Research article

JHEP|Reports

Stellate cell expression of SPARC-related modular calcium-binding protein 2 is associated with human non-alcoholic fatty liver disease severity

Frederik T. Larsen,[1,2] Daniel Hansen,[1,2] Mike K. Terkelsen,[1,2] Sofie M. Bendixen,[1,2] Fabio Avolio,[1,2] Charlotte W. Wernberg,[2,3,6] Mette M. Lauridsen,[2,3] Lea L. Grønkjaer,[3] Birgitte G. Jacobsen,[3] Ellen G. Klinggaard,[1,2] Susanne Mandrup,[1,2] Tina Di Caterino,[4] Majken S. Siersbæk,[1,2] Vineesh Indira Chandran,[2,5] Jonas H. Graversen,[2,5] Aleksander Krag,[2,6] Lars Grøntved,[1,2] Kim Ravnskjaer[1,2,*]

[1]*Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense M, Denmark;* [2]*Center for Functional Genomics and Tissue Plasticity (ATLAS), University of Southern Denmark, Odense M, Denmark;* [3]*Department of Gastroenterology and Hepatology, University Hospital of Southern Denmark, Esbjerg, Denmark;* [4]*Department of Pathology, Odense University Hospital, Odense C, Denmark;* [5]*Department of Molecular Medicine, University of Southern Denmark, Odense C, Denmark;* [6]*Center for Liver Research (FLASH), Department of Gastroenterology and Hepatology, Odense University Hospital, Odense C, Denmark.*

JHEP Reports **2023.** https://doi.org/10.1016/j.jhepr.2022.100615

Background & Aims: Histological assessment of liver biopsies is the gold standard for diagnosis of non-alcoholic steatohepatitis (NASH), the progressive form of non-alcoholic fatty liver disease (NAFLD), despite its well-established limitations. Therefore, non-invasive biomarkers that can offer an integrated view of the liver are needed to improve diagnosis and reduce sampling bias. Hepatic stellate cells (HSCs) are central in the development of hepatic fibrosis, a hallmark of NASH. Secreted HSC-specific proteins may, therefore, reflect disease state in the NASH liver and serve as non-invasive diagnostic biomarkers.
Methods: We performed RNA-sequencing on liver biopsies from a histologically characterised cohort of obese patients (n = 30, BMI >35 kg/m^2) to identify and evaluate HSC-specific genes encoding secreted proteins. Bioinformatics was used to identify potential biomarkers and their expression at single-cell resolution. We validated our findings using single-molecule fluorescence *in situ* hybridisation (smFISH) and ELISA to detect mRNA in liver tissue and protein levels in plasma, respectively.
Results: Hepatic expression of SPARC-related modular calcium-binding protein 2 (*SMOC2*) was increased in NASH compared to no-NAFLD (*p.adj* <0.001). Single-cell RNA-sequencing data indicated that *SMOC2* was primarily expressed by HSCs, which was validated using smFISH. Finally, plasma SMOC2 was elevated in NASH compared to no-NAFLD (p <0.001), with a predictive accuracy of AUROC 0.88.
Conclusions: Increased SMOC2 in plasma appears to reflect HSC activation, a key cellular event associated with NASH progression, and may serve as a non-invasive biomarker of NASH.
Impact and implications: Non-alcoholic fatty liver disease (NAFLD) and its progressive form, non-alcoholic steatohepatitis (NASH), are the most common forms of chronic liver diseases. Currently, liver biopsies are the gold standard for diagnosing NAFLD. Blood-based biomarkers to complement liver biopsies for diagnosis of NAFLD are required. We found that activated hepatic stellate cells, a cell type central to NAFLD pathogenesis, upregulate expression of the secreted protein SPARC-related modular calcium-binding protein 2 (SMOC2). SMOC2 was elevated in blood samples from patients with NASH and may hold promise as a blood-based biomarker for the diagnosis of NAFLD.

Introduction

Obesity is a fast-evolving pandemic driven by a sedentary lifestyle and high-calorie diet, alongside genetic risk factors. In Europe, 23% of the population has a BMI of ≥30 kg/m^2 and the prevalence is increasing.[1] Consequently, non-alcoholic fatty liver (NAFL) is the dominant cause of chronic liver disease.[2] Thus, the healthcare and associated economic burden of NAFLD is expected to increase dramatically with increasing rates of obesity.[3,4] NASH can be accompanied by different degrees of fibrosis. If uncontrolled, NASH may progress to cirrhosis and hepatocellular carcinoma.[5]

Liver biopsy readings are the gold standard for diagnosis and staging of NASH and are thus pivotal for monitoring NAFLD.[6,7] Liver biopsies are accompanied by a risk of complications, such as bleeding and pain, and are subject to sampling bias.[8] Thus, they may not capture the heterogenous distribution of hepatic

Keywords: NAFLD; NASH; SMOC2; non-invasive biomarker; hepatic stellate cells; transcriptomics.
Received 5 July 2022; received in revised form 30 September 2022; accepted 15 October 2022; available online 28 October 2022
* Corresponding author. Address: Department of Biochemistry and Molecular Biology, Campusvej 55, 5230 Odense M, Denmark. Tel.: +45 65508906/+45 93979317.
E-mail address: ravnskjaer@bmb.sdu.dk (K. Ravnskjaer).

fibrosis. In addition, the risk of interobserver variability in histologic scores complicates diagnosis and prognosis of NAFLD severity.[9–11] Several non-invasive biochemical and imaging-based methods exist for the diagnostic evaluation of NAFLD.[12] However, most non-invasive biochemical methods exhibit modest accuracy in independent validation. Imaging-based methods, while having moderate to high accuracy, have limited utility due to cost and require well-equipped centres.[12] Thus, non-invasive methods that fully capture NAFLD severity and dynamics are required to replace liver biopsies for the diagnostic and prognostic evaluation of NASH.

A solid understanding of the cellular changes underlying NAFLD progression is of major importance for improvement of diagnostic and prognostic tools, as well as for the development of future treatment regimens. A consequential event in NASH development is the activation of hepatic stellate cells (HSCs). Activation of quiescent HSCs (qHSCs) is orchestrated by a complex series of cellular events initiated by lipotoxicity-induced necroptosis of hepatocytes.[13] The resulting proinflammatory milieu and infiltrating immune cells directly activate qHSCs, which transdifferentiate into fibrogenic myofibroblasts referred to as activated HSCs (aHSCs).[13] As extracellular matrix (ECM) and matricellular proteins produced by HSCs mirror hepatocellular changes in NAFLD, monitoring their production is clinically relevant. Promising blood biomarkers hence include proteins involved in ECM remodelling such as tissue inhibitor of metalloproteinase-1 (TIMP1), microfibrillar-associated protein 4 (MFAP4), and pro-peptide of type III collagen (PRO-C3).[14–17]

The aim of this study was to identify HSC-expressed secreted proteins such as ECM and matricellular proteins that accurately reflect NAFLD severity. This we did by transcriptomic analysis of human liver biopsies from severely obese patients populating the NAFLD disease spectrum followed by confirmation of cell type-specificity and validation at the protein level.

Materials and methods

Study design and participants

Liver and blood samples were obtained from participants enrolled in an ongoing prospective interventional case-control study, PROMETHEUS. The study is a liver biopsy controlled, single-centre study from Denmark. Inclusion criteria is age 18-70 years and a BMI ≥35 kg/m^2. Exclusion criteria are overuse of alcohol (>12 g for women and >24 g for men per day), other known (or discovered) chronic liver disease, use of hepatotoxic medication (glucocorticoids, tamoxifen, amiodarone), short life expectancy, or contraindication towards liver biopsy.

PROMETHEUS is registered at OPEN.rsyd.dk (OP-551, Odense Patient data Explorative Network) and at ClinicalTrial.gov (NCT03535142). The regional committee on health research ethics approved the study and all participant information (S-20170210). All participants gave written informed consent before study participation.

Study data, such as biometrics (incl. height, weight, transient elastography, and BMI), anthropometrics, and pharmacological treatment data were collected prospectively and managed using REDCap (Research Electronic Data Capture) tools hosted at OPEN.rsyd.dk. REDCap is a secure web-based software platform designed to support data capture for research studies (https://www.sdu.dk/en/forskning/open).

Tissue and blood sampling

Liver biopsies were taken before blood samples and elastography scans. Scans were performed and samples taken at the same day with participants being in a 12 h fasting state.

Liver biopsies were sampled under sterile conditions by two trained clinicians from the right liver lobe with a 16-18G Menghini suction needle (Hepafix, Braun, Germany). Samples were immediately released into sterile saline water. A minimum of 15 mm was used for formaldehyde storage and liver histology. Subsequently, remaining tissue was divided into smaller pieces of 5-10 mm and preserved immediately in RNAlater (Sigma-Aldrich, St. Louis, MO) or snap frozen in liquid nitrogen. Blood was drawn by an experienced lab technician. Biochemical analyses were performed according to standard regional protocols and using commercially available kits. All samples were handled by specialised research biochemical technicians and stored at -80 °C.

