Chronic metabolic stress drives developmental programs and loss of tissue functions in non-transformed liver that mirror tumor states and stratify survival

Constantine N. Tzouanas[1-6,24], Marc S. Sherman[7-8,24], Jessica E.S. Shay[5,9-11,24], Adam J. Rubin[3-6], Benjamin E. Mead[1,3-6], Tyler T. Dao[1,3-6,12], Titus Butzlaff[7], Miyeko D. Mana[5,13-14], Kellie E. Kolb[1,3-6], Chad Walesky[7], Brian J. Pepe-Mooney[2,7-8], Colton J. Smith[7], Sanjay M. Prakadan[1,3-6], Michelle L. Ramseier[1-6], Evelyn Y. Tong[1-6], Julia Joung[3,12,15-18], Fangtao Chi[5], Thomas McMahon-Skates[7], Carolyn L. Winston[7], Woo-Jeong Jeong[2,7,19], Katherine J. Aney[2,7,19], Ethan Chen[2,7,19], Sahar Nissim[2,7,19,20], Feng Zhang[3,12,15-17], Vikram Deshpande[21], Georg M. Lauer[8], Ömer H. Yilmaz[5,13,21,25,*], Wolfram Goessling[2,7,8,19,22,23,25,*], Alex K. Shalek[1-6,25,*]

[1]Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA, USA
[2]Harvard-MIT Program in Health Sciences and Technology, Harvard Medical School, Cambridge, MA, USA
[3]Broad Institute of MIT and Harvard, Cambridge, MA, USA
[4]Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA, USA
[5]David H. Koch Institute for Integrative Cancer Research, MIT, Cambridge, MA, USA
[6]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA
[7]Genetics Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[8]Division of Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts, USA
[9]Alcohol Liver Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
[10]Gastrointestinal Unit, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
[11]Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
[12]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
[13]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA
[14]School of Life Sciences, Arizona State University, Tempe, AZ, USA
[15]Department of Brain and Cognitive Science, MA, Cambridge, MA, USA
[16]McGovern Institute for Brain Research at MIT, Cambridge, MA, USA
[17]Howard Hughes Medical Institute, MIT, Cambridge, MA, USA
[18]Whitehead Institute for Biomedical Research, Cambridge, MA, USA
[19]Dana-Farber Cancer Institute, Boston, MA, USA
[20]Gastroenterology Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[21]Department of Pathology, Massachusetts General Hospital, Boston, MA
[22]Harvard Stem Cell Institute, Cambridge, MA, USA
[23]Developmental and Regenerative Biology Program, Harvard Medical School, Boston, MA, USA
[24]These authors contributed equally
[25]These senior authors contributed equally
*Correspondence: shalek@mit.edu (A.K.S), wgoessling@mgh.harvard.edu (W.G.), ohyilmaz@mit.edu (Ö.H.Y.)

Abstract

Under chronic stress, cells must balance competing demands between cellular survival and tissue function. In metabolic dysfunction-associated steatotic liver disease (MASLD, formerly NAFLD/NASH), hepatocytes cooperate with structural and immune cells to perform crucial metabolic, synthetic, and detoxification functions despite nutrient imbalances. While prior work has emphasized stress-induced drivers of cell death, the dynamic adaptations of surviving cells and their functional repercussions remain unclear. Namely, we do not know which pathways and programs define cellular responses, what regulatory factors mediate (mal)adaptations, and how this aberrant activity connects to tissue-scale dysfunction and long-term disease outcomes. Here, by applying longitudinal single-cell multi-omics to a mouse model of chronic metabolic stress and extending to human cohorts, we show that stress drives survival-linked tradeoffs and metabolic rewiring, manifesting as shifts towards development-associated states in non-transformed hepatocytes with accompanying decreases in their professional functionality. Diet-induced adaptations occur significantly prior to tumorigenesis but parallel tumorigenesis-induced phenotypes and predict worsened human cancer survival. Through the development of a multi-omic computational gene regulatory inference framework and human *in vitro* and mouse *in vivo* genetic perturbations, we validate transcriptional (RELB, SOX4) and metabolic (HMGCS2) mediators that co-regulate and couple the balance between developmental state and hepatocyte functional identity programming. Our work defines cellular features of liver adaptation to chronic stress as well as their links to long-term disease outcomes and cancer hallmarks, unifying diverse axes of cellular dysfunction around core causal mechanisms.

Introduction

Cells must balance their immediate viability with supporting tissue homeostasis and contributing to organismal health[1]. The hepatocytes of the liver, for example, perform wide-ranging professional functions, including nutrient metabolism, protein secretion, and chemical detoxification[2,3]. Moreover, they possess substantial regenerative capacity, helping to restore normal mass and function after acute challenges as dramatic as surgical removal of two-thirds of the liver's mass[4]. However, this intrinsic regenerative ability can prove insufficient during chronic stress, resulting in progressive tissue damage[5,6]. For instance, chronic exposure to high caloric diets can precipitate metabolic dysfunction-associated steatotic liver disease (MASLD, formerly NAFLD/NASH; affecting ~25% of people around the world), which in turn drives progressive fibrosis, cirrhosis, liver failure, and hepatocellular carcinoma (HCC), the second-leading cause of years of life lost to cancer[7–13].

