# SUPPLEMENTARY METHODS

## Allelic imbalance analysis in human biopsies

For homogenization of 5–10 mg frozen tissue and subsequent nucleic acid isolation, tubes with 1.4-mm ceramic beads (Precellys, Villeurbanne, France) and the AllPrep DNA/RNA Mini Kit (Qiagen, Hilden, Germany) were used. Reverse transcription was performed using the Advantage RT-for-PCR Kit (Takara Clontech) with a random primer. Region of interest in the MBOAT7 transcript was amplified using the primers (TCCTTGTGTCTTTCGCTCC; TACACACGGTGACCTGTCA) with the following tag on each primer (ATG, CTG, GTG, ATC, CTC, GTC, AGT, CGT, TGT, CAG, TAG, GAG, (Metabion, Planegg/Steinkirchen). Cycling was done using the Bioline PCR mix (BIOLINE) in 25 µl reaction vol. at 94°C for 1min, 29 cycles at 60°C for 30 s, 72°C for 1 min, 94°C for 30 s, and a final elongation step at 72°C for 5 min. For library preparation and sequencing, Illumina technology was used. All sample-specifically tagged amplicons from a single sample were pooled together. For each pool an indexed sequencing library was prepared using Illumina TruSeq DNA PCR-free library preparation kit following the manufacturer's protocol with the exception that amplicons (around 500 ng) were directly introduced into end-repair skipping the fragmentation step. The libraries were quantified and quality checked using the Agilent Bioanalyzer 2100 and the Agilent DNA 7500 Kit (Agilent Technologies). Libraries were sequenced using the Illumina MiSeq (2 x 300 cycles, paired-end mode). Read information was extracted by MiSeq Control Software (MCS) v2.4.1.3 (Illumina). This includes demultiplexing of reads based on the Illumina indices and trimming of Illumina adapter sequences. Quality control of Sequence data was performed using FastQC. Demultiplexing of reads based on sample-specific PCR tags was done with an inhouse script. Reads were filtered for the expected lengths and the correct primer pair sequences. Only reads derived from molecules that were successfully sequenced in both forward and reverse direction were selected. Chimeras were excluded by filtering the read pairs for identical tags. Reads from each sample were allele-sorted and alleles counted.

#### Human Lipidomic Cohort

280 human liver samples were obtained from patients in Germany, in whom an intraoperative liver biopsy was indicated on clinical grounds such as during scheduled liver resection, exclusion of liver malignancy during major oncologic surgery or assessment of liver histology during bariatric surgery. Samples were snap frozen immediately in liquid nitrogen ensuring an ex vivo time of less than 40 seconds. Patients with evidence of viral hepatitis, hemochromatosis or alcohol consumption >20 g/day (women) and >30 g/day (men) were excluded. For all samples, MBOAT7 rs641738C>T genotype as well as full phenotypic and histological information [1] generated by a single pathologist (C.R.) blinded to the lipidomic analysis was available (Table-1). Phenotypic groups (Table-1) were defined on the basis of clinical and histological parameters as follows [2]: Normal controls (BMI<30) and healthy obese (BMI≥30) patients showed histological fat content ≤5%, no histological inflammation, no ballooning and no fibrosis. NAFL was defined as histological fat contents above 5%, absence of lobular inflammation and ballooning, and with presence or absence of F1 fibrosis. Early NASH was defined as histological fat content above 5%, with presence of lobular inflammation, an NAS score below 5 and with presence or absence of F1 fibrosis. NASH was defined by an NAS score greater or equal to 5. The study protocol abides to the ethical guidelines of the 1975 Declaration of Helsinki and was approved

by the institutional review board (Universität Kiel, D425/07, A111/99), before study commencement.

# Human NAFLD cohort

A total of 846 adult Caucasian NAFLD patients from tertiary referral centers in Austria (n=258), Germany (n=537) and Switzerland (n=51) who underwent percutaneous or surgical liver biopsy were included into this study (Supplementary-Table-1-2). NASH was defined by the NAFLD activity score (NAS). Presence of fibrosis was assessed histologically according to Kleiner classification [1] (stage-F0:no fibrosis; stage-F1:perisinusoidal fibrosis to portal/periportal fibrosis; stage-F2:perisinusoidal and portal/periportal fibrosis; stage-F3:bridging fibrosis). In all patients, infectious (e.g. viral hepatitis, HIV), immunological, drug-induced hepatic-steatosis (e.g. amiodarone, methotrexate, steroids, valproate, etc.) or hereditary causes (hereditary hemochromatosis, Wilson disease) of liver disease were excluded. Alcohol consumption was assessed by self-reporting; subjects with average alcohol consumption of more than 30 g/day (men) or 20 g/day (women) were excluded from this study. The liver biopsies were read by experienced pathologists in a blinded fashion. Portions of this NAFLD cohort have been described previously [3–5]. All patients gave their written consent. Subjects recruited for the genetic association study came from different sites, at which the study protocol was approved by the ethics committees of the participating institutions.

#### DNA preparation and genotyping of the human NAFLD cohort

Of all cases and controls, genomic DNA was extracted from peripheral blood samples according to standard procedures. Genomic DNA (1 μl) was amplified with the GenomiPhi (Amersham) whole-genome amplification kit. Genotyping of MBOAT7 rs641738 (hcv8716820), was performed using the Taqman chemistry (Applied Biosystems, Foster City, CA, USA) on an automated platform with TECAN Freedom EVO and 384well TEMO liquid handling robots (TECAN, Männedorf, Switzerland) as described earlier [6,7]. Reactions were completed and read in a 7900 HT TaqMan sequence detector system (Applied Biosystems, Foster City, CA, USA). All process data were logged and administered with a database-driven LIMS [8].