Histology and staging of NAFLD

All liver biopsies were staged and evaluated by one trained radiologist (T.D.C) blinded to all other data. Scores adhered to the NASH Clinical Research Network (NAS-CRN) classification system for NAFLD: steatosis (0-3), lobular inflammation (0-3) and ballooning (0-2). NAFLD activity score (NAS 0-8) is the sum of these three assessments. Fibrosis was evaluated according to the Kleiner classification,[7] no fibrosis (F0), portal or periportal (F1A-C), perisinusoidal fibrosis in combination with portal and periportal fibrosis (F2), bridging fibrosis (F3), and cirrhosis (F4). Finally, NASH was evaluated according to the steatosis, activity, and fibrosis (SAF) scoring system, no NAFL (SAF 1, steatosis <1), NAFL (SAF 2, steatosis ≥1), and NASH (SAF = 3, steatosis ≥5, ballooning ≥1).[18] From liver biopsies, we defined patients with SAF 1 as no-NAFLD, patients with SAF 2 as NAFL, and patients with SAF 3 as NASH. For predictive modelling we defined patients with NAS ≥4 as severe NAFLD and patients with Kleiner fibrosis grade ≥2 as having fibrosis.

RNA-sequencing and data analysis

Needle biopsies of liver tissue were homogenised using FastPrep-24™ (MP biomedicals, Irvine, CA) and RNA purified using TRIzol-RNA lysis reagent (#T9424, Thermo Fisher, Waltham, MA) according to manufacturer's instructions. Purified RNA was quantified using Qubit 3.0 Fluorometer (Thermo Fisher Scientific) and RNA quality was assessed using Fragment Analyzer 5200 (Agilent, Santa Clara, CA). NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs, San Diego, CA) was used for construction of libraries according to the manufacturer's protocol. RNA was paired-end sequenced using the NovaSeq™ 6000 platform (Illumina, San Diego, CA). Reads were aligned with STAR (v.2.7.8a)[19] to the human genome assembly (GRCh38, Ensembl release 101). FeatureCounts (v.2.0) was employed for exon read quantification.[20] Quality of raw sequencing was assessed using FastQC (v.0.11.9)[21] and MultiQC (v. v1.10.1).[22] Genetic variants were removed from raw sequencing reads using BAMboozle (v.0.5.0).[23] Sanitised reads have been deposited in the NCBI Gene Expression Omnibus repository and are accessible through accession number GSE207310 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE207310).

Table 1. Biometric and biochemistry variables of patient cohort in mean ± SD.

	No-NAFLD	NAFL	NASH
n	5	15	10
Sex (M/F)	0/5	1/14	2/8
Age (years)	37±12	46 ± 12	47 ± 13
BMI (kg/m^2)	46.4 ± 2.8	43.5 ± 4.3	45.3 ± 6.1
ALT (U/L)	20.0 ± 6.8	38.9 ± 26.6	75.1 ± 50.3
AST (U/L)	19.6 ± 4.5	29.3 ± 17.7	75.6 ± 50.3
C-peptide (pmol/L)	1,058.8 ± 365.8	1,461.6 ± 467.5	1,865.6 ± 616.9
Cholesterol (mmol/L)	3.7 ± 0.7	5.1 ± 1.1	4.8 ± 1.0
Triglycerides (mmol/L)	1.0 ± 0.4	1.8 ± 0.7	2.4 ± 1.1
APRI	0.2 ± 0.1	0.3 ± 0.2	0.8 ± 0.6
FIB-4	0.6 ± 0.3	0.7 ± 0.4	1.8 ± 1.3
HOMA-IR	3.8 ± 1.8	8.4 ± 4.9	25.1 ± 26.6
LSM (kPa)	5.24 ± 1.5	9.5 ± 7.0	22.9 ± 15.2
eGFR (ml/min/1.73 m^2)	88.3 ± 17.7	97.6 ± 18.4	96.7 ± 19.9
CRP (mg/L)	17.6 ± 7.0	14.3 ± 15.9	13.4 ± 11.0

ALT, alanine aminotransferase, AST, aspartate aminotransferase, APRI, AST-to-platelet ratio index, CRP, C-reactive protein; eGFR, estimated glomerular filtration rate; FIB-4, fibrosis-4, HOMA-IR, homeostatic model assessment for insulin resistance, LSM, liver stiffness measurement.

Public single-cell integration and annotation

For deconvolution of bulk RNA-seq, three independent public human single-cell RNA-sequencing (scRNA-seq) datasets were retrieved from GEO repositories GSE136103,[24] GSE115469,[25] and GSE158723.[26] Each dataset was initially processed with Seurat (v.4.0.3)[27] to remove low quality cells (200< n <3,000 genes, mitochondrial gene contributions <20%). Moreover, genes expressed in fewer than 50 cells were excluded to remove zero count genes. Following cell removal, normalisation, scaling, and dimensional reduction were performed. Predicted doublets were identified and removed using DoubletFinder (v.2.0.3).[28] Integration was carried out by merging the three processed datasets and correction of the principal component analysis (PCA) embeddings using Harmony (v.0.1.0).[29] Automated cell type annotation using CellTypist (v.0.1.4)[30] was employed for annotation of the complete dataset (trained model reference = *Immune All Low*). Manual correction was done to increase annotation resolution for hepatocytes, cholangiocytes, liver sinusoidal endothelial cells, liver endothelial cells, aHSCs, qHSCs, and vascular smooth muscle cells (VSMCs).

Single-molecule fluorescence *in situ* hybridisation

Single-molecule fluorescence *in situ* hybridisation (smFISH) was performed using the RNAscope Multiplex Fluorescent Reagent Kit v2 assay (#323110, Advanced Cell Diagnostics [ACD], Newark, CA) according to the manufacturer's instructions. In short, formalin-fixed paraffin-embedded liver needle biopsies (n = 6) from obese patients (BMI ≥35 kg/m^2) histological graded as no-NASH (n = 3, 2 × NAS = 0 and 1 × NAS = 1) or NASH (n = 3, 2 × NAS = 7 and 1 × NAS = 8) were sectioned at 3 μm. Tissue sections were deparaffinised using histology-graded xylene and 100% ethanol followed by blockage of endogenous peroxidase using hydrogen peroxide (#322381, ACD). Antigen retrieval (HIER) was performed in 100 °C 1x co-detection target retrieval solution (#323165, ACD) for 30 min followed by protease plus (#322331, ACD) treatment for 40 min. Hs-*SMOC2* (NM_001166412.1, #522921, ACD), Hs-*RGS5* (NM_003617.3, #533421-C2, ACD), and Hs-*LUM* (NM_002345.3, #494761-C4, ACD) probes were then hybridised to the tissue. *SMOC2, RGS5, and LUM* were detected with Opal™ 570, Opal™690, and Opal™520 fluorescent dyes, respectively (1:1,000, #FP1488001KT, #FP1497001KT, #FP1487001KT, Akoya Biosciences, Marlborough, MA). For *FBLN* detection, sections from no-NASH (n = 2, NAS = 0) and NASH (n = 2, NAS = 7/8) were probed with a Hs-*FBLN2* probe (NM_001165035.2, #822761-C2) detected with Opal™ 690 as described above. Sections were counterstained using DAPI (#D9542, stock: 0.5 mg/ml, 1:500, Sigma) and slides were subsequently mounted using Prolong® Diamond Antifade Mountant (#P36961, Thermo Fisher Scientific). Images were acquired on a Nikon confocal A1 microscope (Nikon, Japan) at 20x magnification using NIS-Elements ER version 5.21.03 acquisition software.

Statistical analysis

The Mann-Whitney *U* test was used for multiple comparison of ELISA, normalised count data, and fraction of *SMOC2*$^+$ cells. Bonferroni Hochberg correction was employed to adjust for α-error accumulation. Correlations between two groups (gene expression and clinical variables) were computed using Pearson correlation coefficient (two-tailed *p* value). Poisson regression was employed to model smFISH transcript count data. Predictive modelling of histological grades was assessed by area under the receiver-operating characteristics curve (AUROC) analysis.[31] Optimal cut-off points were estimated using the Youden index. For multiple comparisons, a nominal *p* value ≤0.05 was considered statistically significant. All statistical analyses were performed in R (v. 4.0.3).

Table 2. Histological grading of used obese patient cohort grouped by SAF score.

			Steatosis				Lobular inflammation				Ballooning			Kleiner fibrosis grade				
Group	SAF	n	0	1	2	3	0	1	2	3	0	1	2	0	1	2	3	4
No-NAFLD	1	5	5				4	1			5			4	1			
NAFL	2	15		7	5	3		13	2		15			3	10	2		
NASH	3	10		3	4	3		4	4	2		8	2		4	5	1	

NAFL, non-alcoholic fatty liver; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; SAF, steatosis, activity, and fibrosis.

