Title

Mapping transcriptional heterogeneity and metabolic networks in fatty livers at single-cell resolution.

Journal

iScience, 26(1)

Authors

Coassolo, Laetitia

Liu, Tianyun

Jung, Yunshin

et al.

Publication Date

2023-01-20

DOI

10.1016/j.isci.2022.105802

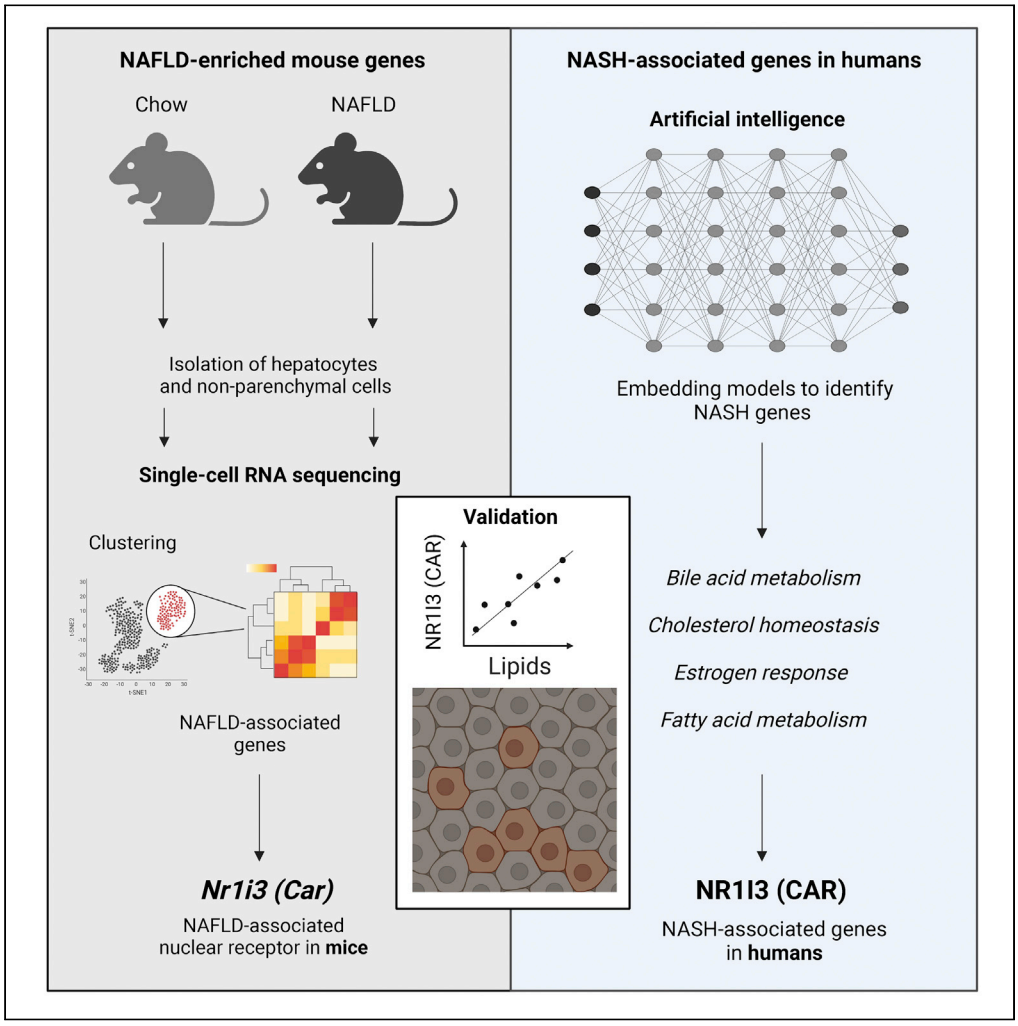
Peer reviewed

iScience

CellPress
OPEN ACCESS

Article

Mapping transcriptional heterogeneity and metabolic networks in fatty livers at single-cell resolution



Laetitia Coassolo, Tianyun Liu, Yunshin Jung, ..., Hannele Yki-Jarvinen, Russ B. Altman, Katrin J. Svensson

katrinjs@stanford.edu

Highlights

Hepatic steatosis is characterized by high hepatocyte heterogeneity

Single-cell analysis reveals *Nr1i3* as a NAFLD-enriched nuclear receptor in mice

Artificial intelligence independently identifies *NR1I3* as a human NASH gene

Nuclear NR1I3 protein expression correlates with steatohepatitis in humans

Coassolo et al., iScience *26*, 105802
January 20, 2023
https://doi.org/10.1016/j.isci.2022.105802

iScience

CellPress
OPEN ACCESS

Article

Mapping transcriptional heterogeneity and metabolic networks in fatty livers at single-cell resolution

Laetitia Coassolo,[1,2,3] Tianyun Liu,[4] Yunshin Jung,[1,2] Nikki P. Taylor,[4] Meng Zhao,[1,2,3] Gregory W. Charville,[1] Silas Boye Nissen,[1,5] Hannele Yki-Jarvinen,[6,7] Russ B. Altman,[8] and Katrin J. Svensson[1,2,3,9,*]

SUMMARY

Non-alcoholic fatty liver disease is a heterogeneous disease with unclear underlying molecular mechanisms. Here, we perform single-cell RNA sequencing of hepatocytes and hepatic non-parenchymal cells to map the lipid signatures in mice with non-alcoholic fatty liver disease (NAFLD). We uncover previously unidentified clusters of hepatocytes characterized by either high or low *srebp1* expression. Surprisingly, the canonical lipid synthesis driver *Srebp1* is not predictive of hepatic lipid accumulation, suggestive of other drivers of lipid metabolism. By combining transcriptional data at single-cell resolution with computational network analyses, we find that NAFLD is associated with high constitutive androstane receptor (CAR) expression. Mechanistically, CAR interacts with four functional modules: cholesterol homeostasis, bile acid metabolism, fatty acid metabolism, and estrogen response. Nuclear expression of CAR positively correlates with steatohepatitis in human livers. These findings demonstrate significant cellular differences in lipid signatures and identify functional networks linked to hepatic steatosis in mice and humans.

INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) and the more severe condition non-alcoholic steatohepatitis (NASH), are now affecting 25% of the global adult population.[1] NAFLD is a progressive metabolic disease characterized by hepatic lipid accumulation, inflammation, fibrosis, and insulin resistance. NAFLD has become the most common form of chronic liver disease in the US and has no pharmacological treatment.[2] In recent years, our molecular understanding of the process of diet-induced hepatic steatosis has emerged.[1,3] Several genes have been identified as risk factors for elevated hepatic lipid accumulation, including a single-nucleotide polymorphism in the *PNPLA3* gene, generating a PNPLA3-I148M variant that increases the risk of developing fatty liver disease[4–6] and an HSD17B13 variant that is protective.[7,8] In addition to genetic factors, environmental factors such as nutrient availability are well-known contributors to hepatic steatosis and NAFLD. Fat consumption,[9–11] excessive fructose intake,[12–14] hyperinsulinemia, and obesity[15,16] positively correlate with NAFLD. However, there is an unmet need to understand the molecular characteristics and changes in cell-type composition and expression in NAFLD.

Interestingly, earlier studies have shown that lipid accumulation is not uniformly induced in all hepatocytes, with most cells having a relatively low accumulation of lipids and fewer cells with very high lipid storage.[17] Elevated intra-hepatocyte lipid content in NAFLD is partly attributable to increased *de novo* lipogenesis,[18] a process largely driven by the transcription factor sterol regulatory element-binding protein-1 (SREBP-1) isoforms *SREBP1a* and *SREBP1c*.[19–21] Yet, whether *SREBP1* is homogenously expressed in hepatocytes, and whether other transcriptional regulators are associated with NAFLD pathogenesis remains underexplored.

Here, by using single-cell analyses, we demonstrate cellular differences in lipid signatures in livers of mice with fatty liver disease. We find that both hepatocyte lipid accumulation and *Srebf1 (Srebp1c)* expression are heterogeneous and identify *Nr1i3*, the constitutive androstane receptor (CAR), as a gene that is enriched in NAFLD livers from mice. While prior reports on CAR have been conflicting,[22] we show that CAR protein levels localized to the nucleus are highly elevated in human livers of patients with steatohepatitis. Our study reveals how NAFLD alters the transcriptomic landscape of hepatocytes during liver pathogenesis and identifies a highly associated NAFLD gene in mice and humans.

[1]Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA

[2]Stanford Diabetes Research Center, Stanford University School of Medicine, Stanford, CA 94305, USA

[3]Stanford Cardiovascular Institute, Stanford University School of Medicine, CA, USA

[4]Department of Bioengineering, Stanford University, Stanford, CA, USA

[5]The Novo Nordisk Foundation Center for Stem Cell Medicine (reNEW), University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen N, Denmark

[6]Department of Medicine, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

[7]Minerva Foundation Institute for Medical Research, Helsinki, Finland

[8]Departments of Bioengineering, Genetics & Medicine, Stanford University, Stanford, CA, USA

[9]Lead contact

*Correspondence: katrinjs@stanford.edu

https://doi.org/10.1016/j.isci.2022.105802

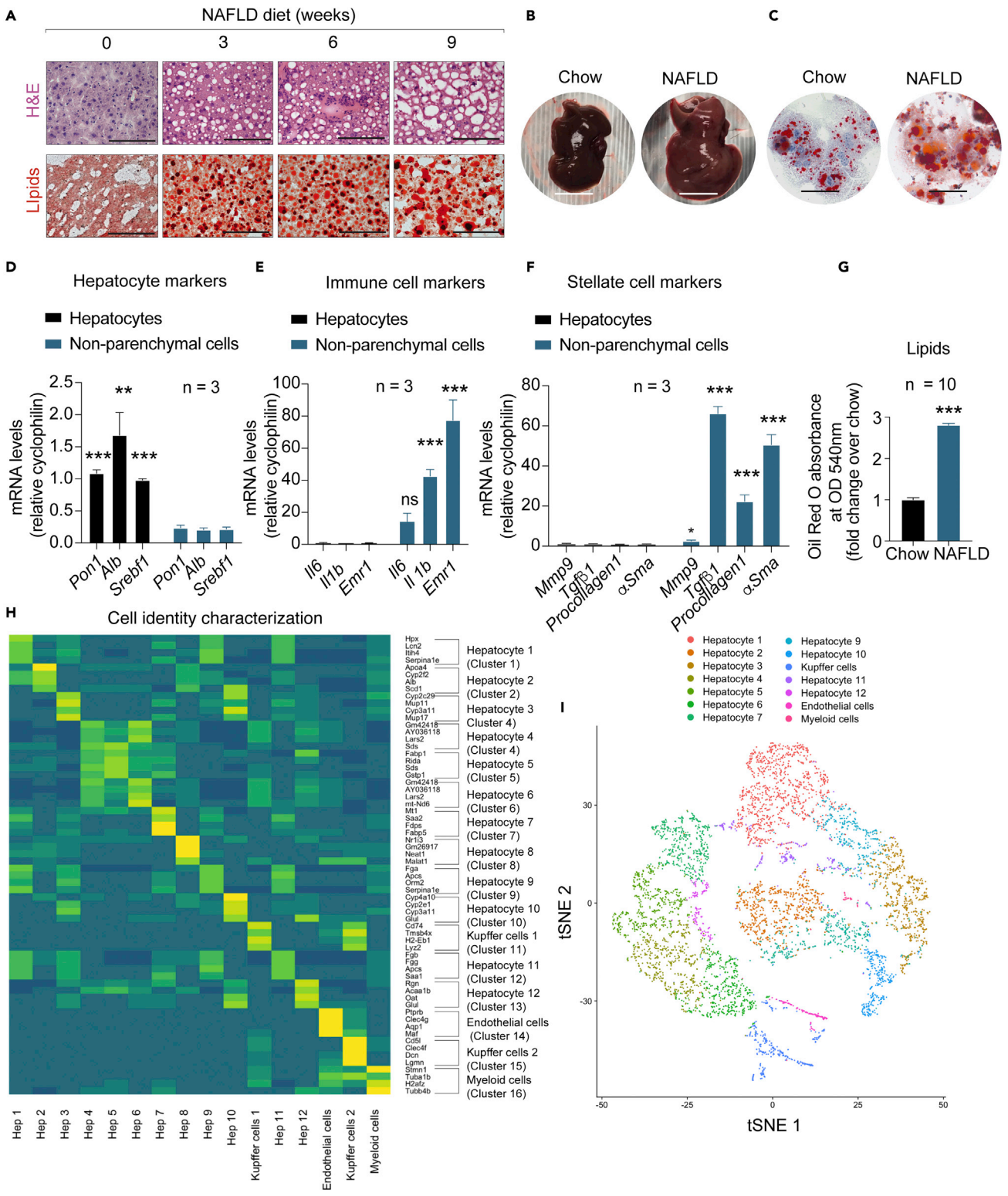


Figure 1. Hepatic steatosis is characterized by high hepatocyte heterogeneity

(A) Histological analyses using H&E and Oil red O (lipids) in livers from mice fed NAFLD diet for 0, 3, 6, and 9 weeks, n = 5 mice per group (10 images per 5 biological replicates in total). Scale bars = 100 μm.