### Quantitative real time-PCR (qRT-PCR)

RNA was isolated using Nucleospin RNA II kit (Macherey-Nagel) and for cDNA synthesis IScript cDNA Synthesis Kit (Bio-Rad) was used. QRT-PCR was performed using SsoFast EvaGreen Supermix (Bio Rad), gene-specific primers (Supplementary table 10) and the Bio-Rad cycler system (Bio-Rad CFX 384 Real-time system). Calculation was based on the threshold cycle method (ΔΔCT) [9]. Expression levels were normalized to those of B2m or 18s.

## Non parenchymal cell (NPC) isolation and flow cytometry

The liver-NPCs were isolated after perfusing the liver with ice-cold PBS. The tissue was minced and subjected to enzymatic digestion (crude collagenase at 2 mg/ml from Clostridium histolyticum and hyaluronidase at 0.5 mg/ml from bovine testis; Sigma Aldrich) in DMEM at 37°C for 30 minutes. Then, the cell suspension was passed through a strainer (100  $\mu$ m) followed by centrifugation at 50 x g, to remove the

Thangapandi VR*, et al. Gut* 2021; 70:940–950. doi: 10.1136/gutjnl-2020-320853

hepatocytes in the pellet. RBC lysis buffer (eBioscience) was used to perform RBC lysis, subsequently the cells were passed through a 40 µm strainer. Subsequently, the cells were washed in buffer containing 0.1% BSA and counted.

NPCs were subjected to flow cytometry analysis after staining with the respective antibodies. Within the CD45<sup>+</sup> population, macrophages were identified as Ly6G- /CD11b<sup>+</sup> /F4/80<sup>+</sup> monocytes as Ly6G-/CD11b+/F4/80and neutrophils as CD11b<sup>+</sup>/Ly6G<sup>+</sup>. Within the CD3<sup>+</sup> population, CD4<sup>+</sup> and CD8<sup>+</sup> cells were identified. The cells were stained with anti-mouse antibodies, against CD3 (clone 145-2C11), CD4 (clone GK1.5), CD8 (clone 53-6.7), CD45 (clone 30-F11), CD11b (clone M1/70), F4/80 (clone BM8) and Ly6G (clone 1A8). Corresponding isotype control stainings were performed using the following antibodies: American hamster IgG for PeCy7, Rat IgG2b κ for APC, Rat IgG2b κ for PerCP, Rat IgG2a κ for PeCy7, Rat IgG2b κ for PerCP, Rat IgG2b κ for APC and Rat IgG2b κ control for Alexa488. Antibodies were purchased from BD Biosciences, eBioscience and Biolegend. Measurement was done using the FACS Canto II flow cytometer and the data were analyzed using FlowJo (Tree Star) software.

#### Histological analysis of mouse liver

PFA-fixed and paraffin-embedded liver sections were stained with hematoxylin/eosin (H&E). Briefly, sections were deparaffinized and rehydrated by subsequent submersions in the following solutions: xylol (twice) for 10 min each, 96% ethanol (twice), 70% ethanol, 40% ethanol and aqua dest for 1 min each. Sections were stained for 5 min in Hematoxylin solution (Thermo Fisher Scientific) and submerged in running tap water for 5 min. Thereafter, the slides were stained with Eosin-Y solution for 10 seconds and washed in tap water for 2 min. Finally, the slides were dehydrated by 1

min incubation in 40% ethanol, 70% ethanol, 96% ethanol (three times) and 100% ethanol (twice). Finally, slides were submerged in xylol (twice) for 2 min each and then mounted with Richard-Allan Scientific Cytoseal XYL (Thermo Fisher Scientific) mounting medium.

For Picro Sirius red staining, the slides were deparaffinized and stained in hematoxylin for 8 min. The slides were then washed in tap water and stained in Picro Sirius red solution for 1 h. Next, the slides were shortly washed in 30% acetic acid and treated with 96% ethanol (twice) for 4 min each, followed by isopropanol for 4 min. Finally, slides were submerged in xylol solution and then mounted with mounting medium.

For Oil red O staining of the liver, cryo-sections were used. First, the sections were fixed in 10% formaldehyde for 30 min and then washed in PBS. The slides were then incubated in Oil red O solution for 10 min. Then, the slides were washed in 60% isopropanol and followed by PBS. Then, the slides were stained in Mayer's Hematoxylin for 30 sec and washing in tap water. The slides were mounted with VectaMount (TM) Mounting Medium. Images were taken using the Axio Observer.Z1 microscope (Carl Zeiss). An analyser slider with lambda-plate (Zeiss) for circular polarized light was used for imaging of Picro Sirius red-stained slides. ImageJ software was used for quantification of Picro Sirius red staining.

#### Genomic DNA isolation and detection of Mboat7 exon 5 deletion

To detect the Mboat7 exon 5 deletion, genomic DNA was isolated from primary mouse hepatocytes using QIAamp DNA micro kit (Qiagen). PCR was performed using DreamTaq Green PCR master mix (Thermo Fisher Scientific) with Mboat7 exon 5 flanking primers 5'CAS-F1 (AAGGCGCATAACGATACCAC) and 3'LOXP-R1 (ACTGATGGCGAGCTCAGACC). After amplification, the PCR product was run on a

2% agarose (Serva) gel together with 50 bp and 100 bp ladder (Generuler DNA ladders, Thermo Fisher Scientific). The gel was imaged on Peqlab "Quantum-ST4- 1100".

