooper GM, Stanaway IB, Gamazon ER, Smith JD, Mirkov S, et al. Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. *PLoS Genet* 2011;7. <https://doi.org/10.1371/journal.pgen.1002078>.
- [33] Strunz T, Grassmann F, Gayán J, Nahkuri S, Souza-Costa D, Maugeais C, et al. A mega-analysis of expression quantitative trait loci (eQTL) provides insight into the regulatory architecture of gene expression variation in liver. *Sci Rep* 2018;8:1–11. <https://doi.org/10.1038/s41598-018-24219-z>.
- [34] Etheridge AS, Gallins PJ, Jima D, Broadaway KA, Ratain MJ, Schuetz E, et al. A new liver expression quantitative trait locus map from 1,183 individuals provides evidence for novel expression quantitative trait loci of drug response, metabolic, and sex-biased phenotypes. *Clin Pharmacol Ther* 2019;0. <https://doi.org/10.1002/cpt.1751>.
- [35] Liu W, Anstee QM, Wang X, Gawrieh S, Gamazon ER, Athinarayanan S, et al. Transcriptional regulation of PNPLA3 and its impact on susceptibility to nonalcoholic fatty liver Disease (NAFLD) in humans. *Aging (Albany NY)* 2017;9:26–40. <https://doi.org/10.18632/aging.101067>.
- [36] Pashos EE, Park YS, Wang X, Raghavan A, Yang W, Abbey D, et al. Large, diverse population cohorts of hiPSCs and derived hepatocyte-like cells reveal functional genetic variation at blood lipid-associated loci. *Cell Stem Cell* 2017;20:558–570. <https://doi.org/10.1016/j.stem.2017.03.017>. e10.
- [37] Thangapandi VR, Knittelfelder O, Brosch M, Patsenker E, Vvedenskaya O, Buch S, et al. Loss of hepatic Mboat7 leads to liver fibrosis. *Gut* 2020. <https://doi.org/10.1136/gutjnl-2020-320853>. epub.
- [38] van den Berg EH, Flores-Guerrero JL, Gruppen EG, de Borst MH, Wolak-Dinsmore J, Connelly MA, et al. Non-alcoholic fatty liver disease and risk of incident type 2 diabetes: role of circulating branched-chain amino acids. *Nutrients* 2019;11. <https://doi.org/10.3390/nu11030705>.
- [39] Hoyles L, Fernández-Real JM, Federici M, Serino M, Abbott J, Charpentier J, et al. Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. *Nat Med* 2018;24:1070–1080. <https://doi.org/10.1038/s41591-018-0061-3>.