Figure 1. *Continued*

(B) Macroscopic photographs of livers from mice fed chow diet or NAFLD diet for 6 weeks. One representative image is shown from an experiment with n = 5 mice per group.

(C) Oil red O staining demonstrating the cellular accumulation of lipids in hepatocytes isolated from mice fed chow or NAFLD diet for 6 weeks. One representative image is shown from an experiment with n = 5 mice per group.

(D) Gene expression analysis of hepatocyte markers in isolated hepatocytes and non-parenchymal cells (n = 3 samples/group). Data are presented as mean ± SEM *p < 0.05, **p < 0.01, ***p < 0.001 by two-tailed Student's *t* test.

(E) Gene expression analysis of immune cell markers in isolated hepatocytes and non-parenchymal cells (n = 3 samples/group). Data are presented as mean ± SEM *p < 0.05, **p < 0.01, ***p < 0.001 by two-tailed Student's *t* test.

(F) Gene expression analysis of fibrosis/stellate cell markers in isolated hepatocytes and non-parenchymal cells (n = 3 samples/group). Data are presented as mean ± SEM *p < 0.05, **p < 0.01, ***p < 0.001 by two-tailed Student's *t* test.

(G) Oil red O staining quantification of hepatocytes isolated from chow of NAFLD livers (n = 10 samples/group). Data are presented as mean ± SEM of biologically independent samples. *p < 0.05, **p < 0.01, ***p < 0.001 by two-tailed Student's *t* test.

(H) Gene set enrichment criteria for the identification of cell types in the dataset of merged chow and NAFLD clusters identified by Seurat. Values are presented as [log2] fold change. Enrichment was calculated for cells in the cluster over all cells outside the cluster.

(I) Aggregated t-SNE plot of merged chow and NAFLD cell clusters identified by sc-RNA sequencing.

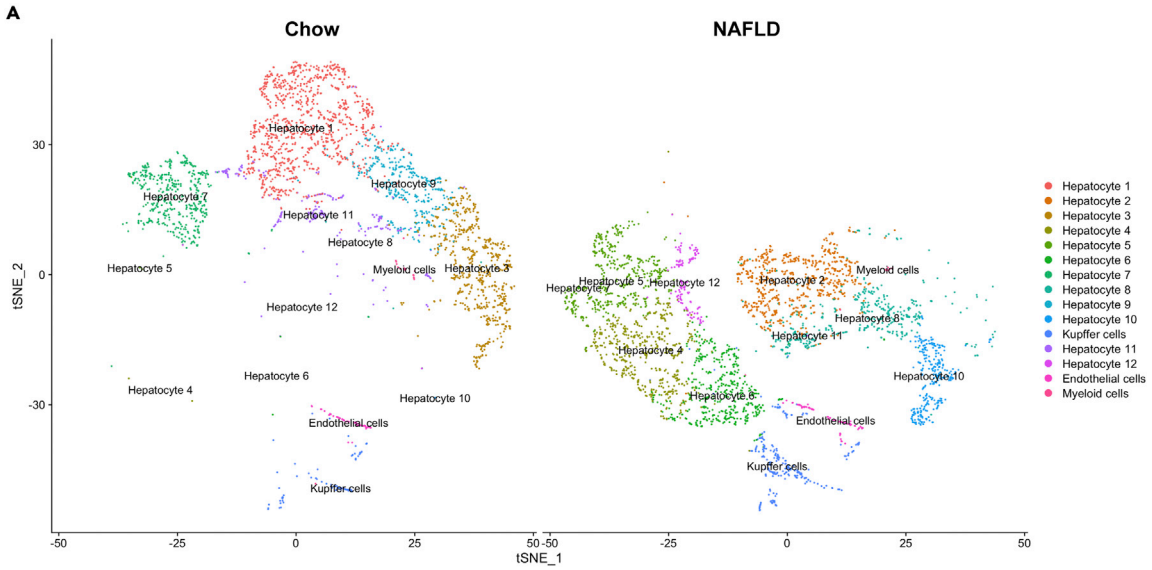
RESULTS

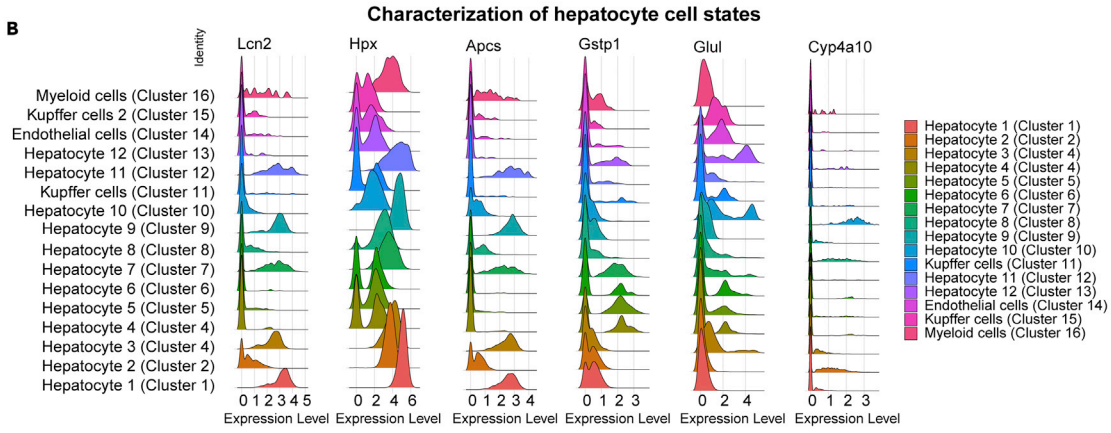
Hepatic steatosis is characterized by high hepatocyte heterogeneity

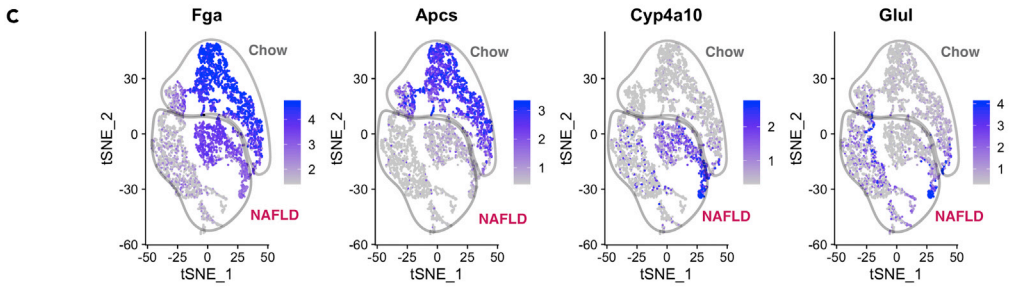
Previous single-cell analyses of livers have shown remarkable heterogeneity in cell populations.[23,24] However, while hepatocyte proteomic analyses have shown dramatic subcellular reorganization of proteins during the development of fatty liver disease,[25] the cellular and metabolic heterogeneity of hepatocytes in NAFLD remains largely uncharacterized.

To dissect the transcriptional changes upon the induction of hepatic steatosis, we first set out to determine the earliest time point of detectable physiological and biochemical changes in liver cell populations isolated from mice fed a high-fructose, high-fat (NAFLD, or amylin) diet[26,27] for 3, 6, and 9 weeks compared with 0 weeks (chow diet). Mice on NAFLD diet demonstrate weight gain starting at 6 weeks (Figure S1A) and induction of ad lib blood glucose levels, indicative of insulin resistance (Figure S1B). Gene expression analyses of livers demonstrate elevated levels of the lipogenesis genes *sterol responsive element binding protein-1c (Srebf1),* demonstrating the activation of lipid synthesis (Figure S1C). In addition, we observe increases in the inflammatory genes *F4/80 (Emr1), tumor necrosis factor alpha (Tnf-α), interleukin-1β (Il1β), transforming growth factor β 1 (Tgfβ1),* and the fibrosis genes *collagen type1 α 1 (Col1α1)* and *matrix metalloproteinase 9* (*Mmp9*), suggestive of the activation of inflammation and fibrosis transcriptional programs in diet-induced NAFLD (Figure S1C). Liver dysfunction was evidenced by a significantly higher liver mass at 9 weeks (Figure S1D) and an increase in plasma ALT levels to 35 U/L at 6 weeks, and 60 U/L at 9 weeks (Figure S1E). Histological analyses of livers confirmed an increase in lipid accumulation by Oil red O staining (Figure 1A) as well as macroscopically larger livers (Figure 1B). Next, we isolated parenchymal hepatocytes and non-parenchymal cell populations in the liver using density gradients.[28] Surprisingly, hepatocytes isolated from mice with established NAFLD demonstrate a large degree of heterogeneity in their ability to accumulate lipids (Figure 1C). The validity of the cell isolation method was further confirmed by the expression of hepatocyte markers *Pon1, Albumin (Alb), and Srebf1* in the hepatocyte fraction (Figure 1D), and the highly enriched immune cell markers *Il-6, Il-1b,* and *Emr1* (Figure 1E), as well as the fibroblast markers *Tgfβ1, Mmp9, Col1α1* and *α-Sma* in the non-parenchymal cell fraction (Figure 1F).

To characterize the molecular changes and heterogeneity at the single-cell level, we next performed single-cell RNA sequencing (scRNA-seq) at 6 weeks, at which time point a >2.5-fold increase in intracellular lipid accumulation was observed (Figure 1G). Freshly isolated hepatocytes and non-parenchymal cells were isolated from chow-or NAFLD-fed mice (n = 5 mice/group), combined, and processed for single-cell RNA sequencing library synthesis preparation after confirming cell viability. Importantly, all samples were sequenced in the same round to avoid the need for batch correction. A total of 5932 cells expressing 17,606 genes were included in the analysis following filtering to remove very high (>0.5) mitochondrial genome transcript ratio, genes detected (UMI count >0) in less than three cells, and cells with very small library size (<1000) (see STAR Methods). Cells were clustered and visualized using aggregated single-cell expression relationship profiles with t-distributed stochastic neighbor embedding (t-SNE) plots using Seurat[29] (Figures 1H and 1I). 16 populations were clustered: 12 distinct hepatocyte populations, and 4 non-parenchymal cell types (Figures 1H and 1I). We identify five main cellular identities in both conditions—hepatocytes, stellate cells, Kupffer cells, myeloid and endothelial cells in line with previously described cell


A
Chow
NAFLD
tSNE_2
tSNE_1
Hepatocyte 1
Hepatocyte 2
Hepatocyte 3
Hepatocyte 4
Hepatocyte 5
Hepatocyte 6
Hepatocyte 7
Hepatocyte 8
Hepatocyte 9
Hepatocyte 10
Kupffer cells
Hepatocyte 11
Hepatocyte 12
Endothelial cells
Myeloid cells


B
Characterization of hepatocyte cell states
Identity
Lcn2
Hpx
Apcs
Gstp1
Glul
Cyp4a10
Myeloid cells (Cluster 16)
Kupffer cells 2 (Cluster 15)
Endothelial cells (Cluster 14)
Hepatocyte 12 (Cluster 13)
Hepatocyte 11 (Cluster 12)
Kupffer cells (Cluster 11)
Hepatocyte 10 (Cluster 10)
Hepatocyte 9 (Cluster 9)
Hepatocyte 8 (Cluster 8)
Hepatocyte 7 (Cluster 7)
Hepatocyte 6 (Cluster 6)
Hepatocyte 5 (Cluster 5)
Hepatocyte 4 (Cluster 4)
Hepatocyte 3 (Cluster 4)
Hepatocyte 2 (Cluster 2)
Hepatocyte 1 (Cluster 1)
Expression Level
Hepatocyte 1 (Cluster 1)
Hepatocyte 2 (Cluster 2)
Hepatocyte 3 (Cluster 4)
Hepatocyte 4 (Cluster 4)
Hepatocyte 5 (Cluster 5)
Hepatocyte 6 (Cluster 6)
Hepatocyte 7 (Cluster 7)
Hepatocyte 8 (Cluster 8)
Hepatocyte 9 (Cluster 9)
Hepatocyte 10 (Cluster 10)
Kupffer cells (Cluster 11)
Hepatocyte 11 (Cluster 12)
Hepatocyte 12 (Cluster 13)
Endothelial cells (Cluster 14)
Kupffer cells (Cluster 15)
Myeloid cells (Cluster 16)


C
Fga
Apcs
Cyp4a10
Glul
Chow
NAFLD
tSNE_2
tSNE_1

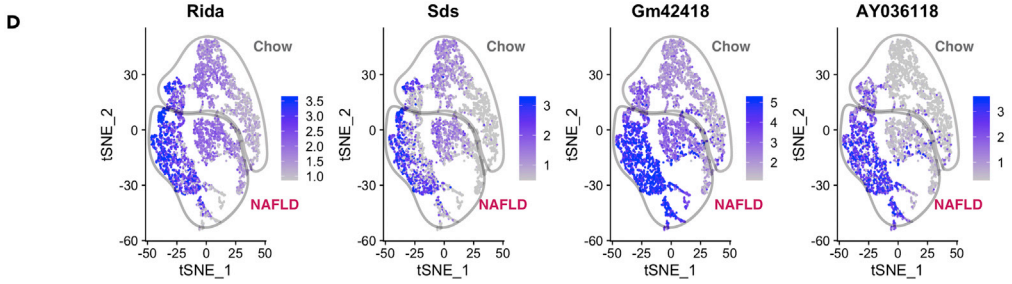

D
Rida
Sds
Gm42418
AY036118
Chow
NAFLD
tSNE_2
tSNE_1

Figure 2. Hepatocytes display dynamic changes during NAFLD progression

(A) t-SNE plot of cell clusters separated by chow and NAFLD identified by Seurat.