#### Hepatic hydroxyproline quantification

Hydroxyproline (HYP) concentration was determined biochemically using a Hydroxyproline Assay Kit (Sigma-Aldrich, Darmstadt, Germany), according to the manufacturer's protocol. Briefly, liver tissue (10-20 mg) was hydrolyzed in 600 µl of 6N HCl at 110 $\degree$ C for 16 h and then cleared from insoluble by passing through a 0.2  $\mu$ m nylon filter. Standard concentrations of all-trans HYP were processed in the same way and used to establish a calibration curve. 50 µl aliquots of hydrolysates or standards were dried under a nitrogen stream and the residues were solubilized in chloramine T for 25 min at RT. This was followed by incubation with freshly prepared Erlich's reagent for 20 min at  $65^{\circ}$  C. Finally, HYP concentration was determined by the reaction of oxidized hydroxyproline with 4-(dimethylamino) benzaldehyde, which results in a colorimetric (560 nm) product proportional to the hydroxyproline present. OD values were plotted on a calibration curve to calculate the hepatic HYP concentrations.

Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) activity assays ALT and AST activity assays (Sigma-Aldrich) were performed in serum samples according to kit protocol.

#### Multi-spot immunoassay

Multi-spot assay was performed using Mouse proinflammatory panel 1 kit (Meso Scale Discovery) to measure IFN $\gamma$ , IL1 $\beta$ , IL2, IL4, IL5, IL6, KC/GRO, IL10, IL12p70 and

TNF $\alpha$  protein levels. Tissue homogenates were used to perform the assay according to kit protocol.

# ANGPTL3 ELISA

ANGPTL3 ELISA (ThermoFisher Scientific) was performed in serum samples according to kit protocol.

# Protein quantification

Whole liver lysates were prepared by disrupting the tissue in RIPA buffer supplemented with cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail and PhosSTOP phosphatase inhibitor cocktail (Sigma Aldrich). Protein concentration was measured using BCA protein assay kit (ThermoFisher Scientific). TIMP1 (Abcam), eIF2 $\alpha$  (Cell Signaling), peIF2 $\alpha$  (Ser51, Cell signaling), CHOP (Santa Cruz), ATF4 (Cell Signaling) and vinculin (Cell Signaling) antibodies were used for immunoblotting. Imaging was performed using the UV Imager FusionFX7 with the Fusion software. Vinculin was used for normalization. Analysis was performed using ImageJ software

# Primary hepatocyte isolation

Primary hepatocyte isolation was performed according to a previously published protocol [10]. Briefly, 8-week old mice were anesthetized and the liver was perfused via the inferior vena cava and portal vein using a pump which allows perfusion at a flow rate of 5 ml/min, while the buffers are maintained at 37° C using a water bath. First, the liver was perfused with HBSS solution (containing 0.5 mM EDTA) for 7 minutes, followed by DMEM (ThermoFisher Scientific) containing 1 mg/ml Collagenase type 1 (ThermoFisher Scientific) for 10 minutes. The digested liver was excised and shaken gently in DMEM during which the hepatocytes dissociate. The isolated

hepatocytes were passed through a 100 μm cell strainer and washed twice at 4° C. Cells were resuspended in DMEM/F-12 medium (ThermoFisher Scientific) supplemented with 10% FBS, 1% Penicillin/Streptomycin and L-glutamine. Cells were seeded in 10% collagen (from calfskin, Sigma Aldrich) coated cell culture plates. For Oil Red O staining of hepatocytes, cells were fixed in 10% formaldehyde. Next cells were then washed in water, 60% isopropanol and allowed to dry. Cells were incubated with freshly prepared Oil Red O solution for 1 h at room temperature. Finally, the cells were washed in tap water three times and imaged.

#### Mouse and Human lipidomic analysis

### Overview of lipidomic analysis

Approximately 25 mg of liver tissue was homogenized in 300 µL of neat isopropanol and the protein concentration was determined by BCA assay (Thermo Fisher Scientific, Rockford IL). Lipids were extracted from aliquots containing an equivalent of 50 µg of total protein and quantified by shotgun lipidomics as described [11,12]; For more detail see below. For oxylipin and free fatty acid measurement, approximately 30 mg of liver tissue was used for sample preparation using LC/ESI-MS/MS at Lipidomix GmbH (Berlin, Germany) as described in a separate section below. The oxylipins and fatty acids were normalized to total protein concentration.

#### Lipid extraction from mouse and human liver tissues

Approximately 25 mg of liver tissue was homogenized in 300 µL of neat isopropanol and the protein concentration was determined by BCA assay (Thermo Fisher Scientific, Rockford IL). Aliquots containing an equivalent of 50 µg of total protein were used for lipid extraction [11,12]. After evaporation of the organic phase, lipid extracts were

reconstituted in 400 µL of 2:1 MeOH / CHCl3 for mouse samples and in 600 µl 2:1 MeOH/CHCl3 for human samples and stored at -20° C. 10 ul of liver lipid extract was diluted with 90  $\mu$ L spray solution (4:2:1 IPA/MeOH/CHCl<sub>3</sub> + 7.5 mM ammonium formate) for mass spectrometric analysis.

#### Human lipidomic data processing

The lipidomics data generated from the human liver samples were quality-checked according to the following steps. First, selected samples were chosen for the pilot run. The goal of the selection was to find the extreme of results (i.e. Extremely fatty and non-fatty samples).