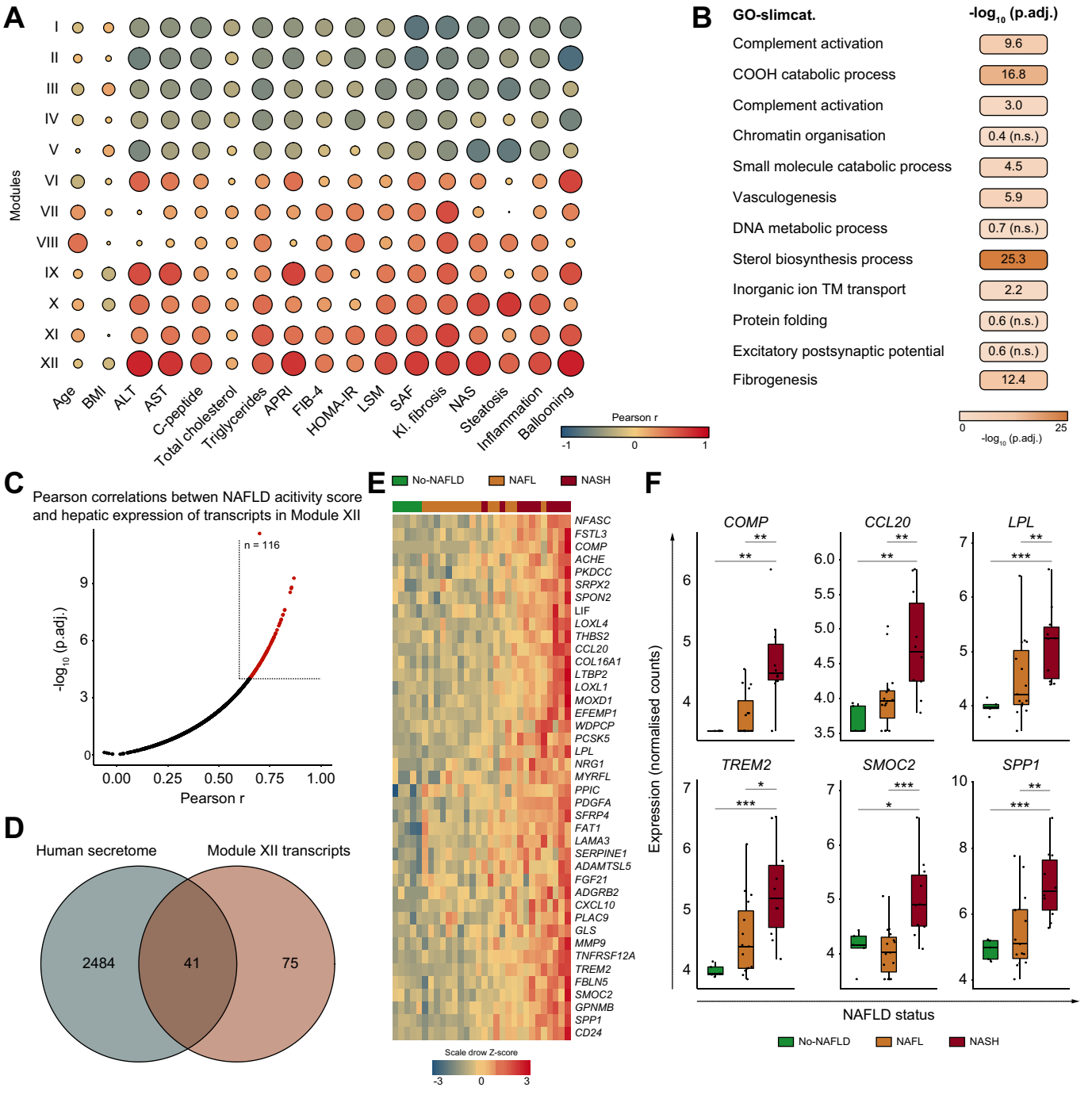

Fig. 1. Hepatic transcriptome profiling and identification of fibrogenesis-related NASH signature transcripts. (A) Pearson correlation (r) between WGCNA identified module eigengenes and biometric, blood chemistry, and histopathologic grades. Size of dos are proportional to the Pearson correlation coefficient. (B) Enriched (*p.adj.* <0.05) and representative (*p.adj.* ≥0.05) GO-slim categories for each module. (C) Pearson correlation between NAS and hepatic expression of transcripts in Module XII in the 30 patients. Transcripts showing high correlation (n = 116, r ≥0.6, -$\log_{10}$(*p.adj.*) ≤4) are shown in red. (D) Venn diagram showing overlap of transcripts highly correlating with NAS and the human secretome obtained from SignalP. (E) Hierarchical clustering of Z-scores of transcripts highly correlating with NAS and overlapping with the human secretome (n = 41). Transcripts found in the human secretome from SignalP are shown on left and NAFLD status of patients (n = 30) is shown as colors on top. (F) Expression of Module XII top 6 transcripts correlating with NAS and overlapping with the human secretome. Hepatic expression of Module XII top 6 transcripts is visualised as boxplots with dots representing biological replicates. Mann-Whitney *U* test was employed to test difference in distribution between groups with Holm-corrected p values. Significance levels are *p <0.05, **p <0.01, ***p <0.001, and **** p ≤0.0001. GO, gene ontology; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; WGCNA, weighted gene co-expression network analysis.

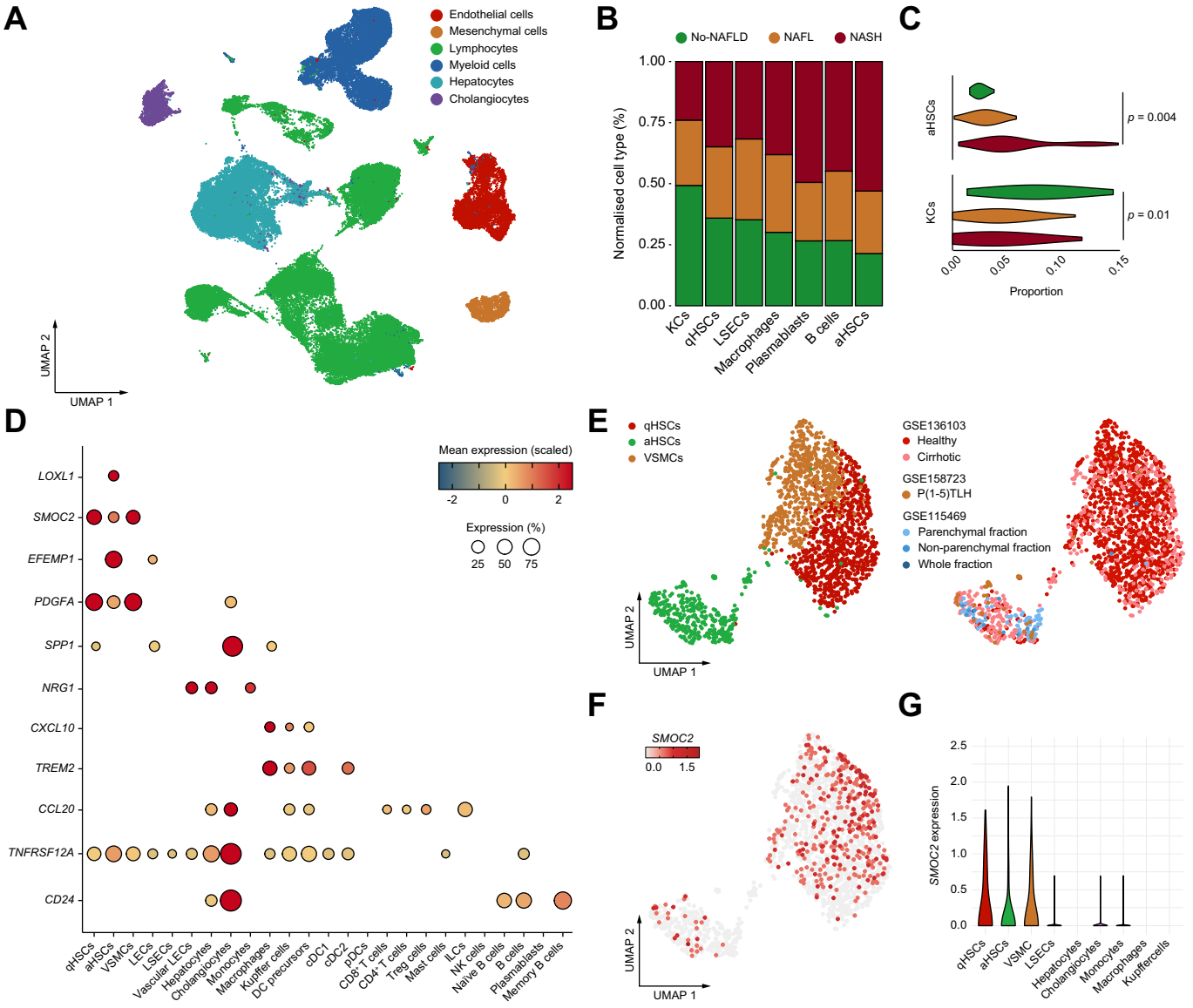

Fig. 2. Identification of cell type-specific gene expression of Module XII secreted proteins. (A) UMAP of human scRNA-seq integrated datasets GSE136103, GSE158723, and GSE115469 (n = 75,632 cells). (B) Estimated abundance of cell types (n = 7, estimated abundance >0.01) from the patient cohort RNA-seq data (n = 30). Estimated abundance is normalised to total estimated abundance for each cell type. Results are represented as stacked barplots and show mean estimated abundance for no-NAFLD, NAFL, and NASH patients. (C) Estimated cell type abundance of aHSCs and Kupffer cells shown as proportions of all estimated cell types. Mann-Whitney *U* test was employed to test difference in distribution between groups with Holm-corrected *p* values. (D) Cell type-resolved expression of Module XII genes encoding secreted proteins ($log_2FC > 2$, expression >5%) shown by dotplot. (E) UMAP showing Leiden clustering of qHSCs, aHSCs, and VSMCs (n = 1,767 cells). Right panel shows the UMAP representation of the human scRNA-seq datasets and treatment groups within each dataset. (F) UMAP showing normalised log_2-expression of *SMOC2* in qHSCs, aHSCs, and VSMCs. (G) Normalised log_2-expression of *SMOC2* in the major hepatic cell types represented as violin plots. UMAP, uniform manifold approximation and projection; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; HSCs, hepatic stellate cells; VSMCs, vascular smooth muscle cells.

Results

Patient characteristics

We generated RNA-seq data from liver needle biopsies obtained from 30 severely obese patients (BMI >35 kg/m^2). The cohort consisted of participants with no-NAFLD (n = 5), NAFL (n = 15), and NASH (n = 10). Clinical biometric and biochemistry features are shown in Table 1. Histological grading of the cohort is shown in Table 2.