(B) Ridgeline plots of the hepatocyte genes *Lcn2, Hpx, Apcs, Gstp1, Glul*, and *Cyp4a10* across all cell clusters identified by Seurat. Values are presented as log2 fold change.

(C) t-SNE plots of hepatocyte-enriched genes *Fga, Apcs, Cyp4a10*, and *Glul* across chow (gray) and NAFLD (red) groups identified by Seurat. Expression in tSNE plots is shown as normalized transcript counts on a color-coded log2 scale.

(D) t-SNE plots of hepatocyte-enriched genes *Rida, Sds, Gm42418*, and *Ay036118* across chow (gray) and NAFLD (red) groups identified by Seurat. Expression in tSNE plots is shown as normalized transcript counts on a color-coded log2 scale.

types in the liver (Figure S1F).[30,31] These data confirm that the liver is characterized by high hepatocyte heterogeneity.

Hepatocytes display dynamic changes during non-alcoholic fatty liver disease progression

To perform more in-depth analyses of the gene signatures in hepatocytes during NAFLD progression, we next clustered cells based on significant cell identity and expression. While Kupffer cells, endothelial cells, and myeloid cells did not markedly change their expression at the six-week time point (Figures S2A–S2D and S3), hepatocytes from NAFLD mice demonstrated large changes in expression as illustrated in the t-SNE plots comparing chow and NAFLD livers (Figure 2A). The overlap between Kupffer cells, endothelial cells, and myeloid cells in chow and NAFLD livers groups argues that these differences are due to batch effects (Figures S2A–S2D). Specifically, we observe distinct gene signatures in the hepatocyte clusters, illustrated in ridgeline plots showing the heterogeneous expression of *Glul* (*glutamine synthetase*) (Figure 2B), a gene encoding for a metabolic enzyme abundantly expressed in hepatocytes.[32,33] Other hepatocyte genes such as *lipocalin-2* (*Lcn2*), *Hpx, Apcs, Gstp1*, and *Cyp4a10* demonstrate similar heterogeneity in expression across hepatocyte clusters, suggesting distinct cellular states (Figure 2B). Furthermore, *fibrinogen alpha* (*Fga*) is enriched in the chow hepatocyte clusters (Figure 2C), while *2-iminibutanoate/2-iminopropanoate deaminase (Rida*), and *serine dehydratase (Sds),* two enzymes involved in amino acid metabolism, are enriched in NAFLD hepatocyte clusters (Figure 2D). Interestingly, two genes, *Gm42418* and *Ay036118* were enriched in NAFLD hepatocyte clusters, indicating a potential function during NAFLD (Figure 2D). These analyses demonstrate that NAFLD is associated with dramatic changes in lipogenic gene expression that are heterogeneously distributed even within a given cell type.

Hepatocyte populations display large heterogeneity in *Srebf1* expression, but *Srebf1* expression does not predict hepatic lipid levels

Srebf1, which is under the direct control of both insulin and fructose is one of the most well-described transcriptional drivers of lipogenesis.[5,34,35] Interestingly, under both chow and NAFLD conditions, there were two main clusters defined by high or low expression of *Srebf1* (*Srebp1c*) (Figures 3A and 3B). Notably, the high expression of *lxra (Nr1h3*) and *Chrebp (Mlxipl)* in *Srebf1*^{high} clusters suggests a broad activation of the classical lipid synthesis pathways (Figures 3A and 3B). The general hepatocyte marker albumin (*Alb)* was equally expressed across hepatocyte clusters, suggesting that the NAFLD state specifically induces the expression of genes involved in lipid synthesis. This finding was unexpected and led us to question whether these changes in expression could explain the differences in lipid accumulation seen in these hepatocytes. Nuclear Srebp1 staining revealed a highly heterogeneous expression in both chow and NAFLD when Srebp1^{high} and Srebp1^{low} cells were quantified (Figures 4A–4C). To correlate lipid accumulation with Srebp1 protein levels in Srebp1^{high} and Srebp1^{low} cells in intact liver sections, we performed co-immunostainings of Srebp1 and the lipid droplet dye LipidSpot. Expectedly, only negligible levels of lipid droplets could be detected under chow conditions, while lipid droplet accumulation and increased lipid droplet size were observed in NAFLD at 6 weeks. Notably, nuclear Srebp1 protein expression did not co-localize with the areas that demonstrated a high density of lipid droplets (Figures 4D and 4E). These data support that both Srebp1^{high} and Srebp1^{low} cells with robust lipid droplet accumulation in NAFLD are present, suggesting that alternative lipogenic drivers might exist.

Identification of *Nr1i3* (*Car*) as a non-alcoholic fatty liver disease-enriched nuclear receptor in mice

To identify other transcriptional regulators that might play a role in NAFLD, we next compared the liver transcriptome between human and mouse livers using a previously published human dataset from Bader et al. in combination with the current mouse single-RNA sequencing dataset. Out of 21,128 detected genes

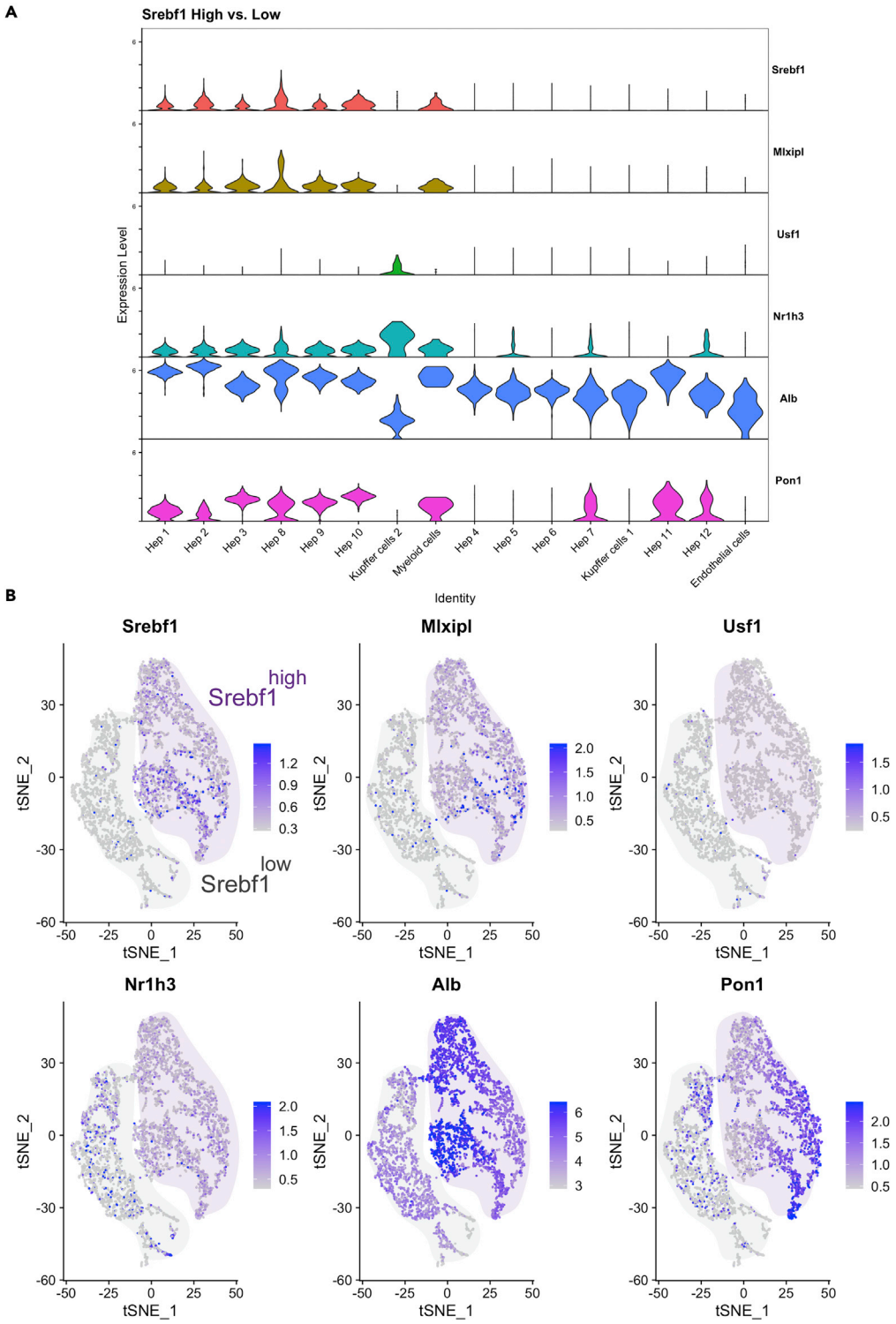


Figure 3. Hepatocyte populations display large heterogeneity in *Srebf1* expression

(A) Violin plots of *Srebf1*, *Chrebp (Mlxipl)*, *Usf1*, *lxrα (Nr1h3)*, *Alb*, and *Pon1* in cell clusters. Values are log2 expression levels.

(B) t-SNE plots of *Srebf1*^{high} and *Srebf1*^{low} clusters displaying the cell type localization of *Srebf1*, *Chrebp (Mlxipl)*, *Usf1*, *lxrα (Nr1h3)*, *Alb*, and *Pon1*.

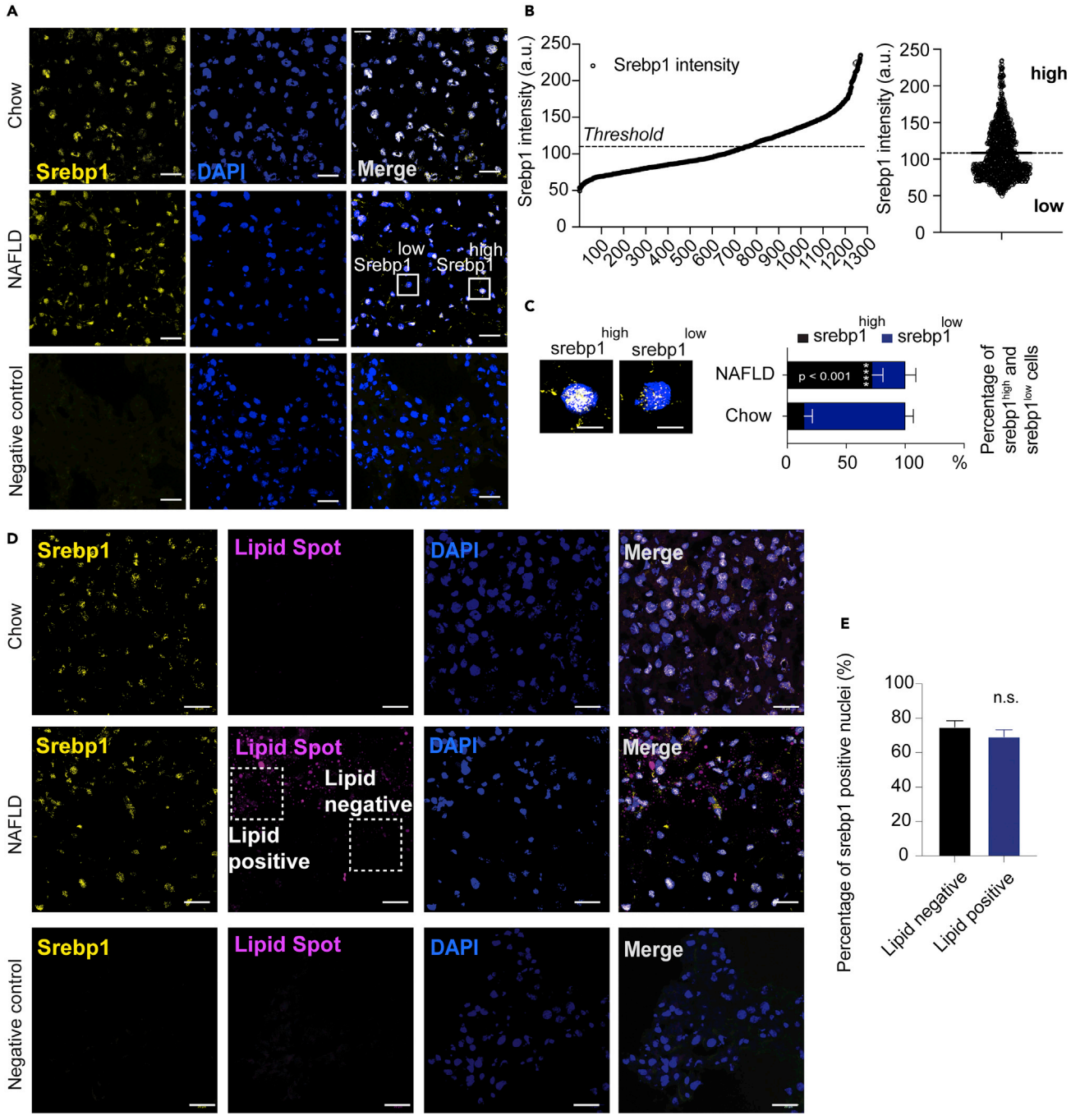


Figure 4. Hepatic Srebp1 levels are heterogeneous and do not correlate with lipid levels

(A) Confocal microscopy analyses of Srebp1 (yellow), and nuclei (blue) in liver tissues in chow and NAFLD (6 weeks) livers. Scale bar denotes 25 μm.