Sample where ranked from fatty to non-fatty samples, in order for the lipid profile of these samples to be used to distinguish between fatty and non-fatty conditions. To consider all the features and not just the fatty feature, a regression analysis was performed, specifically gradient boosted tree as these are well suited for ranking models, the initial data was numerically formatted by covering categorical data into values, e.g. male, female to 1,0 and by removing incomplete entries, additionally a feature was added based on the concurrence of multiple factors, e.g. fatty, inflamed and fibrosis. Later for the full cohort run, 36 samples from the pilot were pooled together in order to create a single quality control (QC) sample. A QS sample was added to each batch to control for inter-batch variation and for the instrument stability. Variation of the measurements across all the batches is considered acceptable (e.g. QC standard deviation was 15% for neutral lipids, and 24% for phospholipids). Each sample was measured twice, and technical replicates were retrieved by splitting a sample lipid extract in running solution into two separate wells. The noise in spectra was reduced based on a repetition rate filtering performed with the PeakStrainer. Lipid

identification was performed with the LipidXplorer v. 1.2.4. The quantification of lipid species, based on known concentrations of internal standards, was performed using in-house scripts. Lipid species for which the internal standard signal was detected but the signal for the species itself was not detected were set to zero. All internal standards were detected in all the samples. The obtained lipidomics dataset underwent two further filters. First, for any given lipid species, when the variation between technical replicates was below 40% the respective mean was taken, otherwise the value was set to NaN. To avoid loss of measurements close to the detection limit, one exception to this first filter has been applied: If a lipid species detected in only one of the technical replicates has a measured value smaller or equal to twice the minimum detected for the respective lipid species (i.e. value  $\leq 2<sup>*</sup>min(species)$ ), this non-zero value has been rescued. Second, a lipid species was set to zero in all samples of one patient group (normal control, healthy obese, NAFL, NASH) if present in less than 15% of the samples of this group.

#### Annotation of lipid classes and species

The measured glycerolipids are triacylglycerols (TG) and diacylglycerols (DG); the glycerophospholipids and lyso-glycerophospholipids are phosphatidic acids (PA), phosphatidylinositols (PI), phosphatidylserines (PS), phosphatidylglycerols (PG), phosphatidylethanolamines (PE), phosphatidylcholines (PC), ether phosphatidylethanolamines (PE O-), ether phosphatidylcholines (PC O-), lysophosphatidic acids (LPA), lyso-phosphatidylinositols (LPI), lyso-phosphatidylcholines (LPC) and lyso-phosphatidylethanolamines (LPE); the sphingolipids are ceramides (Cer) and sphingomyelins (SM); the sterols are cholesterol (Chol) and cholesterol ester (CE). Glycero- and glycerophospholipid species were annotated as previously

described [13] using the total number of carbon atoms: total number of double bonds in both fatty acid / fatty alcohol moieties. Sphingolipid species were annotated by the total number of carbon atoms: double bonds: hydroxyl groups at the ceramide backbone.

#### Common chemicals and lipid standards

All solvents were of LC-MS or better grade. Synthetic lipid standards were purchased from Avanti Polar Lipids (Alabaster AL). Stocks of internal standards were stored in sealed glass ampoules at -20° C until they were used for preparing the internal standard mix in 10:3 methyl-tert-butyl ether (MTBE) / methanol (MeOH). 700 µL of the internal standard mix for mouse samples contained: 1778 pmol of D7-cholesterol; 2080 pmol of D7- CE 16:0; 1000 pmol of D5-TG 50:0; 465 pmol of D5-DG 34:0; 1321 pmol of PC 25:0; 386 pmol of LPC 13:0; 590 pmol of PE 25:0; 85 pmol of LPE 13:0; 240 pmol of PI 25:0, 164 pmol of PG 25:0; 73 pmol of Cer 30:1:2; 185 pmol of PA 25:0; 91 pmol of LPA 13:0; 271 pmol of SM 30:1:2; 32 pmol of LPI 13:0, 160 pmol of PS 25:0, 59 pmol of LPS 13:0 and 75 pmol of LPG 13:0.

700 µL of the internal standard mix for human samples contained: 1778 pmol of D7 cholesterol; 2215 pmol of D7- CE 16:0; 1041 pmol of D5-TG 50:0; 595 pmol of D7-DG 33:1; 1376 pmol of PC 25:0; 386 pmol of LPC 13:0; 589 pmol of PE 25:0; 85 pmol of LPE 13:0; 480 pmol of PI 25:0, 140 pmol of PG 25:0; 73 pmol of Cer 30:1:2; 127 pmol of D7-PA 33:1; 78 pmol of U-<sup>13</sup>C LPA 16:0; 271 pmol of d18-SM 30:1:1; 64 pmol of LPI 13:0, 137 pmol of PS 25:0, 74 pmol of LPS 13:0; 75 pmol of LPG 13:0 and 137 pmol of CL 56:4.

## Lipids quantification by shotgun mass spectrometry

The mass spectrometric analysis was performed on a Q Exactive instrument (Thermo Fischer Scientific, Bremen, Germany) equipped with a robotic nanoflow ion source TriVersa NanoMate (Advion BioSciences, Ithaca NY) using nanoelectrospray chips with a spraying nozzle diameter of 4.1 µm. The ion source was controlled by the Chipsoſt 8.3.1 software (Advion BioSciences). Ionization voltage was + 0.96 kV in positive and − 0.96 kV in negative mode; backpressure was 1.25 psi in both modes [14]. The temperature of the ion transfer capillary was 200° C; S-lens RF level was 50%. Each sample was analyzed for 11 min. FT MS spectra were acquired within the range of  $m/z$  400 – 1000 from 0 min to 0.2 min in positive and within the range of  $m/z$ 350 – 1000 from 6.2 min to 6.4 min in negative mode at the mass resolution of R  $m/z$  $_{200}$ =140000; automated gain control (AGC) of 3  $\times$  10<sup>6</sup> and with the maximal injection time of 3000 ms. t-SIM in positive (1.7 to 6 min) and negative (6.4 to 11 min) mode was acquired with R  $_{m/z}$  200=140000; automated gain control of 5  $\times$  10<sup>4</sup>; maximum injection time of 650 ms; isolation window of 20 Th and scan range of m/z 400 to 1000 in positive and m/z 350 to 1000 in negative mode, respectively. The inclusion list of masses targeted in *t*-SIM analyses started at  $m/z$  355 in negative and  $m/z$  405 in positive ion mode and other masses were computed by adding 10 Th increment (i.e. m/z 355, 365, 375) up to *m*/z 1005.