PCA was conducted based on expression of all protein-coding genes and the first six components were used to stratify the patient cohort. One-way ANOVA showed an effect on variance from sex and tissue-preservation method (Fig. S1A) for which we corrected in the analysis. PCA showed a separation of transcriptional profiles from patients with NAS ≥5 and patients with NAS ≤2 (Fig. S1B). By comparing patients with NAS ≥5 and NAS ≤2, we identified 1,078 differentially expressed genes (DEGs, $q \leq 0.05$) (Table S1) of which 507 were downregulated and 571 were upregulated. PCA based on these DEGs showed a progressive transcriptional change from no-NAFLD to NASH (Fig. S1C). Moreover, the NASH core transcriptomic signature proposed by

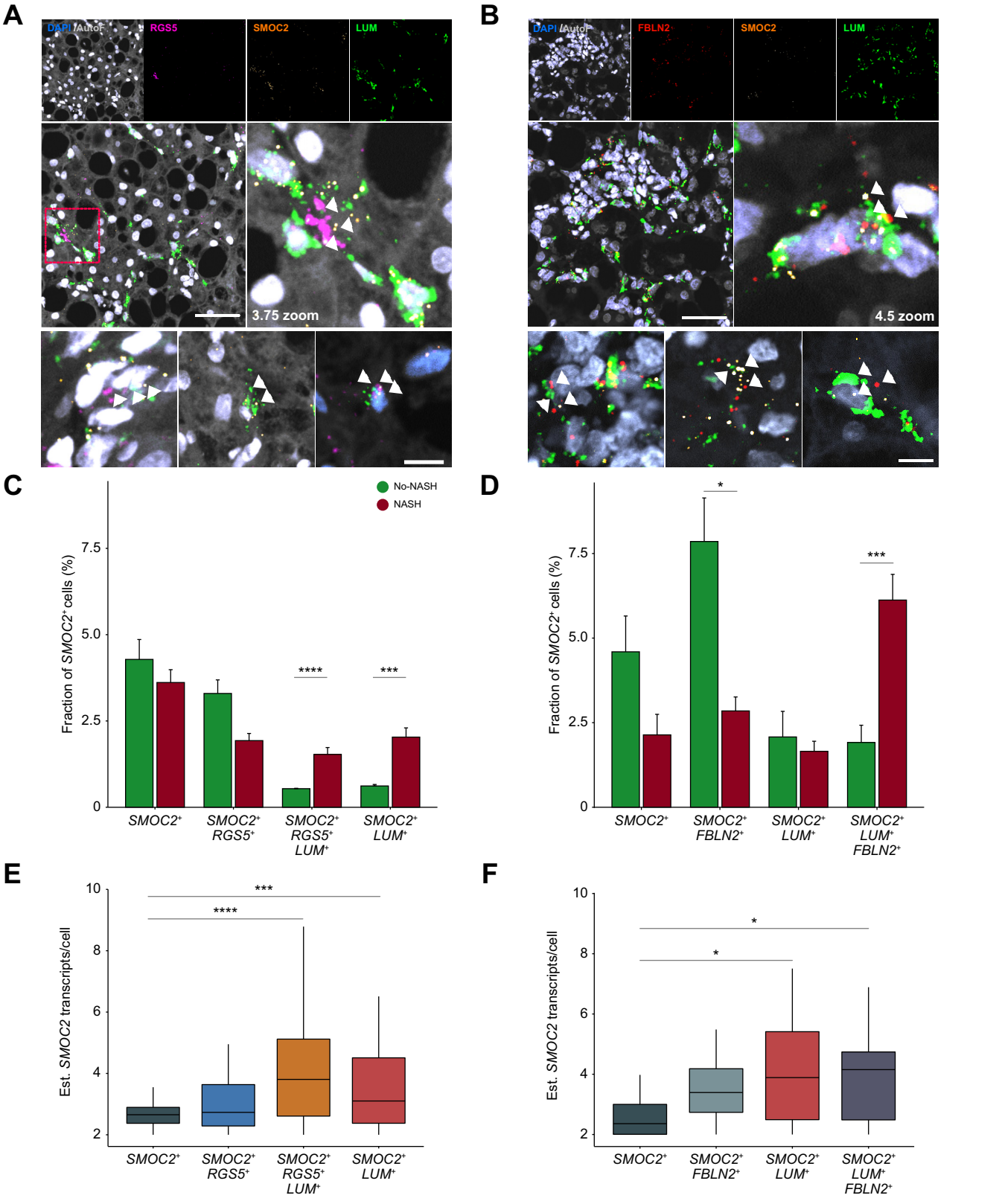

Fig. 3. Single-cell resolution of *SMOC2*, *RGS5*, and *LUM* transcripts show *SMOC2* expression by HSCs in human liver. Confocal images of human liver needle biopsies from severely obese patients histologically graded as NASH (n = 2-3) showing *SMOC2* (orange) co-localised with *RGS5* (magenta) and/or *LUM* (green) (A) and co-localised with *FBLN2* and/or *LUM* (B). Scale bars; middle panels = 50 μm and lower panels = 10 μm. (C) Fraction of total cells/image being *SMOC2*$^+$, *SMOC2*$^+$*RGS5*$^+$, *SMOC2*$^+$*LUM*$^+$, and *SMOC2*$^+$*RGS5*$^+$*LUM*$^+$. (D) Fraction of total cells/image being *SMOC2*$^+$, *SMOC2*$^+$*FBLN2*$^+$, *SMOC2*$^+$*LUM*$^+$, and *SMOC2*$^+$*LUM*$^+$*FBLN2*$^+$. (E) Quantification of *SMOC2* transcripts in *SMOC2*$^+$, *SMOC2*$^+$*RGS5*$^+$, *SMOC2*$^+$*LUM*$^+$, and *SMOC2*$^+$*RGS5*$^+$*LUM*$^+$ cells. (F) Quantification of *SMOC2* transcripts in *SMOC2*$^+$, *SMOC2*$^+$ *FBLN2*$^+$, *SMOC2*$^+$*LUM*$^+$, and *SMOC2*$^+$*LUM*$^+$*FBLN2*$^+$ cells. QuPath was employed to detect and quantify single-, double-, and triple-positive stained cells. Cells

Govaere *et al.* (Fig. S2A) and Terkelsen *et al.* (Fig. S2B) was, in agreement, significantly different in NASH compared to no-NAFLD.[32,33] Together, the transcriptomic profiles stratified the cohort by NAFLD severity, which led us to further analyse fibrogenesis-related NASH signature transcripts.

WGCNA identifies modules of co-expressed genes associated with NAFLD progression

We employed weighted gene co-expression network analysis (WGCNA) for hepatic transcriptome profiling and identification of fibrogenesis-related NASH signature transcripts (Fig. 1). A total of 27 modules of co-expressed genes were identified and merged into 25 modules (Table S2). A significant correlation was found between 12 of the modules and at least one clinical variable (Pearson $r \geq 0.5$, $p \leq 0.05$) (Fig. 1A). High correlation was generally found between modules and histological gradings of biopsies as well as the non-invasive measures LSM and the diagnostic biochemical parameters alanine aminotransferase, aspartate aminotransferase, and aspartate aminotransferase-to-platelet ratio index (Table S3). Gene ontology analysis revealed significant enrichment of pathways in eight of the modules including *sterol biosynthesis process*, *carboxylic acid catabolic process*, and *extracellular structure organisation* (Fig. 1B, Table S4). Module XII had the strongest correlations with histological gradings and diagnostic biochemical parameters (Table S3). Module XII was moreover enriched for genes associated with fibrogenesis pathways (Table S3). Of the 116 Module-XII genes best correlating with NAS ($r \geq 0.6$, *p.adj.* ≤ 0.0001) (Fig. 1C), 41 overlapped with the human secretome (Fig. 1D). Although a recent study has shown GPNMB to be secreted,[34] genes such as *GPNMB*, *ADGRB2*, and *CD24* encode proteins typically not secreted, indicating the human secretome database may contain false positives. Nevertheless, of the 41 genes overlapping with the human secretome, 39 were differentially expressed ($q \leq 0.05$) in NAS ≥5 compared to NAS ≤2 (Table S5) and expression of the 41 genes stratified patients by NAFLD disease status (Fig. 1E). The six genes in Module XII that correlated most strongly with NAS, *COMP*, *CCL20*, *LPL*, *TREM2*, *SMOC2*, and *SPP1*, also present in the human secretome, were significantly induced (*p.adj.* <0.05) in NASH compared to no-NAFLD (Fig. 1F, Table S5). Using WGCNA, we thus identified fibrogenesis-related NASH signature transcripts encoding secreted proteins. We next wanted to resolve expression of the differentially expressed Module XII genes into individual cell types.