(B) Threshold for Srebp1^{high} cells and Srebp1^{low} cells based on intensity.

(C) Representative images of Srebp1^{high} and Srebp1^{low} hepatocytes are shown by enlarged pictures and quantification of Srebp1^{high} and Srebp1^{low} populations (n = 10 images per group). Data are presented as mean ± SEM *p < 0.05 by two-tailed Student's *t* test.

(D) Confocal microscopy analyses of Srebp1 (yellow), lipid droplets (magenta) and nuclei (blue) in liver tissues in chow and NAFLD (6 weeks) livers. Scale bar denotes 25 μm.

(E) Srebp1 positive cells were quantified from ten lipid^{high} or lipid1^{low} representative areas of NAFLD livers. Bar graph shows the fraction of Srebp1-positive nuclei within lipid-negative or lipid positive areas. Data are presented as mean ± SEM of n = 10 images per group and by two-tailed Student's *t* test.

in either dataset, 12,398 genes were present in both humans and mice (Figures 5A and S4A) and correlated well in expression (Figure S4B). In the mouse dataset, we next analyzed the differentially expressed genes between NAFLD and chow across all hepatocyte populations (Figure 5B and Table S1) and identified *Nr1i3* (nuclear receptor subfamily 1, group I, member 3), or constitutive androstane receptor (*Car*) as the only transcription factor among the top NAFLD-enriched genes (Figure S5A and Table S1). These data are consistent with prior studies of the global Car-KO mice, which are protected from hyperlipidemia and diet-induced fibrosis.[36,37] However, the role of CAR in human NAFLD and steatohepatitis is unclear. Analysis of the distribution of *Nr1i3* expression reveals that *Nr1i3* is heterogeneously expressed but present in all hepatocyte populations (Figure S5B). When performing single cluster analysis of the *Nr1i3* expression, two main populations appear to be dominant: $Nr1i3^{high}$ (clusters 2, 3, 8, 9, 10, 12) and $Nr1i3^{low}$ (clusters 1, 4, 5, 6, 7, 11) (Figure 5C). Interestingly, five out of six $Nr1i3^{high}$ hepatocyte clusters overlap in expression with $Srebp1^{high}$. We next sought to identify genes that correlated with *Nr1i3* expression. The analysis revealed that 48 genes positively and significantly correlated with *Nr1i3* expression with a correlation score higher than 0.63 (Figure 5D). Among these genes, some have previously been associated with lipid metabolism and NAFLD progression, including *Tsukushi* (*Tsku*).[38] Future studies on the relationship between *Nr1i3* and these co-expressed genes, will be important to understand the function of *Nr1i3* in NAFLD.

A predictive model independently identifies *NR1I3* (*CAR*) as a human non-alcoholic steatohepatitis gene

Independently, using Artificial Intelligence (AI) embedding methods, we created a predictive model to identify NAFLD/NASH-related genes and pathways with the hypothesis that the predicted genes would modulate relevant liver biology. To identify disease-associated genes, we built a predictive NASH model based on the proximity of the genes to functional modules in our embedded map of protein-protein interactions (PPIs) (see STAR Methods and Figure S6A). We performed 10-fold cross-validation and systematically varied the parameters of the model. We tested 308 NASH genes from the Be-Free dataset from DisGeNET,[39] and 134 NASH gene set identified experimentally (Table S2). Our model achieves AUC of 0.82 and 0.83, respectively in recovering these NASH disease genes (Figure S6B). In seeking key genes, we were particularly interested in transcription factors that might drive the steatohepatitis processes of lipogenesis, fibrosis, and others. To estimate a gene's association with the NASH model, an empirical cutoff of 0.6 was used. Using the above NASH model, *CAR* (human gene name *NR1I3*) was independently identified as a transcription factor with high likelihood (p-value 0.02) to associate with NASH in humans with a score of 0.86. From a group of 1165 transcription factors, 22 are computationally identified as liver expressed, and 12 out of 22 are associated with NASH. Interestingly, *CAR* interacts with four critical functional modules: Cholesterol Homeostasis, Bile Acid Metabolism, Fatty Acid Metabolism, and Estrogen Response (Figure 6).

Nuclear constitutive androstane receptor localization correlates with steatohepatitis in humans

Given that *CAR* was independently identified in both our transcriptional single-cell analysis and AI embedding model, we next sought to determine whether CAR expression levels correlated with NAFLD in human livers. 26 liver sections from 13 patients with no histopathological abnormality and 13 patients with histological steatohepatitis were included in the analysis (see STAR Methods). There was no significant difference in body mass index (BMI) (Figure 7A), age (Figure 7B), or sex (Figure 7C) between the groups. The steatohepatitis group had higher triglyceride levels (Figure 7D), alanine aminotransferase activity (ALT) (Figure 7E), and aspartate transaminase (AST) activity (Figure 7F) and were diagnosed with histologic grades of 1–2 (Figure 7G) and histologic stage of 0–3 (Figure 7H). Using a specific validated antibody for CAR,[40] we quantified the pixel intensity (Figure 7I) of CAR nuclear staining across the 26 liver samples. Representative microscopic images validate the lipid droplet size and nuclear localization of CAR (Figure 7K), with no staining in the negative control (Figure S7A). Interestingly, we found that CAR protein levels were > 2-fold higher in livers with steatohepatitis (Figures 7J and 7K), and positively correlated with lipid droplet size and steatohepatitis ($R^2 = 0.51$, $p = 0.008$) (Figures 7L and 7M). This was in contrast to no correlation in the gene expression in a separate cohort, suggesting that transcript levels of *NR1I3* do not necessarily predict protein expression (Figures S7B–S7I). Lastly, correlation analyses show that CAR demonstrates a weak, but significant, correlation with ALT and AST (Figures S7J and S7K), while BMI did not significantly correlate with CAR levels (Figure 7N). These results show that nuclear CAR protein expression correlates with steatohepatitis in humans independently of BMI and confirm a potential role for CAR in hepatic lipid accumulation in NAFLD. In conclusion, we here use three orthogonal models; single-cell analyses, human liver expression, and computational network prediction to predict and validate new genes

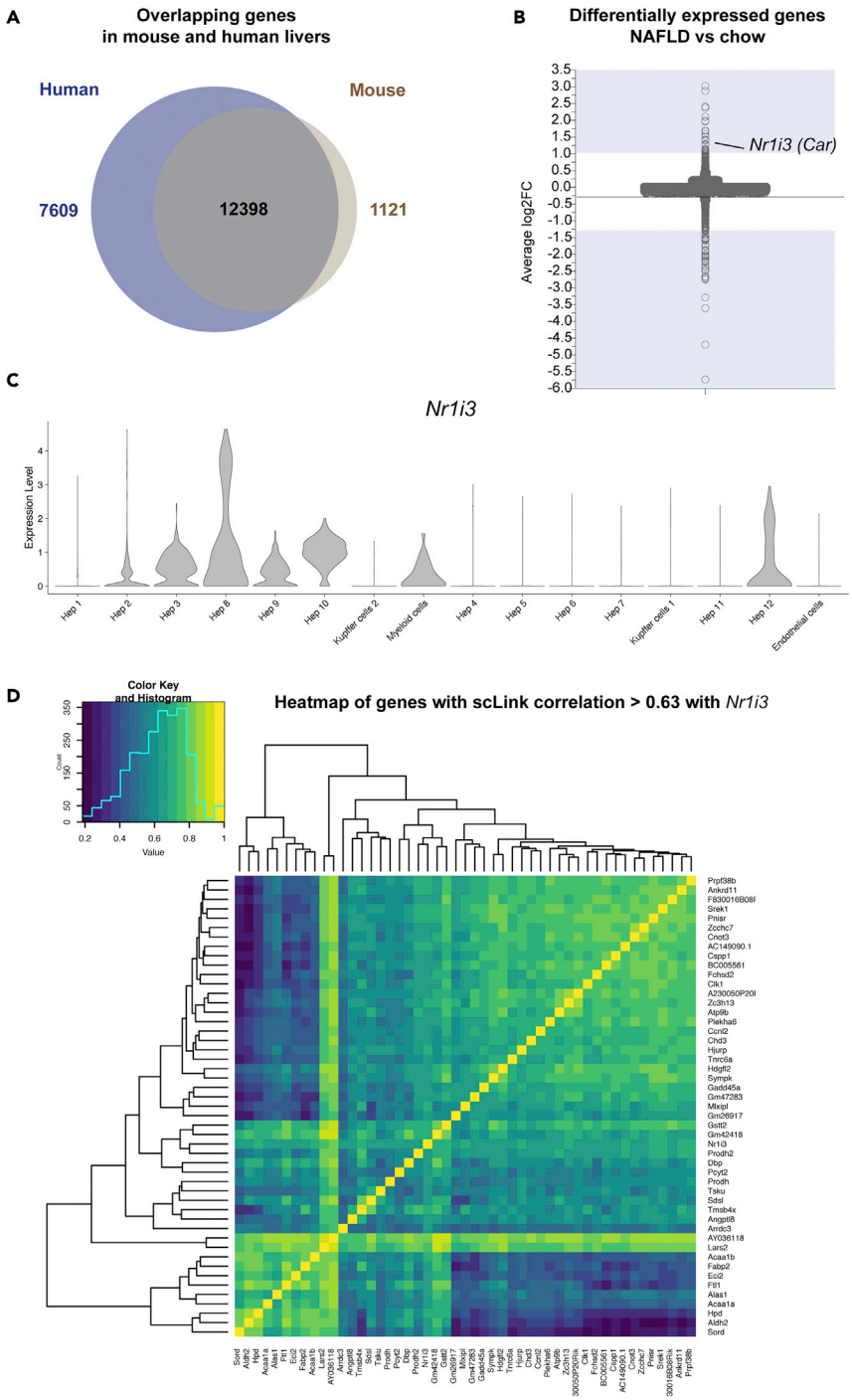


Figure 5. Identification of *Nr1i3* (*Car*) as a NAFLD-enriched nuclear receptor in mice

(A) Overlap of human and mouse genes detected across two independent single-cell RNA seq datasets.

(B) Differentially expressed genes in NAFLD vs chow aggregated across all clusters identified in the single-cell RNA sequencing dataset expressed as average [log2 fold change (FC)].

(C) Violin plots of *Nr1i3 (Car)* in cell clusters. Values are log2 expression levels.

(D) Heatmap of the genes with scLink correlation higher than 0.63 with *Nr1i3*.

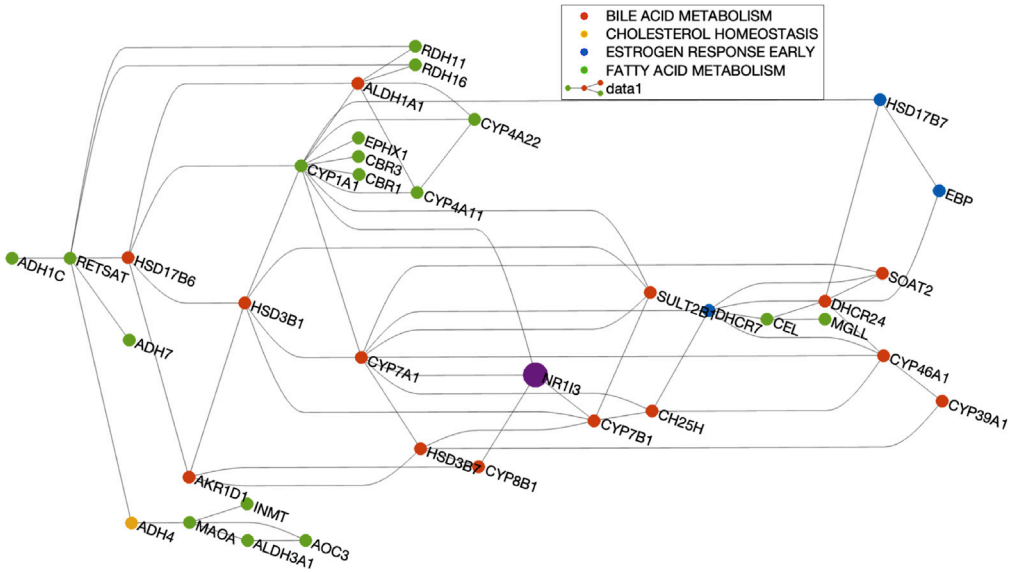


Figure 6. A predictive model independently identifies *CAR* interacting with four NASH-related functional modules

The predicted NASH gene *NR1I3* (large purple dot) directly interacts with three genes from Bile Acid Metabolism Module (*CYP7A1*, *CYP7B1*, and *CYP8B1*), one gene from Fatty Acid Metabolism Module (*CYP1A1*). These four genes further interact with other genes and form a network consisting of fourteen genes from the bile acid metabolism module, three genes from estrogen response early module, seventeen genes from the fatty acid metabolism module, and one gene from the cholesterol homeostasis module. The edges represent the predicted associations between *NR1I3* and these genes.

and their importance in disease pathogenesis providing a resource for identifying new genes for future functional interrogation.