Free cholesterol was quantified by parallel reaction monitoring FT MS/MS within the time range of 0.2 to 1.7 min. For FT MS/MS the number of micro scans was set to 1; precursor isolation window: 0.8 Da, normalized collision energy (nCE):12.5%; AGC: 5 × 10<sup>4</sup> and maximum injection time: 3000 ms. Spectra were pre-processed using repetition rate filtering software PeakStrainer [15] and stitched together by an in-house

developed script [16]. Lipids were identified by LipidXplorer software [17]. Molecular Fragmentation Query Language (MFQL) queries were compiled for PC, PC O-, LPC, LPC O-, PE, PE O-, LPE, PI, LPI, PG, LPG, PA, LPA, PS, LPS, SM, TG, DG, Cer, Chol and CE lipid classes. Lipids were identified by matching of accurately determined intact masses (mass accuracy better than 5 ppm) and quantified by comparing isotopically corrected abundances of their molecular ions with abundances of internal standards of the same lipid class.

#### Oxylipin measurement

#### Sample preparation

Lipidomix GmbH, Berlin performed the oxylipin measurement. Approximately 30 mg of liver tissue was homogenized and spiked with an internal standard consisting of 15- HETE-d8, 14,15-DHET-D11, 14,15-EET-d8, LTB4-d4, 20-HETE-d6, PGE2-d2 (10 ng each; Cayman Chemical, Ann Arbor, USA); this was followed by the addition of 500 µL of methanol and 300 µL of 10 M sodium hydroxide solution and was shook vigorously for 30 min at 60° C for alkaline hydrolysis. The samples were brought to pH 6 with 500 µL of 1 M sodium acetate buffer and acetic acid for hydrolyzed samples. Then, the samples were centrifuged and the supernatant was added to Bond Elute Certify II columns (Agilent Technologies, Santa Clara, USA) for solid-phase extraction; the columns were preconditioned with 3 ml methanol, followed by 3 ml of 0.1 mol/L phosphate buffer containing 5 % methanol (pH 6). The columns were washed with 3 mL methanol/H2O (50/50, vol/vol). For elution, 2 ml of n-hexane: ethyl acetate 25:75 with 1 % acetic acid was used. An SPE Vacuum Manifold was used for extraction. A heating block at 40° C was used to evaporate the eluate under a stream of nitrogen to

obtain solid residues. This was dissolved in 70 µl of acetonitrile/water 50:50 and transferred to an HPLC autosampler vial.

## LC/ESI-MS/MS

The residues were analyzed using an Agilent 1290 HPLC system with a multi-sampler, binary pump, column thermostat (with a Zorbax Eclipse plus C-18, 2.1 x 150 mm), 1.8 µm column using a solvent system of aqueous acetic acid (0.05%) and acetonitrile / methanol (50:50). The elution gradient was started with 5 % organic phase, which was increased within 0.5 min to 56, 5.5 min to 61%, 18.5 min to 87%, 18.6 min to 98% and held there for 6.5 min. The flow rate was 0.3 ml/min and the injection volume were 15 µL. The HPLC was coupled with an Agilent 6490 Triple quadrupole mass spectrometer (Agilent Technologies, Santa Clara, USA) with electrospray ionization source. The following source parameters were applied: drying gas: 115° C / 16 L / min, sheath gas: 390° C / 12 L /min, capillary voltage: 4300 V, nebulizer pressure: 35 psi and nozzle voltage: 1950 V. The analysis was performed with multiple reaction monitoring in negative mode. Further details are given below.









#### Free fatty acid measurement

Free fatty acid measurement was performed by Lipidomix GmbH, Berlin. 30 mg of liver tissue was homogenized in citrate buffer and extracted following Folch's procedure. Extracts were evaporated under nitrogen and dissolved in 100 µL ethanol. An aliquot was diluted 1:10 with isopropanol containing internal standards (C15:0, C21:0 1000 ng/mL, C20:4-d8, C18:2-d4 100 ng/mL, C20:5-d5, C22:6-d5 20 ng/mL Cayman Chemical, Ann Arbor MI). HPLC-measurement was performed using an Agilent 1290 HPLC system with a binary pump, autosampler, and column thermostat equipped with a Phenomenex Kinetex-C18 column 2.6 µm, 2.1 x 150 mm column (Phenomenex, Aschaffenburg, DE), using a solvent system of acetic acid (0.05%) and acetonitrile. All solvents and buffers used were in LC-MS-grade (VWR, Germany). The solvent gradient started at 70 % acetonitrile and was increased to 98 % within 10 min and hold until 14 min with a flow rate of 0.4 mL/min and 5 µL injection volume. The HPLC was coupled with an Agilent 6470 triple quad mass spectrometer with electrospray ionization source operated in negative selected ion mode (parameters are given below).



#### Ion source parameters





The oxylipins and fatty acids were normalized to the total protein concentration. The data was Log2 transformed and the Shapiro test was applied to check for normal distribution. When normality was passed, Welch's t-test was performed and when it did not pass normality Wilcoxon test was performed. Benjamini-Hochberg correction was used to further adjust the P values. The analysis was performed using R 3.6.0.