Hepatic mesenchymal cells express SMOC2

HSCs are the major source of ECM in the NASH liver and are thus key in the development of fibrosis. We thus analysed public scRNA-seq data for HSC-specific expression of Module XII secretome genes. Three public human scRNA-seq datasets were integrated and reannotated for identification of cell type-specific expression of Module-XII secretome genes (Fig. 2 and Fig. S3A). By automated and manual annotation, based on expression of lineage markers, 25 distinct cell types were identified (Fig. 2A, Fig. S3A). The normalised proportions of seven cell types (cell type abundance >1%) in participants with no-NAFLD, NAFL, and NASH were estimated from bulk RNA-seq data (Fig. 2B). The estimated proportions were higher for aHSCs ($p = 0.004$) and lower for Kupffer cells ($p = 0.01$) in participants with NASH compared to no-NAFLD (Fig. 2C). From the 41 Module-XII secretome genes, we ascertained clear cell type enrichment for 11 genes which had a $\log_2$(fold-change) >2 (Fig. 2D). *LOXL1* was exclusively expressed by aHSCs while *SMOC2* was expressed by qHSCs, aHSCs, and VSMCs. *LOXL1* and *SMOC2* were upregulated ($q < 0.05$) in NAS ≥5 compared to NAS ≤2 (Table S2). We detected the highest increase in gene expression for *SMOC2* ($\log_2$[fold-change] = 3.6, $q = 0.0001$). This prompted us to resolve cell type-specific expression of *SMOC2* by subsetting the mesenchymal cell population (Fig. 2E, Fig. S3B-C). Leiden clustering did not separate VSMCs and qHSCs into distinct clusters but subclustering pointed to HSCs as the main *SMOC2*-expressing cell type (Fig. 2F). Furthermore, the major hepatic cell types, liver sinusoidal endothelial cells, hepatocytes, cholangiocytes, monocytes, macrophages, and Kupffer cells did not express *SMOC2* (Fig. 2G).

SMOC2 expression by human HSCs detected by smFISH

Both *SMOC2* expression and the proportion of aHSCs mirrored NAFLD severity. Since scRNA-seq identified mesenchymal cells as the *SMOC2*-expressing cell population, we next sought to establish aHSC/fibroblast expression of *SMOC2 in situ*. We validated *SMOC2* expression in histologically graded liver needle biopsies from no-NASH (Fig. S4) and NASH using a triplex probe smFISH assay (Fig. 3). *RGS5* was chosen as a marker for qHSCs while *LUM* and *FBLN2* were chosen as markers for aHSCs/fibroblasts (Fig. 3A and 3B).[24,35,36] We observed distinct patterns upon confocal microscopy of *RGS5*, *LUM*, and *FBLN2* transcripts in biopsies from patients with no-NASH (Fig. S4A and S4B) and NASH (Fig. 3A and 3B). Moreover, we detected *FBLN2* in cells lining vessel walls (Fig. S4A-B) and in few parenchymal cells in no-NASH biopsies (Fig. S4A). In NASH biopsies, *FBLN2* was co-expressed with *LUM* and *SMOC2* in cells throughout the hepatic parenchyma, as indicated by white arrows (Fig. 3B). *SMOC2*, *RGS5*, and *LUM* were similarly detected in cells throughout the hepatic parenchyma. In very few cells, *SMOC2* was localised in proximity to larger vessels or lining the vessel wall (Fig. S4C-E). We found no significant differences between biopsies from participants with no-NASH and NASH in fractions of *SMOC2*-single-positive cells (Fig. 3C-D) or *SMOC2*$^+$*RGS5*$^+$ cells (Fig. 3C). The fractions of *SMOC2*$^+$*FBLN2*$^+$ cells, however, were lower ($p < 0.05$) in biopsies from individuals with NASH compared to no-NASH (Fig. 3D), whereas fractions of *SMOC2*$^+$*LUM*$^+$ ($p < 0.001$), *SMOC2*$^+$*RGS5*$^+$*LUM*$^+$ ($p < 0.0001$), and *SMOC2*$^+$*LUM*$^+$*FBLN2*$^+$ cells ($p < 0.001$) were higher. Finally, by quantifying transcripts/cell, we found the number of *SMOC2* transcripts to be higher in *SMOC2*$^+$RGS5$^+$*LUM*$^+$ ($p < 0.0001$) and *SMOC2*$^+$*LUM*$^+$ ($p < 0.001$) compared to *SMOC2* single-positive cells (Fig 3D). Taken together, we validated HSC/fibroblast expression of *SMOC2* by smFISH and, furthermore, found that *RGS5*$^+$*LUM*$^+$ and *LUM*$^+$*FBLN2*$^+$ aHSCs/fibroblasts are the main *SMOC2*-expressing cells in the NASH liver.

with ≥2 *SMOC2* transcripts/cell were considered *SMOC2*-positive cells. Fractions of positive cells are shown as mean + SE (n = 6-12 images/biological replicate) and *SMOC2* transcripts/cell are shown as boxplots (n = 6-12 images/biological replicate). Mann-Whitney *U* test with Holm-correction was employed to test difference between cell fractions. Poisson regression with robust standard errors and *p* value calculation was employed to test difference between estimated transcripts/cell. Significance levels are $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p \leq 0.0001$. HSCs, hepatic stellate cells; NASH, non-alcoholic steatohepatitis.

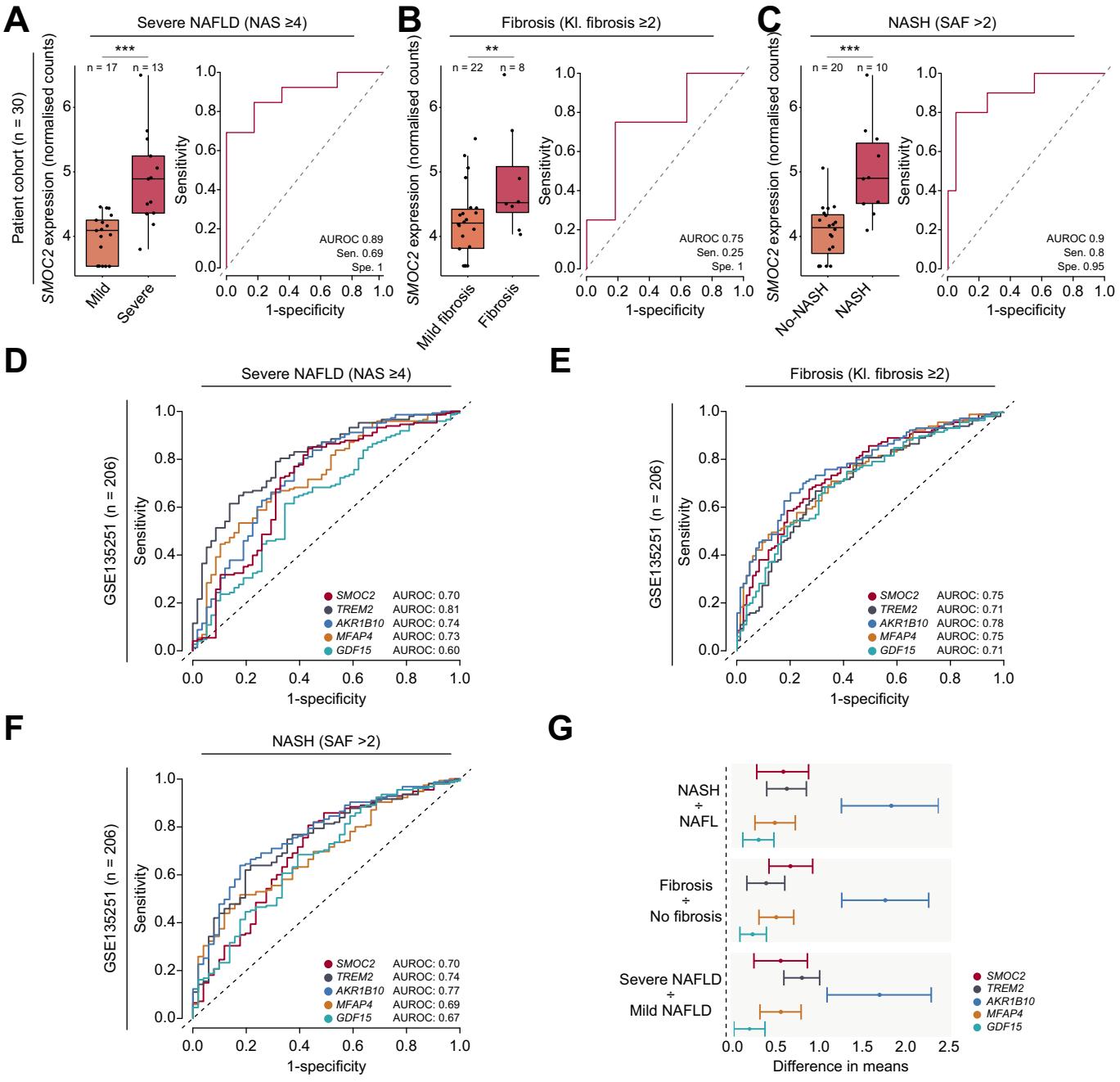

Fig. 4. Predictive modelling of histological grades by hepatic expression of *SMOC2*. (A-C) Association of *SMOC2* expression with NAFLD progression in our patient cohort RNA-seq data (n = 30). (D-F) Association of *SMOC2* expression and previously proposed biomarkers of NAFLD (*TREM2*, *AKR1B10*, *MFAP4*, and *GDF15*) with NAFLD progression in previously described RNA-seq data from a NAFLD multi-centre cohort (GSE135251, n = 206). Patients with NAFLD were categorised by (A and D) severe NAFLD (NAS ≥4), (B and E) fibrosis (Kleiner fibrosis grade ≥2), and (C and F) NASH (SAF >2). Performance of *SMOC2* expression was evaluated using AUROC analysis. Sensitivity and specificity were determined from optimal cut-off points using the Youden index. *SMOC2* expression in our patient cohort segmented into groups are visualised as boxplots with dots representing biological replicates. (G) Expression of *SMOC2*, *TREM2*, *AKR1B10*, *MFAP4*, and *GDF15* in the multi-centre NAFLD cohort are visualised as mean differences between segmented groups with dots representing the mean difference and whiskers representing 95% CIs. Significance levels are $^{**}p$ <0.01 and $^{***}p$ <0.001 (Mann-Whitney *U* test). NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH, non-alcoholic steatohepatitis; SAF, steatosis, activity fibrosis; AUROC, area under the receiver operating characteristic curve; CIs, confidence intervals.