Epidemiological studies indicate that each successive MASLD stage associates with a progressive increase in HCC incidence[14]. However, mutations mainly accumulate after cirrhosis, but not before[15,16]: among pre-cancer liver disease patients, only patients with cirrhotic livers (but not earlier fibrosis stages) exhibited significant increases in mutation rate compared to patients with non-fibrotic livers[17]. Among a MASLD-predominant cohort, convergent somatic mutations were enriched for metabolic enzymes, but largely did not overlap with HCC driver mutations[18]. These epidemiological and mutation cohort studies raise the hypothesis that, in addition to experiencing mutational accumulation, hepatocytes may respond to chronic stress by developing progressively dysfunctional cell states that, while not solely genetically defined, are poised for tumorigenesis. Work in other organs has described environmental stressors driving non-mutational priming for longer-term dysfunction and tumorigenesis, manifesting as transcriptional and epigenetic adaptations in response to inflammation in the pancreatic and skin epithelia and high-fat diets in intestinal stem cells[19–24].

However, prior work in MASLD has largely focused on tissue-level histology or organ-level function in the context of specific gene knockouts or immune subset depletions, or examined broad contributors to cell death, such as reactive oxygen species, unfolded protein response, or lipotoxicity[7–9,25–28]. Comparatively little is known about phenotypic changes in surviving cells and

their dynamics. Outstanding questions include: which pathways and functional tradeoffs are induced with progressive exposure to environmental stressors? How do early adaptations connect to long-term consequences like tumor outcomes? And, which decision-making circuits causally mediate cellular (mal)adaptations? Knowledge of the temporal hierarchy of stress adaptations (and their accompanying disease repercussions) would help elucidate how the liver coordinates homeostatic functions while buffering stresses affecting constituent cells[29,30]. Furthermore, the discovery of cell-extrinsic and cell-intrinsic drivers of these processes could lead to novel therapeutic targets and improved patient stratification for individuals with MASLD or HCC.

Here, we examine how chronic metabolic stress drives functional tradeoffs between cellular identity and homeostatic function among hepatocytes to precipitate cancer-associated states. We conduct longitudinal single-cell multi-omics analyses of a diet-only mouse model of chronic stress via metabolic overload, spanning the stages of early steatosis to spontaneous tumorigenesis. With these resources and extensions to human MASLD/HCC cohorts, we define the progression of hepatocyte adaptation, including upregulation of early developmental markers, anti-apoptotic/pro-survival effectors, and WNT signaling. These shifts occur at the expense of core identity and professional functions, resulting in reduced expression of lineage-determining transcription factors, rate-limiting enzymes, and immunomodulatory secreted proteins. Through the development of a computational framework to discover putative regulators of disease-associated gene programs and experimental genetic perturbations (human *in vitro* and mouse *in vivo*), we validate RELB, SOX4, and HMGCS2 as causally driving hepatocyte dysregulation and inducing early shifts towards development- and cancer-associated states. Our results define principles of cellular response to chronic stress in non-transformed liver tissue and connect them to cancer-associated sequelae, suggesting avenues by which initial stress adaptations perturb cellular states, priming them for long-term tissue dysfunction and disease outcomes.

Results

Long-term high fat diet in wild-type mice mimics aspects of human MASLD and spontaneous tumorigenesis

As an exemplar of chronic stress, we studied a high-fat diet (HFD)-mediated liver injury model. In agreement with prior work, we found HFD C57BL/6 mice developed obesity, elevated serum cholesterol, increased alanine aminotransferase (ALT) levels, decreased serum albumin, and impaired glucose tolerance, indicative of hepatocellular damage, synthetic dysfunction, and diet-induced systemic insulin resistance (Fig. 1A-C, S1)[31]. Histologically, HFD livers exhibited (initially pericentral) steatosis, lobular inflammation, pericellular "chicken-wire" fibrosis, and hepatocellular ballooning (Fig. 1D-H, S1; see Supplementary Note 1 for additional contextualization against human disease progression). Spontaneous HCC developed without additional genotypic or chemical manipulation, which we validated in a second mouse cohort at a different institution (see Methods; 16/18 HFD mice vs. 4/19 CD mice developing liver tumors by 18 months; Fig. S1). These systemic and histologic findings mimic aspects of human MASLD progression: simple steatosis at 6 months, followed by certain features of inflammation, fibrosis, and spontaneous HCC tumorigenesis at 12 and 15 months.