#### RNA sequencing

### Library preparation

Total RNA with an integrity number between 8-9 was used. mRNA was isolated from 1 µg of total RNA by poly-dT enrichment using the NEBNext Poly(A) mRNA Magnetic Isolation module, according to the manufacturer's instructions. The final elution was done in 15 µl of 2x first-strand cDNA synthesis buffer (NEBnext, NEB). After chemical fragmentation, by incubating the samples for 15 min at 94° C, the samples were subjected to the workflow for strand-specific RNA-Seq library preparation (Ultra Directional RNA Library Prep, NEB). For ligation, custom adaptors were used (Adaptor-Oligo 1: 5'-ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT-3', Adaptor-Oligo 2: 5'-P-GAT CGG AAG AGC ACA CGT CTG AAC TCC AGT CAC-3'). After ligation, adapters were depleted using an XP bead purification system (Beckman Coulter) by adding beads in a ratio of 1:1, followed by an index PCR (15 cycles) using Illumina compatible index primer. After double XP bead purification (with beads added in a ratio 1:1), libraries were quantified on a Fragment Analyser run with an NGS Assay

Kit (Agilent) and loaded on a Nextseq500 flow cell with 75 cycles by pooling the samples in respect of their molarity aiming at 30 million reads per sample.

### Data processing and analysis of RNA sequencing data

Reads were trimmed with flexbar v3.5.0 [18] for adapter contamination. The reads were then aligned to the mouse genome reference GRCm38 with the STAR v2.6.1d [19] and the per-sample 2-pass mapping strategy processing all reads in both passes as described in the documentation. Read counts were summarized to Gencode gene models vM21 [20] with featureCounts v1.6.1 [21] counting primary alignments only. EdgeR (v3.16.5) was used to detect differentially expressed genes with maximal FDR of 0.05 and minimal absolute logFC of 0.5. We discarded genes for which fewer than three samples had counts per million value above 1, calculated normalization factors and robustly estimated the dispersion. Gene Trail 2 1.6 was used to check the enrichment pathways of the deregulated genes.

To perform gene set enrichment analysis (GSEA), differential gene expression data (containing all expressed genes) were ranked using the -log10 transform of the p-value and then signed as positive or negative based on the direction of fold change. Then, the GSEA software (Broad Institute) was used to perform the GSEA pre-ranked analysis (1000 permutations, minimum term size of 15 and maximum term size of 500) [22,23]. As input, the annotated gene sets from Molecular Signatures Database (MSigDB) were used, specifically the GO\_Extra\_Celluar\_Matrix gene set and the Hallmark (v6.2) gene sets [24–26]. Heat maps were generated using the web-based tool Morpheus (https://software.broadinstitute.org/morpheus).

# Statistical analysis

All statistical analyses were performed using GraphPad Prism (GraphPad Inc., La Jolla, CA, USA) software. For the comparison of quantitative measurements, the Mann-Whitney U test or Unpaired t test was used. Data are expressed as mean  $\pm$ standard error of the mean (SEM).

Oxylipins and free fatty acids were normalized to total protein concentration and analysed using R 3.6.0. Data was Log2 transformed and Shapiro test was applied to check for normal distribution. When normality was passed, Welch's t-test was performed and if not, then Wilcoxon test was performed. Benjamini-Hochberg correction was used to further adjust the P values.

Genetic analyses were calculated using an additive model. Differences between the groups were compared by logistic regression analysis adjusted for sex, age, BMI and presence of T2DM (Supplementary Table 1).

## SUPPLEMENTARY TABLES

# Supplementary Table 1: Patient samples used for genetic analysis (NAFLD Cohort)

Demographic, clinical and histological characteristics of the patient samples for the analysis of the phenotypic impact of MBOAT7 rs641738 genotype on NAFLD phenotype. The data for the two patient strata based on BMI are provided. Quantitative parameters are provided as median with interquartile range in brackets.



# Supplementary Table 2: Patient samples used for genetic analysis (NAFLD Cohort)

Additional demographic, clinical and histological characteristics of the patient samples for the analysis of the phenotypic impact of MBOAT7 rs641738 genotype on NAFLD phenotype. Values are given as mean with standard error in parentheses. \*ANOVA; \*\*Pearson Chi-Square; \*\*\* Test of deviation from Hardy-Weinberg-equilibrium



# Supplementary Tables 3: High fat choline deficient methionine low diet (HFCDD)

# composition





# Supplementary Table 4: Differential mRNA expression in 10-week old Mboat7<sup>Δhep</sup> mice livers

The table provides a list of differentially expressed transcripts between Mboat7<sup>Δhep</sup> and Mboat7WT mice using a false discovery rate (FDR) of 0.05. The table content is ordered by FDR in decreasing order. logFC: log fold change.

# Supplementary Table 5: Differential mRNA expression in Mboat7<sup>Δhep</sup> mice livers after 6 weeks of HFCDD feeding

The table provides a list of differentially expressed transcripts between Mboat7<sup>Δhep</sup> and Mboat7WT mice using a false discovery rate (FDR) of 0.05. The table content is ordered by FDR in decreasing order. logFC: log fold change.

#### Supplementary Table 6: Hepatic oxylipin measurements on a chow diet

Hepatic-oxylipin concentrations in Mboat7<sup>WT</sup> and Mboat7<sup>Δhep</sup> mice fed a chow diet for 10 weeks. Abbreviations are provided at the end of the table. The data were analyzed using R and the oxylipins were normalized to the total protein concentration. The data was Log2 transformed and Shapiro test was applied to check for normal distribution. When normality was passed, then Welch's t-test was performed and when it did not pass normality then Wilcoxon test was performed. The Benjamini-Hochberg correction was used to further adjust the P values [27].