DISCUSSION

Our study uncovers several aspects of liver biology that were previously unknown. First, we find that hepatocytes are highly heterogeneous in their capacity to store lipids and metabolic profiles. By capturing these lipid-laden hepatocytes for single-cell analyses, we find that this heterogeneity is associated with lipid metabolism in a subset of hepatocytes. Elegant previous work analyzing transcriptional changes in hepatocytes isolated from mice fed a 60% high-fat diet showed that hepatic steatosis sensitizes cells to hepatocyte inflammation,[41] suggesting that the stored lipid content is a principal determinant of hepatocyte function. How the lipid heterogeneity of hepatocytes contributes to NAFLD progression will be an area for future exploration.

Second, we find that high Srebp1 expression is not directly associated with higher lipid accumulation, indicating the involvement of other driving factors in a subset of hepatocytes. The finding of a subset of genes in hepatocytes co-expressed with *srebf1* warrants further studies into the mechanisms regulating lipid metabolism *in vivo*.

Third, using experimental and AI predictive models, we identify *CAR* as a gene highly associated with steatohepatitis in mice and humans. *CAR* is abundantly expressed in hepatocytes, but the reported role in nonalcoholic fatty liver disease is controversial.[36,42–44] Here, we find that CAR is localized to the nucleus and overexpressed in patient livers diagnosed with steatohepatitis. Future studies should assess the association and functional correlation between CAR and NAFLD/NASH, as well as explore targeting CAR using specific agonists. Moving forward, it will be important to investigate whether CAR regulates lipogenesis, inflammation, or fibrosis during NASH using relevant models.

Collectively, our results uncover an unexpected heterogeneity in hepatocyte steatosis and identify lipogenic cell populations. In conclusion, this study facilitates the discovery of previously unrelated genes involved in lipid metabolism, which may be used to better understand how to target fatty liver disease.

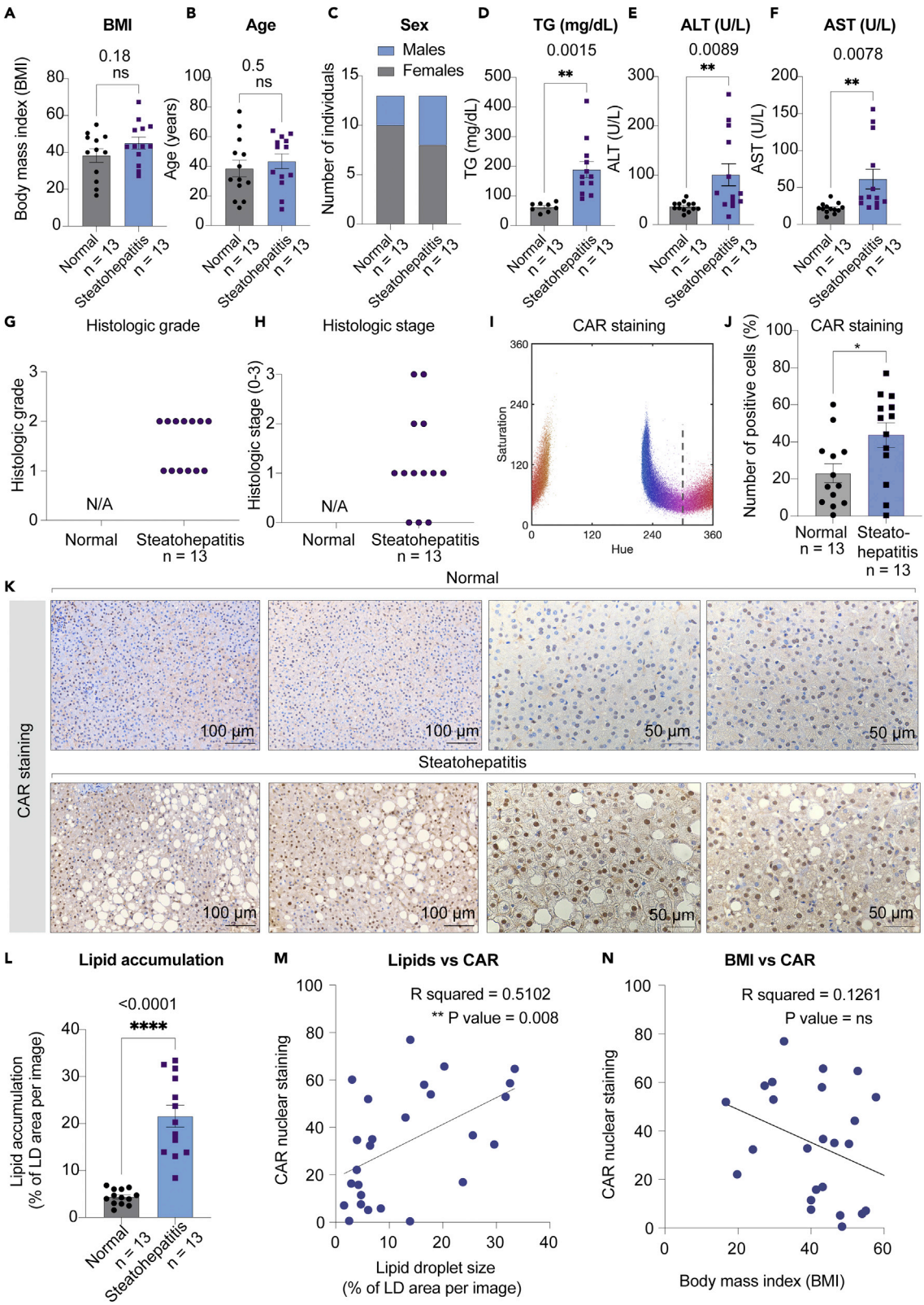

A
BMI
0.18
ns
Body mass index (BMI)
Normal n = 13
Steatohepatitis n = 13
B
Age
0.5
ns
Age (years)
C
Sex
Males
Females
Number of individuals
D
TG (mg/dL)
0.0015
**
TG (mg/dL)
E
ALT (U/L)
0.0089
**
ALT (U/L)
F
AST (U/L)
0.0078
**
AST (U/L)
G
Histologic grade
Histologic grade
N/A
Normal
Steatohepatitis n = 13
H
Histologic stage
Histologic stage (0-3)
N/A
Normal
Steatohepatitis n = 13
I
CAR staining
Saturation
Hue
J
CAR staining
*
Number of positive cells (%)
Normal n = 13
Steato-hepatitis n = 13
K
Normal
CAR staining
100 µm
50 µm
Steatohepatitis
L
Lipid accumulation
<0.0001
****
Lipid accumulation (% of LD area per image)
M
Lipids vs CAR
R squared = 0.5102
** P value = 0.008
CAR nuclear staining
Lipid droplet size (% of LD area per image)
N
BMI vs CAR
R squared = 0.1261
P value = ns
CAR nuclear staining
Body mass index (BMI)

Figure 7. Nuclear CAR localization correlates with steatohepatitis in humans

(A) Body mass index (BMI) of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per group). Data are presented as mean ± SEM p value = ns (non-significant) by two-tailed Student's *t* test.

(B) Age of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per group). Data are presented as mean ± SEM p value = ns (non-significant) by two-tailed Student's *t* test.

(C) Sex of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per group). Data are presented as mean ± SEM p value = ns (non-significant) by two-tailed Student's *t* test.

(D) Triglyceride (TG) levels in mg/dL of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per group). Data are presented as mean ± SEM **p < 0.01 by two-tailed Student's *t* test.

(E) Alanine transferase (ALT) activity in U/L of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per group). Data are presented as mean ± SEM **p < 0.01 by two-tailed Student's *t* test.

(F) Aspartate transaminase (AST) activity (in U/L of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per group). Data are presented as mean ± SEM **p < 0.01 by two-tailed Student's *t* test.

(G) Histological grade of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per group).

(H) Histological stage of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per group).

(I) Segmentation of positive (Red) and negative (blue) cells based on pixel intensity in the nuclei from all images combined to determine the positive threshold for quantification. Each dot represents one cell.

(J) Quantification of the % number of CAR-positive nuclei of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per condition, 4 images per sample). Data are presented as mean ± SEM *p < 0.05 by two-tailed Student's *t* test.

(K) Representative histological images of CAR staining (n = 13 patient samples per condition, 4 images per sample) in normal livers or livers with steatohepatitis. Bar scale: 100 μm or 50 μm.

(L) Quantification of the % of Lipid Droplet area of individuals with normal livers vs. livers with steatohepatitis (n = 13 patient samples per condition, 4 images per sample). Data are presented as mean ± SEM ****p < 0.0001 by two-tailed Student's *t* test.

(M) Correlation plot of CAR levels and lipid droplet size across 26 individuals. **p value <0.01 by two-tailed Student's *t* test.

(N) Correlation plot of CAR levels and BMI across 26 individuals. p value = ns (non-significant by two-tailed Student's *t* test.

Limitations of the study

There are limitations to this work. While our study demonstrates heterogeneous hepatocyte cell state signature in NAFLD, the functions of these distinct hepatocyte populations remain to be determined. In addition, single-cell analyses of hepatocytes should be complemented with single nuclei analyses to confirm the hepatocyte cell states identified in this article.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human samples
  - Animals
- METHOD DETAILS
  - Isolation of cells for single-cell RNA sequencing
  - Library preparation, single-cell RNA sequencing and data preprocessing
  - Cluster identification and expression analysis of single-cell RNA sequencing data
  - AI embedding models for predicting pathogenesis genes
  - Immunohistochemistry
  - CAR quantification
  - Oil red O staining and quantification
  - Immunofluorescence staining and quantification
  - Plasma alanine transaminase (ALT) activity measurement
  - RNA expression analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis of cellular and animal experiments

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105802.

ACKNOWLEDGMENTS

K.J.S. was supported by NIH grants DK125260, DK111916, Merck & Co., Inc. (Rahway, NJ, USA) SEEDS award, the Stanford Diabetes Research Center P30DK116074, the Jacob Churg Foundation, the McCormick and Gabilan Award, the Weintz Family COVID-19 research fund, American Heart Association (AHA), the Stanford School of Medicine, and the Stanford Cardiovascular Institute (CVI). M.Z. was supported by the American Heart Association (AHA) postdoctoral fellowship (905674). L.C. was supported by the Stanford School of Medicine Dean's Postdoctoral Fellowship. S.B.N. was supported by the Novo Nordisk Foundation (grant award NNF20OC0059462) and the Stanford Bio-X Program. reNEW was also supported by the Novo Nordisk Foundation (grant award NNF21CC0073729). RBA and TL were supported by NIH GM102365 and Merck & Co., Inc. (Rahway, NJ, USA) SEEDS award. This work used the Genome Sequencing Service Center by Stanford Center for Genomics and Personalized Medicine Sequencing Center, supported by the grant award NIH S10OD020141, and the Diabetes Genomics and Analysis Core of the Stanford Diabetes Research Center supported by grant award NIH/NIDDK P30DK116074. We thank Pratima Nallagatla and Meng Wang for assistance with bioinformatic analyses and Aila Karioja-Kallio for the collection of human liver samples. We thank the Stanford University Pathology Histology core for histological processing and the Pathology Department for microscopy equipment. The graphical abstract was created with Biorender.com.

AUTHOR CONTRIBUTIONS

Conceptualization, L.C., Y.J. K.J.S.; Methodology, L.C., Y.J. S.B.N., K.J.S.; Validation, Y.J., M.Z., L.C., K.J.S.; Formal Analysis, L.C., Y.J., S.B.N., K.J.S; Investigation, L.C., Y.J., M.Z., T.L., L.C., S.P.; Resources, G.C., R.A., A, K-K, H. Y-J., K.J.S.; Writing – Original Draft, L.C., K.J.S.; Writing – Review & Editing, L.C., M.Z., T.L., R.A., K.J.S.; Funding Acquisition, K.J.S.; Supervision, K.J.S.

DECLARATION OF INTERESTS

The authors do not declare any conflict of interest.

Received: August 29, 2022
Revised: November 15, 2022
Accepted: December 9, 2022
Published: January 20, 2023

REFERENCES

1. Diehl, A.M., and Day, C. (2017). Cause, pathogenesis, and treatment of nonalcoholic steatohepatitis. N. Engl. J. Med. *377*, 2063–2072. https://doi.org/10.1056/nejmra1503519.