Hepatic expression of SMOC2 discriminates NASH from no-NASH

To evaluate the diagnostic potential of *SMOC2* as a biomarker for NASH, we employed predictive modelling of histological grades NAS, Kleiner fibrosis, or SAF scores using hepatic expression of *SMOC2* (Fig. 4). To validate the predictive accuracy of *SMOC2* in our cohort, we included RNA-seq data from the, currently, most comprehensive multi-centre NAFLD cohort published (GSE135251, n = 206 patients with NAFLD).[32] Moreover, using this multi-centre NAFLD cohort, we benchmarked the

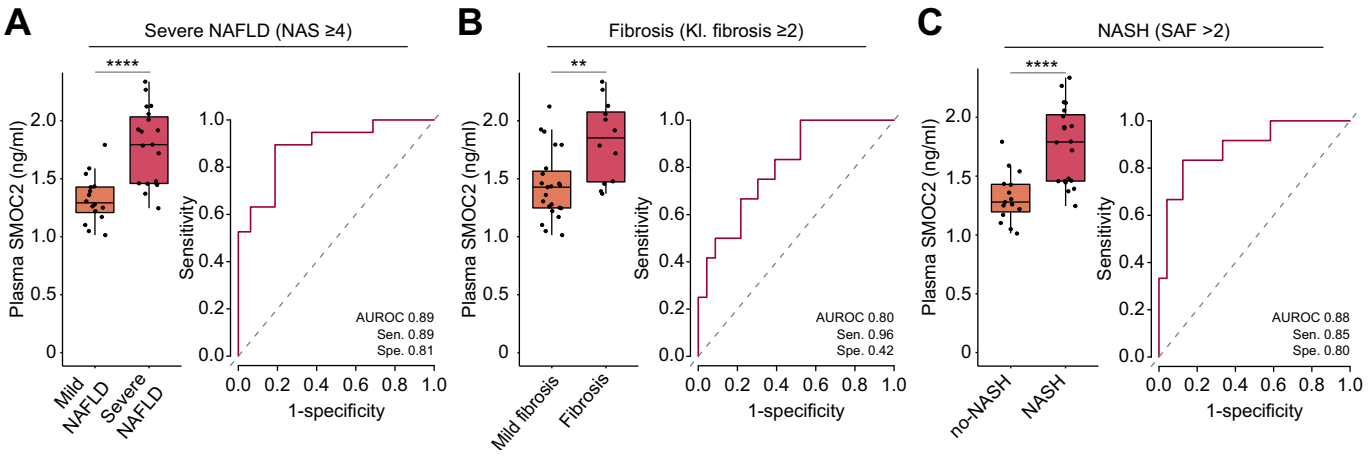

Fig. 5. Predictive modelling of histological grades by plasma SMOC2. (A) Plasma SMOC2 levels in patients with severe NAFLD (NAS ≥4, n = 19) and mild NAFLD (NAS <4, n = 16) (left panel) and predictive accuracy of severe NAFLD (right panel). (B) Plasma SMOC2 levels in patients with fibrosis (Kleiner fibrosis ≥2, n = 12) and mild fibrosis (Kleiner fibrosis <2, n = 23) (left panel) and predictive accuracy of fibrosis (right panel). (C) Plasma SMOC2 levels in patients with NASH (SAF > 2, n = 20) and no-NASH (SAF <2, n = 15) (left panel) and predictive accuracy of NASH (right panel). Plasma SMOC2 performance was evaluated using AUROCs. Sensitivity and specificity were determined from optimal cut-off points using the Youden index. Significance levels are $^{**}p$ <0.01 and **** p ≤0.0001 (Mann-Whitney *U* test). NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH, non-alcoholic steatohepatitis; SAF, steatosis, activity fibrosis; AUROC, area under the receiver operating characteristic curve.

performance of *SMOC2* against recently proposed single-gene biomarkers of NASH and hepatic fibrosis (*TREM2*, *AKR1B10*, *MFAP4*, and *GDF15*) and determined predictive performance of gene combinations (Table S6).[32,37–39] Patients were grouped into severe NAFLD (NAS ≥4) and mild NAFLD (NAS <4), fibrosis (Kleiner fibrosis grade ≥2) and mild fibrosis (Kleiner fibrosis grade <2), and finally NASH (SAF = 3) and no-NASH (SAF ≤2) (Table S7). We sequenced three male needle biopsies, which was insufficient to determine sex-specific differences in hepatic *SMOC2* expression. Thus, we segmented the multi-centre NAFLD cohort into predicted females and males by *XIST* expression (Fig. S5). No significant differences were found between sexes in the segmented groups. In patients with severe NAFLD compared to mild NAFLD, *SMOC2* expression was elevated (our patient cohort; p <0.001, multi-centre NAFLD cohort; p <0.0001) with a predictive accuracy (AUROC) for severe NAFLD of 0.89 (se 0.69, sp 1) in our cohort and 0.7 (se 0.84, sp 0.57) in the multi-centre cohort (Fig. 4A, 4D, and 4G). In patients with fibrosis compared to mild fibrosis, expression of *SMOC2* was elevated (our patient cohort; p <0.01, multi-centre cohort; p <0.0001), with an AUROC for fibrosis of 0.75 (se 0.25, sp 1) in our cohort and 0.75 (se 0.86, sp 0.51) in the multi-centre cohort (Fig. 4B, 4E, and 4G). Finally, in participants with NASH compared to no-NASH, expression of *SMOC2* was elevated (our patient cohort; p <0.001, multi-centre cohort; p <0.0001) with an AUROC for NASH of 0.9 (se 0.8, sp 95) in our cohort and 0.70 (se 0.86, sp 0.51) in the multi-centre cohort (Fig. 4C, 4F, and 4G). Predictive performance of hepatic *SMOC2* expression was similar to those of *TREM2*, *AKR1B10*, *MFAP4*, and *GDF15* (Fig. 4D-F).

The high predictive accuracy of hepatic *SMOC2* expression for histological grades in both our cohort and the multi-centre NAFLD cohort suggested SMOC2 as a potential diagnostic biomarker of NASH.

Plasma SMOC2 levels are associated with NASH severity

Finally, we investigated if the NASH-induced elevation of *SMOC2* expression translated into increased SMOC2 protein in plasma. To first assess the potential contribution of *SMOC2* expression in adipose tissue we quantified *SMOC2* expression by reverse-transcription quantitative PCR in subcutaneous adipose tissue from our cohort (Fig. S6). We found no effect of NAFLD status on the variance in *SMOC2* expression in subcutaneous adipose tissue (p = 0.33). This prompted us to quantify levels of SMOC2 in plasma from a histologically characterised severely obese cohort (n = 35, BMI >35 kg/m^2). This cohort was segmented as described above (Table S8). We did not find any significant differences in plasma SMOC2 levels between sexes in the segmented groups (Fig. S7). Plasma SMOC2 levels were elevated in severe NAFLD compared to mild NAFLD (p <0.0001) with an AUROC for severe NAFLD of 0.89 (se 0.89, sp 0.81) (Fig. 5A). Plasma SMOC2 levels were elevated in patients with fibrosis compared to mild fibrosis (p <0.01) with an AUROC for fibrosis of 0.80 (se 0.96, sp 0.42) (Fig. 5B). Finally, plasma SMOC2 levels were elevated in participants with NASH compared to no-NASH (p <0.0001), with an AUROC for NASH of 0.88 (se 0.85, sp 0.80) (Fig. 5C). Our findings strongly suggest that circulating SMOC2 protein could be a good diagnostic biomarker for NASH in obese patients.

Discussion

Histological grading of liver biopsies is the gold standard for diagnosis and prognostication in NASH.[6,7] Proteins involved in liver ECM remodelling are promising non-invasive biomarkers of fibrosis.[14,15,17] Activated HSCs are believed to be the main contributors to ECM deposition in the NASH liver. Our aim, therefore, was to identify HSC-expressed ECM proteins related to NAFLD progression, using transcriptomic analysis of human liver needle biopsies from severely obese patients populating the NAFLD disease spectrum. Using the powerful WGCNA approach, we identified a module of co-expressed genes relating to fibrogenesis, which correlated with histological grades. Within this module, *SMOC2* encoding a secreted protein was upregulated in the livers of patients with advanced NAFLD. We integrated

publicly available scRNA-seq data to deconvolve our bulk RNA-seq data and resolve *SMOC2* expression in individual cell types. We found that SMOC2 was exclusively expressed in mesenchymal cells. Using smFISH on human liver needle biopsies, we validated *SMOC2* expression in HSCs and found aHSCs/fibroblasts to be the main source of *SMOC2* in the NASH liver. Finally, we found elevated plasma protein levels of SMOC2 in participants with NASH compared to no-NASH, pointing to plasma SMOC2 as a potential non-invasive biomarker of NASH.