30

Oxylipin abbreviations: 13-HODE, 13-Hydroxyoctadecadienoic acid; 9.10-EpOME, 9(10)-Epoxy-12Z-octadecenoic acid; 12.13-EpOME, 12,13-epoxy-9(Z)-octadecenoic acid; 9.10-DiHOME, 9,10-dihydroxy-12Z-octadecenoic acid; 12.13-DiHOME, 12,13 dihydroxy-9Z-octadecenoic acid; 5.6-EET, 5,6-epoxy-8Z,11Z,14Z-eicosatrienoic acid; 8.9-EET, 8,9-epoxy-5Z,11Z,14Z-eicosatrienoic acid; 11.12-EET, 11,(12)-epoxy-5Z,8Z,14Z-eicosatrienoic acid; 14.15-EET, 14(15)-epoxy-5Z,8Z,11Z-eicosatrienoic acid; 5.6-DHET, 5,6-Dihydroxy-8,11,14-icosatrienoic acid; 8.9-DHET, 8,9-dihydroxy-5Z,11Z,14Z-icosatrienoic acid; 11.12-DHET, 11,12-dihydroxy-5Z,8Z,14Zeicosatrienoic acid; 14.15-DHET, 14,15-dihydroxy-5Z,8Z,11Z-eicosatrienoic acid; 5.6- EEQ, 5,6-epoxy-8Z,11Z,14Z,17Z-eicosatetraenoic acid; 8.9-EEQ, 8,9-epoxy-5Z,11Z,14Z,17Z-eicosatetraenoic acid; 11.12-EEQ, 11,12-epoxy-5Z,8Z,14Z,17Zeicosatetraenoic acid; 14.15-EEQ, 14,15-epoxy-5Z,8Z,11Z,17Z-eicosatetraenoic acid; 17.18-EEQ, 17,18-epoxy-5Z,8Z,11Z,14Z-eicosatetraenoic acid; 5.6-DiHETE, 5,6 dihydroxy-8Z,11Z,14Z,17Z-eicosatetraenoic acid; 8.9-DiHETE, 8,9-dihydroxy-5Z,11Z,14Z,17Z-eicosatetraenoic acid; 11.12-DiHETE, 11,12-dihydroxy-5Z,8Z,14Z,17Z-eicosatetraenoic acid; 14.15-DiHETE, 14,15-dihydroxy-5Z,8Z,11Z,17Z-eicosatetraenoic acid; 17.18-DiHETE, 17,18-dihydroxy-5Z,8Z,11Z,14Z-eicosatetraenoic acid; 7.8-EDP, 7,8-epoxy-4Z,10Z,13Z,16Z,19Zdocosapentaenoic acid; 10.11-EDP, 10,11-epoxy-4Z,7Z,13Z,16Z,19Zdocosapentaenoic acid; 13.14-EDP, 13,14-epoxy-4Z,7Z,10Z,16Z,19Zdocosapentaenoic acid; 16.17-EDP, 16,17-epoxy-4Z,7Z,10Z,13Z,19Zdocosapentaenoic acid; 19.20-EDP, 19,20-epoxy-4Z, 7Z,10Z,13Z,16Zdocosapentaenoic acid; 7.8-DiHDPA, 7,8-dihydroxydocosa-4Z,10Z,13Z,16Z,19Zpentaenoic acid; 10.11-DiHDPA, 10,11-dihydroxy-4Z,7Z,13Z,16Z,19Zdocosapentaenoic acid; 13.14-DiHDPA, 13,14-dihydroxy-4Z,7Z,10Z,16Z,19Zdocosapentaenoic acid; 16.17-DiHDPA, 16,17-dihydroxy-4Z,7Z,10Z,13Z,19Zdocosapentaenoic acid; 19.20-DiHDPA, 19,20-dihydroxy-4Z,7Z,10Z,13Z,16Zdocosapentaenoic acid; 5-HETE, 5-hydroxy-6E,8Z,11Z,14Z-eicosatetraenoic acid; 8- HETE, 8-hydroxy-5Z,9E,11Z,14Z-eicosatetraenoic acid; 9-HETE, 9-hydroxy-5Z,7E,11Z,14Z-eicosatetraenoic acid; 11-HETE, 11-hydroxy-5Z,8Z,11E,14Zeicosatetraenoic acid; 12-HETE, 12-hydroxy-5,8,10,14-eicosatetraenoic acid; 15- HETE, 15-hydroxy-5Z,8Z,11Z,13E-eicosatetraenoic acid; 16-HETE, 16-hydroxy-5Z,8Z,11Z,14Z-eicosatetraenoic acid; 17-HETE, 17-hydroxy-5Z,8Z,11Z,14Z-





4Z,7Z,10Z,13Z,16Z,19Z-docosahexaenoic acid; NPDx- nonenzymatic isomer of NPD1 (Neuroprotectin D1).

#### Supplementary Tables 7: Hepatic oxylipin measurements on HFCDD diet

Hepatic-oxylipin concentrations in Mboat7<sup>WT</sup> and Mboat7<sup>Δhep</sup> mice fed a HFCDD for 6 weeks. Abbreviations are provided at the end of Supplementary Table 5. The data was analyzed using R and the oxylipins were normalized to the total protein concentration. The data was Log2 transformed and the Shapiro test was applied to check for normal distribution. When normality was passed, then Welch's t-test was performed and when it did not pass normality then Wilcoxon test was performed. Benjamini-Hochberg correction was used to further adjust the P values [27].





# Supplementary Table 8: Patient samples used for lipidomic analysis (Lipidomic Cohort)

Overview of the patients in the lipidomic cohort. Patients are grouped by MBOAT7 rs641738 genotype and histology using the NAS score. Values are given as counts with percentage and as mean with standard error in parentheses. \*ANOVA; \*\*Pearson Chi-Square; \*\*\* Test of deviation from Hardy-Weinberg-equilibrium.



Thangapandi VR*, et al. Gut* 2021; 70:940–950. doi: 10.1136/gutjnl-2020-320853



Thangapandi VR*, et al. Gut* 2021; 70:940–950. doi: 10.1136/gutjnl-2020-320853

# Supplementary Table 9: Patient samples used for lipidomic analysis (Lipidomic Cohort)

Overview of the patients in the lipidomic cohort. Patients are grouped by histology using the NAS score as the main classification criterion. Values are given as counts and percentage for categorical parameters and as mean with standard error (SE) for continuous parameters. N/A: data not available.