2. Sanyal, A.J. (2019). Past, present and future perspectives in nonalcoholic fatty liver disease. Nat. Rev. Gastroenterol. Hepatol. *16*, 377–386. https://doi.org/10.1038/s41575-019-0144-8.

3. Brunt, E.M., Wong, V.W.S., Nobili, V., Day, C.P., Sookoian, S., Maher, J.J., Bugianesi, E., Sirlin, C.B., Neuschwander-Tetri, B.A., and Rinella, M.E. (2015). Nonalcoholic fatty liver disease. Nat. Rev. Dis. Primers *1*, 15080. https://doi.org/10.1038/nrdp.2015.80.

4. Romeo, S., Kozlitina, J., Xing, C., Pertsemlidis, A., Cox, D., Pennacchio, L.A., Boerwinkle, E., Cohen, J.C., and Hobbs, H.H. (2008). Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat. Genet. *40*, 1461–1465. https://doi.org/10.1038/ng.257.

5. Huang, Y., He, S., Li, J.Z., Seo, Y.K., Osborne, T.F., Cohen, J.C., and Hobbs, H.H. (2010). A feed-forward loop amplifies nutritional regulation of PNPLA3. Proc. Natl. Acad. Sci. USA *107*, 7892–7897. https://doi.org/10.1073/pnas.1003585107.

6. Luukkonen, P.K., Nick, A., Hölttä-Vuori, M., Thiele, C., Isokuortti, E., Lallukka-Brück, S., Zhou, Y., Hakkarainen, A., Lundbom, N., Peltonen, M., et al. (2019). Human PNPLA3-I148M variant increases hepatic retention of polyunsaturated fatty acids. JCI Insight *4*, e127902. https://doi.org/10.1172/jci.insight.127902.

7. Abul-Husn, N.S., Cheng, X., Li, A.H., Xin, Y., Schurmann, C., Stevis, P., Liu, Y., Kozlitina, J., Stender, S., Wood, G.C., et al. (2018). A protein-truncating HSD17B13 variant and protection from chronic liver disease. N. Engl. J. Med. *378*, 1096–1106. https://doi.org/10.1056/NEJMoa1712191.

8. Luukkonen, P.K., Tukiainen, T., Juuti, A., Sammalkorpi, H., Haridas, P.A.N., Niemelä, O., Arola, J., Orho-Melander, M., Hakkarainen, A., Kovanen, P.T., et al. (2020). Hydroxysteroid 17-β dehydrogenase 13 variant increases phospholipids and protects against fibrosis in nonalcoholic fatty liver disease. JCI Insight *5*, e132158. https://doi.org/10.1172/jci.insight.132158.

9. Tavares De Almeida, I., Cortez-Pinto, H., Fidalgo, G., Rodrigues, D., and Camilo, M.E. (2002). Plasma total and free fatty acids composition in human non-alcoholic steatohepatitis. Clin. Nutr. *21*, 219–223. https://doi.org/10.1054/clnu.2001.0529.

10. Puri, P., Wiest, M.M., Cheung, O., Mirshahi, F., Sargeant, C., Min, H.K., Contos, M.J., Sterling, R.K., Fuchs, M., Zhou, H., et al. (2009). The plasma lipidomic signature of nonalcoholic steatohepatitis. Hepatology *50*, 1827–1838. https://doi.org/10.1002/hep.23229.

11. Donnelly, K.L., Smith, C.I., Schwarzenberg, S.J., Jessurun, J., Boldt, M.D., and Parks, E.J. (2005). Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. J. Clin. Invest. *115*, 1343–1351. https://doi.org/10.1172/JCI23621.

12. Abdelmalek, M.F., Suzuki, A., Guy, C., Unalp-Arida, A., Colvin, R., Johnson, R.J., and Diehl, A.M.; Nonalcoholic Steatohepatitis Clinical Research Network (2010). Increased fructose consumption is associated with fibrosis severity in patients with nonalcoholic fatty liver disease. Hepatology *51*, 1961–1971. https://doi.org/10.1002/hep.23535.

13. Stanhope, K.L., Schwarz, J.M., Keim, N.L., Griffen, S.C., Bremer, A.A., Graham, J.L., Hatcher, B., Cox, C.L., Dyachenko, A., Zhang, W., et al. (2009). Consuming fructose-sweetened, not glucose-sweetened, beverages increases visceral adiposity and lipids and decreases insulin sensitivity in overweight/obese humans. J. Clin. Invest. *119*, 1322–1334. https://doi.org/10.1172/JCI37385.

14. Thuy, S., Ladurner, R., Volynets, V., Wagner, S., Strahl, S., Königsrainer, A., Maier, K.-P., Bischoff, S.C., and Bergheim, I. (2008). Nonalcoholic fatty liver disease in humans is associated with increased plasma endotoxin and plasminogen activator inhibitor 1 concentrations and with fructose intake. J. Nutr. *138*, 1452–1455. https://doi.org/10.1093/jn/138.8.1452.

15. Gaggini, M., Morelli, M., Buzzigoli, E., DeFronzo, R.A., Bugianesi, E., and Gastaldelli, A. (2013). Non-alcoholic fatty liver disease (NAFLD) and its connection with insulin resistance, dyslipidemia, atherosclerosis and coronary heart disease. Nutrients *5*, 1544–1560. https://doi.org/10.3390/nu5051544.

16. Luyckx, F.H., Lefebvre, P.J., and Scheen, A.J. (2000). Non-alcoholic steatohepatitis: association with obesity and insulin resistance, and influence of weight loss. Diabetes Metabol. *26*, 98–106.

17. Herms, A., Bosch, M., Ariotti, N., Reddy, B.J.N., Fajardo, A., Fernández-Vidal, A., Alvarez-Guaita, A., Fernández-Rojo, M.A., Rentero, C., Tebar, F., et al. (2013). Cell-to-cell heterogeneity in lipid droplets suggests a mechanism to reduce lipotoxicity. Curr. Biol. *23*, 1489–1496. https://doi.org/10.1016/j.cub.2013.06.032.

18. Softic, S., Gupta, M.K., Wang, G.X., Fujisaka, S., O'Neill, B.T., Rao, T.N., Willoughby, J., Harbison, C., Fitzgerald, K., Ilkayeva, O., et al. (2017). Divergent effects of glucose and fructose on hepatic lipogenesis and insulin signaling. J. Clin. Invest. *127*, 4059–4074. https://doi.org/10.1172/JCI94585.

19. Maxwell, K.N., Soccio, R.E., Duncan, E.M., Sehayek, E., and Breslow, J.L. (2003). Novel putative SREBP and LXR target genes identified by microarray analysis in liver of cholesterol-fed mice. J. Lipid Res. *44*, 2109–2119. https://doi.org/10.1194/jlr.M300203-JLR200.

20. Moon, Y.-A., Liang, G., Xie, X., Frank-Kamenetsky, M., Fitzgerald, K., Koteliansky, V., Brown, M.S., Goldstein, J.L., and Horton, J.D. (2012). The Scap/SREBP pathway is essential for developing diabetic fatty liver and carbohydrate-induced hypertriglyceridemia in animals. Cell Metab. *15*, 240–246. https://doi.org/10.1016/j.cmet.2011.12.017.

21. Kim, J.B., and Spiegelman, B.M. (1996). ADD1/SREBP1 promotes adipocyte differentiation and gene expression linked to fatty acid metabolism. Genes Dev. *10*, 1096–1107. https://doi.org/10.1101/gad.10.9.1096.

22. Stern, S., Kurian, R., and Wang, H. (2022). Clinical relevance of the constitutive androstane receptor. Drug Metab. Dispos. *50*, 1010–1018. https://doi.org/10.1124/DMD.121.000483.

23. Xiong, X., Kuang, H., Ansari, S., Liu, T., Gong, J., Wang, S., Zhao, X.Y., Ji, Y., Li, C., Guo, L., et al. (2019). Landscape of intercellular crosstalk in healthy and NASH liver revealed by single-cell secretome gene analysis. Mol. Cell *75*, 644–660.e5. https://doi.org/10.1016/j.molcel.2019.07.028.

24. MacParland, S.A., Liu, J.C., Ma, X.Z., Innes, B.T., Bartczak, A.M., Gage, B.K., Manuel, J., Khuu, N., Echeverri, J., Linares, I., et al. (2018). Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat. Commun. *9*, 4383. https://doi.org/10.1038/s41467-018-06318-7.

25. Krahmer, N., Najafi, B., Schueder, F., Quagliarini, F., Steger, M., Seitz, S., Kasper, R., Salinas, F., Cox, J., Uhlenhaut, N.H., et al. (2018). Organellar proteomics and phospho-proteomics reveal subcellular reorganization in diet-induced hepatic steatosis. Dev. Cell *47*, 205–221.e7. https://doi.org/10.1016/j.devcel.2018.09.017.

26. Boland, M.L., Oró, D., Tølbøl, K.S., Thrane, S.T., Nielsen, J.C., Cohen, T.S., Tabor, D.E., Fernandes, F., Tovchigrechko, A., Veidal, S.S., et al. (2019). Towards a standard diet-induced and biopsy-confirmed mouse model of non-alcoholic steatohepatitis: impact of dietary fat source. World J. Gastroenterol. *25*, 4904–4920. https://doi.org/10.3748/wjg.v25.i33.4904.

27. Jiang, Z., Zhao, M., Voilquin, L., Jung, Y., Aikio, M.A., Sahai, T., Dou, F.Y., Roche, A.M., Carcamo-Orive, I., Knowles, J.W., et al. (2021). Isthmin-1 is an adipokine that promotes glucose uptake and improves glucose tolerance and hepatic steatosis. Cell Metab. *33*, 1836–1852.e11. https://doi.org/10.1016/j.cmet.2021.07.010.

28. Jung, Y., Zhao, M., and Svensson, K.J. (2020). Isolation, culture, and functional analysis of hepatocytes from mice with fatty liver disease. STAR Protoc. *1*, 100222. https://doi.org/10.1016/j.xpro.2020.100222.

29. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420. https://doi.org/10.1038/nbt.4096.

30. Halpern, K.B., Shenhav, R., Matcovitch-Natan, O., Tóth, B., Lemze, D., Golan, M., Massasa, E.E., Baydatch, S., Landen, S., Moor, A.E., et al. (2017). Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature *542*, 352–356. https://doi.org/10.1038/nature21065.

31. Ben-Moshe, S., Shapira, Y., Moor, A.E., Manco, R., Veg, T., Bahar Halpern, K., and Itzkovitz, S. (2019). Spatial sorting enables comprehensive characterization of liver zonation. Nat. Metab. *1*, 899–911. https://doi.org/10.1038/s42255-019-0109-9.

32. Pettinelli, P., Arendt, B.M., Teterina, A., McGilvray, I., Comelli, E.M., Fung, S.K., Fischer, S.E., and Allard, J.P. (2018). Altered hepatic genes related to retinol metabolism and plasma retinol in patients with non-alcoholic fatty liver disease. PLoS One *13*, e0205747. https://doi.org/10.1371/journal.pone.0205747.

33. Ma, R., Martínez-Ramírez, A.S., Borders, T.L., Gao, F., and Sosa-Pineda, B. (2020). Metabolic and non-metabolic liver zonation is established non-synchronously and requires sinusoidal Wnts. Elife *9*, e46206. https://doi.org/10.7554/eLife.46206.

34. Gosis, B.S., Wada, S., Thorsheim, C., Li, K., Jung, S., Rhoades, J.H., Yang, Y., Brandimarto, J., Li, L., Uehara, K., et al. (2022). Inhibition of nonalcoholic fatty liver disease in mice by selective inhibition of mTORC1. Science (New York, N.Y.) *376*, eabf8271. https://doi.org/10.1126/science.abf8271.

35. Softic, S., Cohen, D.E., and Kahn, C.R. (2016). Role of dietary fructose and hepatic de novo lipogenesis in fatty liver disease. Dig. Dis. Sci. *61*, 1282–1293. https://doi.org/10.1007/s10620-016-4054-0.

36. Maglich, J.M., Lobe, D.C., and Moore, J.T. (2009). The nuclear receptor CAR (NR1I3) regulates serum triglyceride levels under conditions of metabolic stress. J. Lipid Res. *50*, 439–445. https://doi.org/10.1194/jlr.M800226-JLR200.

37. Yamazaki, Y., Kakizaki, S., Horiguchi, N., Sohara, N., Sato, K., Takagi, H., Mori, M., and Negishi, M. (2007). The role of the nuclear receptor constitutive androstane receptor in the pathogenesis of non-alcoholic steatohepatitis. Gut *56*, 565–574. https://doi.org/10.1136/gut.2006.093260.

38. Xiong, X., Wang, Q., Wang, S., Zhang, J., Liu, T., Guo, L., Yu, Y., and Lin, J.D. (2019). Mapping the molecular signatures of diet-induced NASH and its regulation by the hepatokine Tsukushi. Mol. Metab. *20*, 128–137. https://doi.org/10.1016/j.molmet.2018.12.004.