SMOC2 encodes a matricellular protein (MCP) of the secreted protein acidic and cysteine-rich (SPARC) family of MCPs.[40] MCPs are non-structural components of the ECM, which bind growth factors, cytokines, and chemokines, thereby playing pivotal roles in ECM-cell signal transduction.[41] Following injury, MCPs are secreted into the ECM to facilitate cell signalling, migration, and adhesion.[42] SMOC2 is associated with fibrosis, inflammation and cell growth, acting downstream of TGF-β1 signalling.[43–45] *SMOC2* has been shown to play an important role in renal fibrosis and has been suggested as a potential biomarker.[46,47] Whole-body *Smoc2* ablation in mice ameliorated diet-induced obesity, hepatic steatosis, and fibrosis.[48] The study also showed elevated *SMOC2* mRNA and protein levels in human steatotic liver biopsies and proposed a fibrogenic role for hepatocyte SMOC2 through direct intracellular interaction with TGF-β1. Using available scRNA-seq data, we identified HSCs and a proportion of VSMCs as the *SMOC2*-expressing cell types of the human liver. Using smFISH, we further pinpointed aHSCs/fibroblasts as the main *SMOC2*-expressing cell types. Expression of *SMOC2* increased with NAFLD severity in our cohort and we identified *LUM*$^+$ and *LUM*$^+$*FBLN2*$^+$ aHSCs/fibroblasts as the main source of this increase in the NASH liver. The fraction of *SMOC2*$^+$*FBLN2*$^+$ cells decreased significantly in NASH compared to no-NASH, which indicates that portal fibroblasts express *SMOC2* in the healthy human liver as expected. In contrast, the fraction of *SMOC2*$^+$*LUM*$^+$*FBLN2*$^+$ cells increased significantly in NASH compared to no-NASH, indicating that aHSCs/fibroblasts express *SMOC2* in the human NASH liver. In the cirrhotic human liver, aHSCs are the main fibrogenic cells,[24] and in different murine fibrosis models aHSCs give rise to 80% to 95% of myofibroblasts.[49] This suggests that increased plasma levels of SMOC2 reflect an expansion of aHSCs during NAFLD progression rather than *SMOC2* expression in hepatocytes, VSMCs or portal fibroblasts. SMOC2 derived from aHSC/fibroblasts could moreover contribute to NAFLD progression. Recently, SMOC2 was shown to promote lung fibroblast-to-myofibroblast transformation *in vitro* through activation of ERK and AKT pathways.[50] Through similar mechanisms, *SMOC2* may also play a role in priming activation and transdifferentiation of HSCs during NAFLD progression, which depends on MEK-ERK signalling.[51]

We tested the performance of *SMOC2* expression in discriminating NASH from no-NASH in our cohort and benchmarked the predictive performance of *SMOC2* against previously proposed biomarkers using a comprehensive public RNA-seq study from a multi-centre NAFLD cohort.[32] In our cohort, *SMOC2* expression exhibited good performance (AUROC 0.79-0.90) in discriminating NAFLD severity defined by NAS score, Kleiner fibrosis grade, and SAF score. In the multi-centre cohort, *SMOC2* expression exhibited modest-to-good performance (AUROC 0.67-0.83), which was comparable to previously proposed biomarkers. Measuring non-invasive biomarkers in liquid biopsies, however, is of high value. In discriminating NASH from no-NASH, we found plasma SMOC2 exhibited good performance in predicting fibrosis (AUROC 0.80, Kleiner fibrosis grade ≥2) and excellent performance in predicting severe NAFLD (AUROC 0.89, NAS ≥4) and NASH (AUROC 0.88, SAF = 3). Variations in specificity ranges (0.80-0.89) and high variations in sensitivity ranges (0.42-0.81) were observed. Restriction on tissue volume in liver biopsies and subsequent underestimation of disease severity as well as sampling bias is a known problem.[10] Predictive performance of plasma SMOC2 may, consequently, be restricted by a discordance with histological scores. Of note, plasma SMOC2 had a low specificity (0.42) in predicting fibrosis. Several patients (n = 8) in the mild fibrosis group were graded with steatosis ≥2, lobular inflammation ≥2, and ballooning ≥1. Thus, in line with a potential role of priming activation of HSCs, plasma SMOC2 may reflect cellular changes related to active fibrogenesis rather than fibrosis *per se*. Our evaluation of plasma SMOC2 as a non-invasive biomarker is, moreover, restricted by the size of our cohort and distribution of histological grades and sex. Thus, a larger cohort is required for further analysis and to establish plasma SMOC2 as a non-invasive biomarker for diagnosis of NASH and, moreover, to exclude potential sex-specific differences in plasma SMOC2 levels.

NASH co-morbidities and overall low tissue-specificity of SMOC2 may compromise the specificity of SMOC2 as a NASH biomarker. Dyslipidaemia-related cardiovascular disease is the leading cause of death in patients with NASH[52,53] and atherosclerosis-related fibrosis[54] may contribute to elevated plasma SMOC2 levels in NASH. An aspect of SMOC2 biology that deserves further investigation. The core fibrosis signalling pathway induced by persistent injury-induced inflammation involves activation and transdifferentiation of mesenchymal cells into scar-forming myofibroblasts.[55] Murine SMOC2 has been implicated in the development of renal,[46] skeletal muscle,[56] and pulmonary fibrosis,[43] suggesting that at least in mice, SMOC2 is part of a core regenerative signalling pathway. Human SMOC2 may have a similar function in tissue regeneration and elevated plasma SMOC2 levels in patients with NASH may, thus, derive from fibrosis in other tissues. In the current study, the estimate glomerular filtration rate of the cohort did not indicate chronic kidney disease and the C-reactive protein levels did not differ between groups. Moreover, we excluded adipose tissue expression of *SMOC2* as a source of elevated plasma SMOC2.

In conclusion, we have identified increased hepatic *SMOC2* expression and concomitant elevated plasma SMOC2 level as a novel diagnostic biomarker for NASH. We described cell type-specific expression of *SMOC2* by HSCs/fibroblasts, thereby linking *SMOC2* to a key cell type in NAFLD and fibrosis progression. Plasma SMOC2 in severely obese individuals may hence reflect liver fibrogenesis and be useful as a diagnostic tool to stratify patients requiring further examination and treatment for NASH. The combination of SMOC2 with other proposed biomarkers, such as TREM2, may aid in the diagnosis and prognostication of patients with NAFLD and deserves further investigation.

Abbreviations

aHSC, activated HSC; AUROC, area under the receiver operating characteristic curve; ECM, extracellular matrix; HSC, hepatic stellate cells; LSM, liver stiffness measurement; MCP, matricellular protein; NAFL, non-alcoholic fatty liver; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; PCA, principal component analysis; qHSC, quiescent HSC; SAF, steatosis, activity, and fibrosis; SE, sensitivity; smFISH, single-molecule fluorescence in situ hybridisation; SMOC2, SPARC-related modular calcium-binding protein 2; SP, specificity; SPARC, secreted protein acidic and cysteine-rich; VSMCs, vascular smooth muscle cells; WGCNA, weighted gene co-expression network analysis.

Financial support

This work was supported by the Danish National Research Foundation (grant DNRF141) to Center for Functional Genomics and Tissue Plasticity (ATLAS), and a Fellowship (D.H.) from the Danish Diabetes Academy, which is funded by the Novo Nordisk Foundation, grant number NNF17SA0031406.

Conflict of interest

All authors declare that they have no conflicts of interest.

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

Authors' contributions

Study conceptualization: A.A.K, M.E.M.L, C.W.W., L.G., and K.R.; *Manuscript preparation*: F.T.L. and K.R.; *NGS data analysis*: F.T.L., M.K.T., and K.R.; *Statistical analysis*: F.T.L.; *Imaging data analysis*: F.T.L. and D.H.; *Laboratory and practical work*: F.T.L., D.H., S.M.B., V.I.C., J.H.G., M.E.M.L, C.W.W., M.S., E.G., L.L.G, and B.G.J.

Data availability statement

Data are accessible through GEO Series accession number GSE207310.

Acknowledgment

We thank Tenna P. Mortensen, Maibrith Wishoff, and Signe Marie Andersen for expert technical assistance. We thank Casimiro Gerarduzzi (affiliation) for valuable intellectual input. Bioimaging was performed at DaMBIC, a bioimaging research core facility at University of Southern Denmark (SDU), established by an equipment grant from the Danish Agency for Science, Technology and Innovation and internal funding from SDU.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhepr.2022.100615.