YES 0 0 42 (31%) 24 (49%) 23 (77%)

Supplementary Table 10: Sequences of gene-specific mouse primers used in qRT-PCR assays.



Primer sequences for detecting ER stress was derived from Oslowski et. al [28]

# REFERENCES

- 1 Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology Published Online First: 2005. doi:10.1002/hep.20701
- 2 Brosch M, Kattler K, Herrmann A, et al. Epigenomic map of human liver reveals principles of zonated morphogenic and metabolic control. Nat Commun 2018;9:1–11. doi:10.1038/s41467-018-06611-5
- 3 Strnad P, Buch S, Hamesch K, et al. Heterozygous carriage of the alpha1 antitrypsin Pi∗Z variant increases the risk to develop liver cirrhosis. Gut Published Online First: 2019. doi:10.1136/gutjnl-2018-316228
- 4 Stättermayer AF, Traussnigg S, Aigner E, et al. Low hepatic copper content and PNPLA3 polymorphism in non-alcoholic fatty liver disease in patients without metabolic syndrome. J Trace Elem Med Biol Published Online First: 2017. doi:10.1016/j.jtemb.2016.08.006
- 5 Horvath S, Erhart W, Brosch M, et al. Obesity accelerates epigenetic aging of human liver. Proc Natl Acad Sci U S A Published Online First: 2014. doi:10.1073/pnas.1412759111
- 6 Buch S, Stickel F, Trépo E, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. Nat Genet 2015;47:1443–8. doi:10.1038/ng.3417
- 7 Stickel F, Buch S, Nischalke HD, et al. Genetic variants in PNPLA3 and TM6SF2 predispose to the development of hepatocellular carcinoma in individuals with alcohol-related cirrhosis. Am J Gastroenterol Published Online First: March

2018. doi:10.1038/s41395-018-0041-8

- 8 Franke A, Wollstein A, Teuber M, et al. GENOMIZER: An integrated analysis system for genome-wide association data. Hum Mutat Published Online First: 2006. doi:10.1002/humu.20306
- 9 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using realtime quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402–8. doi:10.1006/meth.2001.1262
- 10 Severgnini M, Sherman J, Sehgal A, et al. A rapid two-step method for isolation of functional primary mouse hepatocytes: Cell characterization and asialoglycoprotein receptor based assay development. Cytotechnology Published Online First: 2012. doi:10.1007/s10616-011-9407-0
- 11 Sales S, Graessler J, Ciucci S, et al. Gender, Contraceptives and Individual Metabolic Predisposition Shape a Healthy Plasma Lipidome. Sci Rep 2016;6:27710. doi:10.1038/srep27710
- 12 Sales S, Knittelfelder O, Shevchenko A. Lipidomics of human blood plasma by high-resolution shotgun mass spectrometry. In: Methods in Molecular Biology. 2017. doi:10.1007/978-1-4939-7057-5\_16
- 13 Liebisch G, Vizcaino JA, Kofeler H, et al. Shorthand notation for lipid structures derived from mass spectrometry. J Lipid Res 2013;54:1523–30. doi:10.1194/jlr.M033506
- 14 Schuhmann K, Almeida R, Baumert M, et al. Shotgun lipidomics on a LTQ Orbitrap mass spectrometer by successive switching between acquisition polarity modes. J Mass Spectrom 2012;47:96–104. doi:10.1002/jms.2031

- 15 Schuhmann K, Thomas H, Ackerman JM, et al. Intensity-Independent Noise Filtering in FT MS and FT MS/MS Spectra for Shotgun Lipidomics. Anal Chem 2017;89:7046–52. doi:10.1021/acs.analchem.7b00794
- 16 Schuhmann K, Srzentic K, Nagornov KO, et al. Monitoring Membrane Lipidome Turnover by Metabolic (15)N Labeling and Shotgun Ultra-High-Resolution Orbitrap Fourier Transform Mass Spectrometry. Anal Chem 2017;89:12857–65. doi:10.1021/acs.analchem.7b03437
- 17 Herzog R, Schuhmann K, Schwudke D, et al. LipidXplorer: a software for consensual cross-platform lipidomics. PLoS One 2012;7:e29851. doi:10.1371/journal.pone.0029851
- 18 Roehr JT, Dieterich C, Reinert K. Flexbar 3.0 SIMD and multicore parallelization. Bioinformatics 2017;33:2941–2. doi:10.1093/bioinformatics/btx330
- 19 Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21. doi:10.1093/bioinformatics/bts635
- 20 Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 2012;22:1760–74. doi:10.1101/gr.135350.111
- 21 Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30:923– 30. doi:10.1093/bioinformatics/btt656
- 22 Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.

Proc Natl Acad Sci U S A 2005;102:15545–50. doi:10.1073/pnas.0506580102

- 23 Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003;34:267–73. doi:10.1038/ng1180
- 24 Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst 2015;1:417–25. doi:10.1016/j.cels.2015.12.004
- 25 Gene T, Consortium O. Gene Ontology : tool for the. Gene Expr 2000;25:25–9. doi:10.1038/75556
- 26 Carbon S, Douglass E, Dunn N, et al. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res 2019;47:D330–8. doi:10.1093/nar/gky1055
- 27 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Published Online First: 1995. doi:10.1111/j.2517-6161.1995.tb02031.x
- 28 Oslowski CM, Urano F. Measuring ER stress and the unfolded protein response using mammalian tissue culture system. In: Methods in Enzymology. 2011. doi:10.1016/B978-0-12-385114-7.00004-0