39. Piñero, J., Ramírez-Anguita, J.M., Saüch-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., and Furlong, L.I. (2020). The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Res. *48*, D845–D855. https://doi.org/10.1093/NAR/GKZ1021.

40. Petryszak, R., Keays, M., Tang, Y.A., Fonseca, N.A., Barrera, E., Burdett, T., Füllgrabe, A., Fuentes, A.M.P., Jupp, S., Koskinen, S., et al. (2016). Expression Atlas update - an integrated database of gene and protein expression in humans, animals and plants. Nucleic Acids Res. *44*, D746–D752. https://doi.org/10.1093/nar/gkv1045.

41. Sheng, L., Jiang, B., and Rui, L. (2013). Intracellular lipid content is a key intrinsic determinant for hepatocyte viability and metabolic and inflammatory states in mice. Am. J. Physiol. Endocrinol. Metab. *305*, E1115–E1123. https://doi.org/10.1152/ajpendo.00401.2013.

42. di Masi, A., De Marinis, E., Ascenzi, P., and Marino, M. (2009). Nuclear receptors CAR and PXR: molecular, functional, and

biomedical aspects. Mol. Aspects Med. *30*, 297–343. https://doi.org/10.1016/j.mam.2009.04.002.

43. Dong, B., Saha, P.K., Huang, W., Chen, W., Abu-Elheiga, L.A., Wakil, S.J., Stevens, R.D., Ilkayeva, O., Newgard, C.B., Chan, L., and Moore, D.D. (2009). Activation of nuclear receptor CAR ameliorates diabetes and fatty liver disease. Proc. Natl. Acad. Sci. USA *106*, 18831–18836. https://doi.org/10.1073/pnas.0909731106.

44. Lynch, C., Pan, Y., Li, L., Heyward, S., Moeller, T., Swaan, P.W., and Wang, H. (2014). Activation of the constitutive androstane receptor inhibits gluconeogenesis without affecting lipogenesis or fatty acid synthesis in human hepatocytes. Toxicol. Appl. Pharmacol. *279*, 33–42. https://doi.org/10.1016/J.TAAP.2014.05.009.

45. Weisenfeld, N.I., Kumar, V., Shah, P., Church, D.M., and Jaffe, D.B. (2017). Direct determination of diploid genome sequences. Genome Res. *27*, 757–767. https://doi.org/10.1101/gr.214874.116.

46. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

47. Franzén, O., Gan, L.-M., and Björkegren, J.L.M. (2019). PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database *2019*, baz046. https://doi.org/10.1093/database/baz046.

48. Suo, S., Zhu, Q., Saadatpour, A., Fei, L., Guo, G., and Yuan, G.-C. (2018). Revealing the critical regulators of cell identity in the mouse cell Atlas. Cell Rep. *25*, 1436–1445.e3. https://doi.org/10.1016/j.celrep.2018.10.045.

49. Zhang, X., Lan, Y., Xu, J., Quan, F., Zhao, E., Deng, C., Luo, T., Xu, L., Liao, G., Yan, M., et al. (2019). CellMarker: a manually curated resource of cell markers in human and mouse. Nucleic Acids Res. *47*, D721–D728. https://doi.org/10.1093/nar/gky900.

50. Liu, T., Wang, S., Wornow, M., and Altman, R.B. (2022). Construction of disease-specific cytokine profiles by associating disease genes with immune responses. PLoS Comput. Biol. *18*, e1009497. https://doi.org/10.1371/JOURNAL.PCBI.1009497.

51. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. *47*, D607–D613. https://doi.org/10.1093/NAR/GKY1131.

52. Grover, A., and Leskovec, J. (2016). node2vec: scalable feature learning for networks. KDD *2016*, 855–864. proceedings. International Conference on Knowledge Discovery & Data Mining. https://doi.org/10.1145/2939672.2939754.

53. Rieckmann, J.C., Geiger, R., Hornburg, D., Wolf, T., Kveler, K., Jarrossay, D., Sallusto, F., Shen-Orr, S.S., Lanzavecchia, A., Mann, M., and Meissner, F. (2017). Social network architecture of human immune cells unveiled by quantitative proteomics. Nat. Immunol. *18*, 583–593. https://doi.org/10.1038/ni.3693.

54. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. *1*, 417–425. https://doi.org/10.1016/J.CELS.2015.12.004.

55. Wishart, D.S., Tzur, D., Knox, C., Eisner, R., Guo, A.C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., et al. (2007). HMDB: the human metabolome database. Nucleic Acids Res. *35*, D521–D526. https://doi.org/10.1093/NAR/GKL923.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat anti-Mouse IgG (H+L) Cross-absorbed Secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	Cat# A11001; RRID:AB_2534069
Mouse monoclonal anti-SREBP1 (2A4)	Santa Cruz Biotechnology	Cat# sc-13551; RRID:AB_628282
Mouse monoclonal anti-NR1I3	Thermo Fisher Scientific	Cat# CF805306
Bacterial and virus strains		
One Shot Top10 Chemically Competent *E. coli*	ThermoFisher	Cat# C404003
Chemicals, peptides, and recombinant proteins		
HBSS buffer	Gibco	Cat# 14175-095
Potassium chloride	Sigma	Cat# 112-033-101
D-(+)-Glucose	Sigma	Cat# G8270
Sodium bicarbonate	Sigma	Cat# S6297
UltraPure™ 0.5M EDTA, pH8.0	Invitrogen	Cat# 15575-038
William E media	Quality Biological	Cat# 112-033-101
Corning™ Regular Fetal Bovine Serum	Corning	Cat# 35-010-CV
Sodium pyruvate	ThermoFisher	Cat# 11360070
Dexamethasone	Sigma	Cat# D1756
HEPES buffered saline	Sigma-Aldrich	Cat# 51558
DMEM high glucose	Sigma	Cat# D6429
Insulin, Human Recombinant	Millipore Sigma	Cat# 91077C
Penicillin/Streptomycin	Gibco	Cat# 15140-122
Collagenase IV	Sigma	Cat# C5138
PBS	Gibco	Cat# 10010-023
10x PBS	Gibco	Cat# 70011044
Percoll	Sigma	Cat# P1644
Ethanol	FisherScientific	Cat# 22-032-601
BSA	Sigma	Cat# A7906
Glutamax	ThermoFisher	Cat# 35050061
Isopropanol	FisherChemical	Cat# BP2632-4
Oil Red O solution	Sigma	Cat# O1391
Formalin	SIP Brand	Cat# C4320
Rat tail Collagen I	Corning	Cat# 354236
Trypan Blue Stain (0.4%)	Invitrogen	Cat# T10282
High-Capacity cDNA Reverse Transcription Kit	ThermoFisher	Cat# 4368814
2X SYBR GREEN qPCR MASTER MIX	Bimake	Cat# B21203
Trizol	ThermoFisher	Cat# 15-596-026
Trichrome Stain (Masson) Kit	Sigma	Cat# HT15
LipidSpot™610	Biotium	Cat# 70069
ImmPACT DAB substrate	Vector laboratories	Cat# SK-4105
Hematoxylin solution	Sigma	Cat# MHS32

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
Alanine Transminase Colorimetric Activity Assay Kit	Cayman	Cat# 700260
Countess II FL Automated Cell Counter	Life Technology	Cat# AMQAF1000
ImmPRESS® HRP Horse Anti-Mouse IgG Polymer Detection Kit	Vector laboratories	Cat# MP-7402
Deposited data		
Raw and analyzed sequencing data	This paper	Gene Expression Omnibus GSE210501
Experimental models: Cell lines		
Primary hepatocytes from C57BL/6J *Mus musculus*	This paper	This paper
Experimental models: Organisms/strains		
C57BL/6J *Mus musculus*	The Jackson Laboratory	Cat# 000664; RRID: IMSR_JAX:000664
Oligonucleotides		
See Table S3		
Software and algorithms		
SPOT Advanced	SPOT IMAGING™	SPOT56A RRID:SCR_016613
10x Genomics Cell Ranger	10x Genomics	Version 3.1.0 RRID:SCR_017344
10x Genomics Loupe Browser	10x Genomics	Version 3.3.1 RRID:SCR_018555
R Core	https://www.r-project.org/	Version 4.0.2 RRID:SCR_001905
R Studio	https://rstudio.com/	Version 1.3
Seurat	Butler et al., Nature Biotechnology 2018	Version 3.2.0 RRID:SCR_007322
ImageJ	N/A	ImageJ RRID:SCR_003070
GraphPad Prism	GraphPad Prism	RRID:SCR_002798
DisGeNET	https://www.disgenet.org/ Piñero et al., 2020	DisGeNET v7.0 RRID:SCR_006178
Node2vec	Grover and Leskovec, 2016	
HMDB	https://hmdb.ca/ Wishart et al., 2007	Version 5.0 RRID:SCR_007712
MSigDB	Liberzon et al., 2015	Molecular Signatures Database v7.5.1 RRID:SCR_016863
ImmProt	http://www.immprot.org/ Rieckmann et al., 2017	
STRING	https://string-db.org/ Szklarczyk et al., 2019	STRING V11.5 RRID:SCR_005223
Other		
Blood Glucose meter	OneTouch UltraMini meter	N/A
Blood Glucose Strips	GenUltimate	Cat# 100-50
Rodent NASH diet	Research Diets	Cat# D09100310

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
70um cell strainer	BD Falcon	Cat# 352350
25G x 7/8 Needle	BD	Cat# 305124
Petri dish Fisherbrand™	FisherScientific	Cat# FB0875713
Fisherbrand™ Variable-Flow Peristaltic Pump	FisherScientific	Cat# 13-876-2
Perfusion tubing connectors 3/32″ Polypropylene Coupler	United States plastic Corp	Cat# 65600
Perfusion tubing 1/32″ ID x 3/32″ OD Silicone	United States plastic Corp	Cat# 57286
Corning Costar Flat bottomed cell culture plates	Fisher Scientific	Cat# 07-200-82
Countess II FL Automated Cell Counter	Life Technology	Cat# AMQAF1000
Millex-GS Syringe Filter Unit, 0.22 µm	Millipore Sigma	Cat# SLGSV255F
BD Syringes without Needle, 50 mL	Fisher Scientific	Cat# BD309653
LEICA TCS SP8 X Confocal microscope	Leica	N/A
QuantStudio™5 – 384 – Well Block qpcr machine	Thermo Fisher	Cat# 272530639

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Katrin J. Svensson (katrinjs@stanford.edu).

Materials availability

All reagents used in this study are commercially available.

Data and code availability

- All raw and processed single-cell RNA sequencing data have been deposited to Gene Expression Omnibus (GSE210501).
- The script for immunohistochemistry quantification is available at https://github.com/Svensson-Lab/Coassolo2022.
- All information required to reanalyze the data is reported in this paper.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human samples

The human liver samples used for histological characterization of CAR levels were obtained at Stanford University under IRB protocol #58373 using excess/archival material. The subjects were identified by searching the pathology archive database for liver disease diagnoses. Exclusion criteria were liver cancers and tumors as identified by information stored in the pathology database. Clinical and laboratory data were obtained by retrieving existing data in the electronic medical record. The human liver cDNA samples were obtained from de-identified NAFLD/non-NAFLD patients from Hannele Yki-Jarvinen, and approved by the Ethics Committee in Helsinki, Finland. The human single-cell RNA sequencing data was re-analyzed from a published dataset by Macparland et al.[24]

Animals

Animal experiments were performed per procedures approved by the Institutional Animal Care and Use Committee of the Stanford Animal Care and Use Committee (APLAC) protocol #32982. C57BL/6J mice were purchased from the Jackson Laboratory (#000664). Unless otherwise stated, mice were housed in a temperature-controlled (20–22°C) room on a 12-hour light/dark cycle. All experiments were performed with age-matched male mice housed in groups of five unless stated otherwise. For single-cell RNA sequencing experiments, 12-week-old male mice were fed with either chow diet (Envigo, #2018) or NAFLD diet (40% fat, 20% kcal fructose, and 2% cholesterol, #D09100310 ResearchDiets) for 6 weeks and sacrificed at 18 weeks of age.