References

Author names in bold designate shared co-first authorship

[1] Organization WH. WHO European regional obesity report 2022. Copenhagen: WHO Regional Office for Europe; 2022. World Health Organization; 2022.
[2] Younossi ZM, Golabi P, de Avila L, Paik JM, Srishord M, Fukui N, et al. The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: a systematic review and meta-analysis. J Hepatol 2019;71:793–801.
[3] Hagstrom H, Nasr P, Ekstedt M, Hammar U, Widman L, Stal P, et al. Health care costs of patients with biopsy-confirmed nonalcoholic fatty liver disease are nearly twice those of matched controls. Clin Gastroenterol Hepatol 2020;18:1592–1599 e1598.
[4] O'Hara J, Finnegan A, Dhillon H, Ruiz-Casas L, Pedra G, Franks B, et al. Cost of non-alcoholic steatohepatitis in Europe and the USA: the GAIN study. JHEP Rep 2020;2:100142.
[5] Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, McCullough AJ. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. Gastroenterology 1999;116:1413–1419.
[6] Campos GM, Bambha K, Vittinghoff E, Rabl C, Posselt AM, Ciovica R, et al. A clinical scoring system for predicting nonalcoholic steatohepatitis in morbidly obese patients. Hepatology 2008;47:1916–1923.
[7] Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–1321.
[8] McGill DB, Rakela J, Zinsmeister AR, Ott BJ. A 21-year experience with major hemorrhage after percutaneous liver biopsy. Gastroenterology 1990;99:1396–1400.
[9] Vuppalanchi R, Unalp A, Van Natta ML, Cummings OW, Sandrasegaran KE, Hameed T, et al. Effects of liver biopsy sample length and number of readings on sampling variability in nonalcoholic Fatty liver disease. Clin Gastroenterol Hepatol 2009;7:481–486.
[10] Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005;128:1898–1906.
[11] Davison BA, Harrison SA, Cotter G, Alkhouri N, Sanyal A, Edwards C, et al. Suboptimal reliability of liver biopsy evaluation has implications for randomized clinical trials. J Hepatol 2020;73:1322–1332.
[12] Wong VW, Adams LA, de Ledinghen V, Wong GL, Sookoian S. Noninvasive biomarkers in NAFLD and NASH - current progress and future promise. Nat Rev Gastroenterol Hepatol 2018;15:461–478.
[13] Tsuchida T, Friedman SL. Mechanisms of hepatic stellate cell activation. Nat Rev Gastroenterol Hepatol 2017;14:397–411.
[14] Abdelaziz R, Elbasel M, Esmat S, Essam K, Abdelaaty S. Tissue inhibitors of metalloproteinase-1 and 2 and obesity related non-alcoholic fatty liver disease: is there a relationship. Digestion 2015;92:130–137.
[15] Boyle M, Tiniakos D, Schattenberg JM, Ratziu V, Bugianessi E, Petta S, et al. Performance of the PRO-C3 collagen neo-epitope biomarker in non-alcoholic fatty liver disease. JHEP Rep 2019;1:188–198.
[16] Luo Y, Oseini A, Gagnon R, Charles ED, Sidik K, Vincent R, et al. An evaluation of the collagen fragments related to fibrogenesis and fibrolysis in nonalcoholic steatohepatitis. Sci Rep 2018;8:12414.
[17] Saekmose SG, Mossner B, Christensen PB, Lindvig K, Schlosser A, Holst R, et al. Microfibrillar-associated protein 4: a potential biomarker for screening for liver fibrosis in a mixed patient cohort. PLoS One 2015;10:e0140418.
[18] Nascimbeni F, Bedossa P, Fedchuk L, Pais R, Charlotte F, Lebray P, et al. Clinical validation of the FLIP algorithm and the SAF score in patients with non-alcoholic fatty liver disease. J Hepatol 2020;72:828–838.
[19] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
[20] Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30:923–930.
[21] Andrews S. FastQC. 2019 [cited; Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
[22] Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016;32:3047–3048.
[23] Ziegenhain C, Sandberg R. BAMboozle removes genetic variation from human sequence data for open data sharing. Nat Commun 2021;12:6216.
[24] Ramachandran P, Dobie R, Wilson-Kanamori JR, Dora EF, Henderson BEP, Luu NT, et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. Nature 2019;575:512–518.
[25] **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;9:4383.
[26] Payen VL, Lavergne A, Alevra Sarika N, Colonval M, Karim L, Deckers M, et al. Single-cell RNA sequencing of human liver reveals hepatic stellate cell heterogeneity. JHEP Rep 2021;3:100278.
[27] Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell 2021;184:3573–3587 e3529.
[28] McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst 2019;8:329–337 e324.
[29] Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 2019;16:1289–1296.
[30] **Dominguez Conde C**, **Xu C**, **Jarvis LB**, **Rainbow DB**, **Wells SB**, Gomes T, et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. Science 2022;376:eabl5197.
[31] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.
[32] Govaere O, Cockell S, Tiniakos D, Queen R, Younes R, Vacca M, et al. Transcriptomic profiling across the nonalcoholic fatty liver disease

spectrum reveals gene signatures for steatohepatitis and fibrosis. Sci Transl Med 2020;12.
[33] Terkelsen MK, Bendixen SM, Hansen D, Scott EAH, Moeller AF, Nielsen R, et al. Transcriptional dynamics of hepatic sinusoid-associated cells after liver injury. Hepatology 2020;72:2119–2133.
[34] **Gong XM**, **Li YF**, **Luo J**, **Wang JQ**, **Wei J**, Wang JQ, et al. Gpnmb secreted from liver promotes lipogenesis in white adipose tissue and aggravates obesity and insulin resistance. Nat Metab 2019;1:570–583.
[35] Wang ZY, Keogh A, Waldt A, Cuttat R, Neri M, Zhu S, et al. Single-cell and bulk transcriptomics of the liver reveals potential targets of NASH with fibrosis. Sci Rep 2021;11:19396.
[36] Dobie R, Wilson-Kanamori JR, Henderson BEP, Smith JR, Matchett KP, Portman JR, et al. Single-cell transcriptomics uncovers zonation of function in the mesenchyme during liver fibrosis. Cell Rep 2019;29:1832–1847 e1838.
[37] Bracht T, Schweinsberg V, Trippler M, Kohl M, Ahrens M, Padden J, et al. Analysis of disease-associated protein expression using quantitative proteomics-fibulin-5 is expressed in association with hepatic fibrosis. J Proteome Res 2015;14:2278–2286.
[38] **Chandran VI**, **Wernberg CW**, Lauridsen MM, Skytthe MK, Bendixen SM, Larsen FT, et al. Circulating TREM2 as a non-invasive diagnostic biomarker for NASH in patients with elevated liver stiffness. Hepatology 2022;00:1–15.
[39] Hendrikx T, Porsch F, Kiss MG, Rajcic D, Papac-Miličević N, Hoebinger C, et al. Soluble TREM2 levels reflect the recruitment and expansion of $TREM2^+$ macrophages that localize to fibrotic areas and limit NASH. J Hepatol 2022;77:1373–1385.
[40] Vannahme C, Gosling S, Paulsson M, Maurer P, Hartmann U. Characterization of SMOC-2, a modular extracellular calcium-binding protein. Biochem J 2003;373:805–814.
[41] Bornstein P. Matricellular proteins: an overview. J Cel Commun Signal 2009;3:163–165.
[42] Bornstein P, Sage EH. Matricellular proteins: extracellular modulators of cell function. Curr Opin Cel Biol 2002;14:608–616.
[43] Luo L, Wang CC, Song XP, Wang HM, Zhou H, Sun Y, et al. Suppression of SMOC2 reduces bleomycin (BLM)-induced pulmonary fibrosis by inhibition of TGF-beta1/SMADs pathway. Biomed Pharmacother 2018;105:841–847.
[44] Su JR, Kuai JH, Li YQ. Smoc2 potentiates proliferation of hepatocellular carcinoma cells via promotion of cell cycle progression. World J Gastroenterol 2016;22:10053–10063.
[45] Liu P, Lu J, Cardoso WV, Vaziri C. The SPARC-related factor SMOC-2 promotes growth factor-induced cyclin D1 expression and DNA synthesis via integrin-linked kinase. Mol Biol Cel 2008;19:248–261.
[46] Gerarduzzi C, Kumar RK, Trivedi P, Ajay AK, Iyer A, Boswell S, et al. Silencing SMOC2 ameliorates kidney fibrosis by inhibiting fibroblast to myofibroblast transformation. JCI Insight 2017;2.
[47] Schmidt IM, Colona MR, Kestenbaum BR, Alexopoulos LG, Palsson R, Srivastava A, et al. Cadherin-11, Sparc-related modular calcium binding protein-2, and Pigment epithelium-derived factor are promising non-invasive biomarkers of kidney fibrosis. Kidney Int 2021;100:672–683.
[48] Yuting Y, Lifeng F, Qiwei H. Secreted modular calcium-binding protein 2 promotes high fat diet (HFD)-induced hepatic steatosis through enhancing lipid deposition, fibrosis and inflammation via targeting TGF-beta1. Biochem Biophys Res Commun 2019;509:48–55.
[49] Mederacke I, Hsu CC, Troeger JS, Huebener P, Mu X, Dapito DH, et al. Fate tracing reveals hepatic stellate cells as dominant contributors to liver fibrosis independent of its aetiology. Nat Commun 2013;4:2823.
[50] Wang Y, Yang H, Su X, Cao A, Chen F, Chen P, et al. TGF-beta1/SMOC2/AKT and ERK axis regulates proliferation, migration, and fibroblast to myofibroblast transformation in lung fibroblast, contributing with the asthma progression. Hereditas 2021;158:47.
[51] **Marcher AB**, **Bendixen SM**, Terkelsen MK, Hohmann SS, Hansen MH, Larsen BD, et al. Transcriptional regulation of hepatic stellate cell activation in NASH. Sci Rep 2019;9:2324.
[52] Ekstedt M, Hagstrom H, Nasr P, Fredrikson M, Stal P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology 2015;61:1547–1554.
[53] Samuel VT, Shulman GI. Nonalcoholic fatty liver disease as a nexus of metabolic and hepatic diseases. Cell Metab 2018;27:22–41.
[54] Lan TH, Huang XQ, Tan HM. Vascular fibrosis in atherosclerosis. Cardiovasc Pathol 2013;22:401–407.
[55] Henderson NC, Rieder F, Wynn TA. Fibrosis: from mechanisms to medicines. Nature 2020;587:555–566.
[56] Schuler SC, Kirkpatrick JM, Schmidt M, Santinha D, Koch P, Di Sanzo S, et al. Extensive remodeling of the extracellular matrix during aging contributes to age-dependent impairments of muscle stem cell functionality. Cel Rep 2021;35:109223.