METHOD DETAILS

Isolation of cells for single-cell RNA sequencing

All cells were cultured in a humidified atmosphere containing 5% CO2 at 37°C. Primary hepatocyte isolation was performed using gradient separation.[28] 12-week-old male C57BL/6J mice fed with either chow diet or NAFLD diet for 6 weeks were sacrificed. The inferior vena cava was cannulated with a 25G needle connected to tubing for perfusion with 20 ml of 37°C pre-heated pH 7.4 HBSS buffer (#14175-095, Gibco) containing 5.4 mM KCl, 30 mM Sodium bicarbonate and 0.285 mM EDTA at a rate of 1 ml per minute. After 2 minutes of perfusion, the hepatic portal vein was cut. After perfusion, 20 ml of liver digestion medium containing 1 mg/ml collagenase type IV ((#C5138, Sigma), 10% FBS and 1 mM HEPES was added to the perfusion buffer tube. After 7–10 minutes, the liver was dissected and transferred to a petri dish and mechanically dissociated by gently swirling the tissue in 10 ml Williams E medium containing Glutamax (#112-033-101, Quality Biological), 10% FBS, 2 mM sodium pyruvate, 1uM dexamethasone and 0.1uM insulin. The dissociated liver cells were filtered through a 70 μm cell strainer and centrifuged at 50 x*g* for 4 minutes to separate hepatocytes and non-parenchymal cells (NPC). For isolation of hepatocytes, cell pellets were resuspended in plating medium and mixed with 90% or 25% Percoll (#P1644, Sigma) followed by centrifugation at 100 x*g* for 10 minutes and 50 x*g* for 3 minutes. For isolation of NPCs, cell suspensions were centrifuged at 100 x*g* for 5 minutes to reduce the number of erythrocytes. NPCs were isolated and washed as follows: 300 x*g* for 7 minutes, 650 x*g* for 4 minutes, 240 x*g* for 5 minutes, and 650 x*g* for 4 minutes. Pellets were then combined with hepatocytes. Cells were combined as one sample (chow and NAFLD, n = 5 mice/group) and processed for single-cell RNA sequencing library synthesis preparation after confirming cell numbers and viability. For culture of hepatocytes and non-parenchymal cells, cell pellets were resuspended in plating medium. 4 hours after seeding on collagen-coated plates, cells were washed with PBS, followed by the addition of maintenance medium Williams E supplemented with 0.2% BSA, 2 mM sodium pyruvate, 0.1 μM dexamethasone.

Library preparation, single-cell RNA sequencing and data preprocessing

All samples were sequenced in the same round to avoid the need for batch correction. The libraries were prepared using the 10X Genomics 3′ version 3 single cell gene expression kit. Then they were sequenced on the Illumina HiSeq 4000 with a 2x101 bp reads. Dual indexed libraries of isolated mouse liver cells were then pooled and sequenced on an Illumina HiSeq 4000 sequencer at the Stanford Genome Sequencing Service Center in a 100-bp paired-end configuration. The read structure was dual indexed sequencing run with Read 1 starting from a Read 1 being 28 bases including cell barcode and unique molecular identifier (UMI), index i7 of 10 bases, index i5 of 10 bases and Read 2 being 90 bases containing transcript information. The libraries were processed and decomplexed using the pipeline from 10x Genomics Cell Ranger 3.1.0.[45]

Cluster identification and expression analysis of single-cell RNA sequencing data

The data was analyzed using Seurat[29] in R Studio. Raw count tables were loaded into Seurat for both datasets (chow and NAFLD) and the Seurat "merge" function was applied to perform pooled analysis. After removing cells that were either less than 200 genes or more than 50,000 genes and filtering out over 5% mitochondrial content, gene expression was normalized by global-scaling normalization method, "LogNormalize", merged and clustered following the standard Seurat package procedures.[29,46] The combined dataset of chow and NAFLD mice liver identified populations consistent with prior reports.[47–49] The cluster definition heatmap illustrates three representative markers of gene expressions shown in log2 value for each cluster to define the population. For comparisons between mouse and human sc-RNA sequencing data, the Rshiny app provided by the Bader lab was used to generate and analyze the hepatocyte clusters. In Bader et al., clusters 1, 3, 4, 6, 14, and 15 were identified as hepatocytes. The following analyses were performed: (1) scatter plot for all mouse and human genes (12,398 genes found in common between 13,519 genes in the mouse and 20,007 genes in the human datasets) expressed as average gene expression across all samples in the human dataset (8444 cells) vs. mouse (5932 cells, NAFLD and chow), (2) venn diagram of overlap in mouse and human genes, and (3) extraction of the clustering information for the human dataset to determine if genes were heterogeneously expressed in the human hepatocyte clusters. P values were adjusted using Bonferroni correction for multiple testing.

AI embedding models for predicting pathogenesis genes

To identify disease associated genes, we have built predictive models based on the proximity of the genes to functional modules in an embedded map of PPIs.[50] We applied node2vec to capture network topology

features of the 14,707 genes (from STRING-2019) in a 64-dimensional embedding space.[51] Node2vec is a graph algorithm that uses the local connectivity around a gene to summarize its interactions in a low dimensional space;[52] neighboring nodes have similar interactions and are "close" in the embedding space, as measured by cosine distance. We then estimated a gene's role in pathogenesis by computing its embedding distance from a set of 220 functional modules (47 immune response modules from ImmProt,[53] 50 signaling pathway modules from MsigDB,[54] 123 metabolic modules from Human Metabolic Reaction Database[55] to those of known disease genes. Methods for constructing the NASH predictive model are described in Figure S6. Given a module consisting of a set of genes, we summed the individual gene embedding vectors to generate a summary vector for the functional module. Cosine similarity measures proximity between gene embeddings and module vectors. For a given gene, we can calculate its similarity to each of the above functional modules; the vector of values describes the genes functional relationships. Given a gold standard set of genes involved in disease, we can predict additional disease genes using machine learning on these embedded representations of functional relationships. As a proof of concept, we collected 70 NASH genes from DisGeNET[39] as a gold-standard list. We identified the functional modules that are closest to these 70 NASH genes. We used a Linear SVM to create a classifier for defining additional genes with a similar profile as the known 70. A held out set of randomly chosen genes were used as negative examples. We performed 10-fold cross validation and systematically varied the parameters of the model (Figure S6). Using large-scale gene embeddings,[50] module embeddings were calculated by summing the individual gene embedding vectors within a given functional module to generate a summary vector.

Immunohistochemistry

Immunohistochemistry on mouse livers was performed on OCT-embedded 6 μm thick frozen sections. Tissue sections were fixed in 3% formalin in PBS for 1h at 20°C. For hematoxylin and eosin (H&E) staining, slides were stained with hematoxylin, washed with water and 95% ethanol, and stained with eosin for 30 min. Sections were then dehydrated with ethanol and xylene and mounted with mounting medium. Trichrome staining was performed according to the manufacturer's instructions using Trichrome Stain Kit (Sigma, #HT15). For CAR quantification, immunostaining was performed on paraffin-embedded human liver tissues. In brief, the paraffin blocks were sliced into 5-μM thick sections, deparaffinized with xylene and rehydrated with decreasing concentrations of ethanol in water. Antigen retrieval was achieved by incubating slides in sodium citrate buffer (pH 6.0) for 25 min at 95°C, followed by 20 min of cooling at room temperature. Endogenous peroxidases were quenched by incubating the slides in 3% hydrogen peroxide for 10 min. The sections were then washed with phosphate-buffered saline (PBS) for 5 min. Endogenous avidin and biotin were blocked using a blocking kit according to manufacturer instructions (Vector Laboratories, Burlington, ON, Canada). Primary antibody (1/100, #CF805306, ThermoFisher) was applied for 2h at room temperature in a humidified chamber. After rinsing the slides in PBS, they were incubated in polymer-HRP secondary antibody (Vector Laboratories) for 30 min at room temperature. After washing with PBS for 5 minutes, the slides were incubated with Vectastain ABC reagent (Vector Laboratories) for 30 min. After washing with PBS for five minutes, color development was achieved by applying diaminobenzidine tetrahydrochloride (DAB) solution (Vector Laboratories) for two minutes. After washing in distilled water, the sections were counterstained with hematoxylin and blue in ammonia water, dehydrated through ethanol and xylene, and cover-slipped using a xylene-based mounting medium. All slides were reviewed at the same time with a Nikon eclipse e1000 upright light microscope and images were captured with a Nikon Digital Sight DS-Fi1 color camera and Spot Advanced software.

CAR quantification

Segmentation was done independently for each image by converting an RGB image into grayscale and stretching the converted pixel intensities such that the top 1% and the bottom 1% of the pixels were saturated. A threshold was introduced at 40% of the maximum pixel intensity. Any pixel with a grayscale intensity below that threshold was considered part of a cell nucleus. To improve the segmentation further, a flood-filled operation and a 5-pixel disk-shaped morphological opening were applied. The average RGB pixel color inside each segmented nuclei was calculated, which was considered the most likely color of that cell nuclei. To classify nuclei into positive and negative cells, the RGB pixel color was converted into HSV space. Any cell nuclei with a hue inside the interval [120; 300] was considered negative; otherwise, the cell was considered positive. The script is available at https://github.com/Svensson-Lab/Coassolo2022.

Oil red O staining and quantification

For Oil red O staining on slides or in cells, samples were fixed with 3% formalin in PBS, washed twice with water, incubated with 60% isopropanol for 5 min and then incubated with premixed Oil Red O solution (Sigma, #O1395) at the ratio of 3:2 of Oil Red O:H_2O for 20 min followed by washing with H_2O. For quantification of Oil Red O staining in cells, Oil Red O staining was extracted by adding 0.5 ml isopropanol to stained cells in a 24-well plate and transferred to a 96-well clear flat-bottomed plate. Absorbance was read at 540 nm. Values are expressed as absorbance normalized by cell number.

Immunofluorescence staining and quantification

For Srebp1 (Santa Cruz, sc-13551) and LipidSpot dye (Biotium, 70069, 1:1000 dilution in PBS) co-staining, the staining was performed sequentially. The Srebp1 antibody (1:200 dilution in PBS/1% BSA) was applied for 24 h at 4°C, followed by a secondary fluorescent antibody (Thermo Fisher Scientific, A11001, 1:500 dilution in PBS/1% BSA) for 1h. LipidSpot dye was then applied for 2h on the same liver tissue, according to the manufacturer's protocol. The fluorescent images of Srebp1 and LipidSpot were taken using a Leica SP8 confocal microscope using a 63x objective. Quantification of nuclear-localized Srebp1 was performed using immunofluorescence images with the ImageJ software. Ten images were analyzed in each group. For quantification of Srebp1 high and low cells, the number of nucleus per image were quantified using the DAPI staining. The intensity of Srebp1 staining was measured in each nucleus and the average pixel intensity of Srebp1 staining was used as a threshold to determine positive or negative expression. The cut-off value was set to 108,5 as defined as the average pixel intensity of Srebp1 staining (Figure 4B). Srebp1 expression was interpreted as positive when expressed above the 50th percentile. The values are expressed as the percentage of $Srebp1^{high}$ and $Srebp1^{low}$ cells. The statistical analysis was performed by comparisons between Chow $Srebp1^{high}$ and NAFLD $Srebp1^{high}$ using two tailed Student's t-test and a P-value of < 0.05 was considered significant. For quantification of Srebp1 expression within lipid-positive or lipid-negative areas, 10 representative pictures (90 μm × 90 μm) of lipid positive and negative areas were selected from the 6 weeks of NAFLD diet mice. Lipid positive areas were determined by quantifying the Lipid Spot staining. The co-localized area (μm^2) between Srebp1 and DAPI staining was quantified in each area. The values are expressed as percentage of Srebp1 positive nuclei. In total, 10 lipid-positive and lipid-negative areas were quantified. The statistical analysis was performed by comparisons between lipid positive and negative areas using two tailed Student's t-test and a P-value of < 0.05 was considered significant.

Plasma alanine transaminase (ALT) activity measurement

Animals fed with either chow or NAFLD diet were sacrificed, and blood was drawn by cardiac puncture and collected in heparinized tubes. Blood was centrifuged at 6000 x*g* for 10 minutes at 4°C to collect plasma and the supernatant was transferred to a new tube and kept at −80°C. ALT activity measurements and analyses were performed according to the manufacturer's protocol using Alanine Transaminase Colorimetric Activity Assay Kit (Cayman, #700260). Briefly, 150 μl of substrates, 20 μl of cofactor and 20 μl of 2x diluted plasma samples were loaded onto 96 well plates in duplicates including positive controls that were provided in the kit. The plates were incubated for 15 minutes at 37°C. After incubation, the reactions were initiated by adding 20 μl of ALT initiator followed by immediate measurement of the absorbance at 340 nm every 10 minutes at 37°C. The absolute difference in absorbance value between two time points was divided by the time point difference, which then was multiplied by the extinction coefficient of 0.21/(4.11 x 0.02) and the dilution factor. Values are expressed as units (U)/L.

RNA expression analysis

Total RNA from cultured cells or tissues was isolated using TRIzol (#15596018 Thermo Fisher Scientific) and Rneasy mini kits (# 74104 Qiagen). RNA was reverse transcribed using the ABI high-capacity cDNA synthesis kit. For qRT-PCR analysis, cDNA, primers and SYBR-green fluorescent dye (ABI) were used. Relative mRNA expression was determined by normalization with *cyclophilin* levels using the ΔΔCt method. The primer sequences used are listed in Table S3.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis of cellular and animal experiments

All values in graphs are presented as mean ± SEM P-values represent ****: $p < 0.001$, ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$, non-significant: $p > 0.1$. Two-way ANOVA was used for repeated measures (*$p < 0.05$, **$p < 0.01$,

***$p < 0.001$). Student's t test was used for single comparisons. Values for n represent biological replicates for cultured cell experiments or individual animals for *in vivo* experiments. Specific details for n values are noted in each figure legend. For cellular assays, n corresponds to the number of experimental replicates using cells isolated from individual mice. Each animal experiment was repeated using at least two cohorts of mice. For animal experiments, n corresponds to the number of animals per condition. Mice were randomly assigned to treatment groups for *in vivo* studies.