To understand longitudinal adaptations to chronic metabolic stress, we conducted immune-biased live tissue single-cell RNA-seq (scRNA-seq; N = 17 mice, n = 23,819 cells), epithelial-biased frozen tissue single-nucleus RNA-seq (snRNA-seq; N = 18 mice, n = 79,408 cells), and frozen tissue snATAC-seq (N = 13 mice, n = 97,113 cells) for HFD and CD mice at each timepoint, recovering all major parenchymal, stromal, and immune subsets (Fig. S2-4). To further examine the fidelity of our MASLD murine model, we queried select non-parenchymal cell observations from the literature. For example, prior work has implicated human $CXCR6^+CD8^+$ T cells (expressing *TOX* and *PDCD1*) as drivers of autoaggressive hepatocyte killing in MASLD[32,33]. $Cxcr6^+Cd8^+$ T cells in our model recapitulated markers of these previously-described human cells, and T cells expressing exhaustion-related markers were compositionally enriched with HFD in our scRNA-seq dataset (Fig. 1I-K). Multiplexed immunofluorescence on mouse liver revealed that TOX^+ T cells were strongly enriched *in situ* as early as 6 months (Fig. 1L-O). Likewise, prior work has identified human scar-associated macrophages (SAMac) as being both localized to cirrhotic

regions and capable of activating collagen expression in stellate cells[34]. SAMac markers aligned with a cluster of macrophages in our scRNA-seq data; these SAMac-like mouse macrophages were enriched even at our earliest timepoint before overt fibrosis deposition (Fig. S5A-D). Collectively, these results indicate that our mouse model mirrors key molecular features of human MASLD.

Metabolic stress drives tradeoffs between pro-survival pathways and core homeostatic functions

Given the diversity of homeostatic roles performed by hepatocytes, we focused on their dynamic stress responses and adaptations. Reasoning that genes with temporally-coordinated expression trajectories might capture linked biological processes, cellular responses, or regulatory targets, we defined four expression programs demonstrating progressive or rapid alteration in response to high-fat diet: 1) "Longitudinal Increase" and 2) "Sustained Upregulation", which describe genes progressively or consistently elevated with long-term metabolic stress, respectively; and, 3) "Longitudinal Decrease" and 4) "Sustained Downregulation", whose genes progressively decline or consistent lessening, respectively (Fig. 2A; Table S1).

Examining the Longitudinal Increase program, long-term metabolic stress increased expression of genes involved in pro-survival pathways, WNT signaling, cholesterol metabolism, and intercellular signaling (Fig. 2B-D). For instance, *Igfbp1*, *Bcl2l1*, *Jun*, and *Klf6* have well-established roles in promoting hepatocyte survival and inhibiting apoptosis[35–38]. Increases in key WNT effectors (e.g., *Tcf4*) and regeneration-elevated regulators of cell cycle and senescence (e.g., *Cdkn1a*) suggest connections between stress responses and hepatocyte regeneration[39,40]. Linking stress responses and altered metabolism, hepatocytes also progressively upregulated the rate-limiting enzyme of cholesterol synthesis (*Hmgcr*) and key cholesterol uptake regulators (e.g., *Ldlr*), aligning with increases in cholesterol accumulation in hepatocytes[41,42]. Towards immune influences on hepatocyte phenotypes, we also observed upregulation of inflammation-responsive, regeneration-associated, and HCC-elevated genes like *Lcn2* (whose hepatocyte-specific knockout impairs acute regenerative capacity) and *Tm4sf4* (whose liver-specific overexpression increases inflammatory mediators like TNF and toxin-mediated acute liver

179 damage)[43,44]. Globally, externally-defined gene sets related to cytokine signaling, cell cycle
180 regulation, and WNT signaling were enriched among the Longitudinal Increase program (Fig. 2D;
181 Table S2).

182 The Sustained Upregulation program, capturing aspects of the hepatocyte response to
183 metabolic stress that are maintained over time, provided further examples of hepatocytes
184 shifting to prioritize pro-survival responses under metabolic stress (Fig. 2E-G; Table S3).
185 Hepatocytes increased expression of receptors mediating immune interactions and exerting pro-
186 survival or anti-apoptotic effects, including *Cd74* (which promotes hepatocyte survival, in
187 addition to antigen presentation, through binding to MIF)[45,46], *Cd1d1* (which mitigates iNKT-
188 mediated inflammation and hepatocyte apoptosis)[47], and *Cd59a* (which prevents formation of
189 complement membrane attack complexes on hepatocytes)[48]. Specific cholesterol synthesis and
190 lipid oxidation enzymes were also upregulated, including *Hmgcs1* (catalyzing cytoplasmic
191 acetoacetyl-CoA to HMG-CoA) and *Cpt1a* (rate-limiting enzyme for fatty acid beta-oxidation)[49,50],
192 as were metabolic regulators (e.g., *Mtor* and *Gpat4*) and enzymes catalyzing signal transduction
193 metabolites (e.g., *Hexa*, *Smpd2*, *Pde4d*)[51–55].

194 However, upregulation of pro-survival, WNT-associated, and immune interaction-related
195 responses came at the expense of processes normally associated with the professional
196 homeostatic responsibilities of hepatocytes (Fig. 2H-J; Table S4). *Hnf4a* was progressively
197 downregulated with long-term metabolic stress; this is particularly notable given its role as a
198 master regulator of hepatocyte identity, with wide-ranging effects on hepatocyte functions and
199 fate specification during development[56]. Concordantly, the Longitudinal Decrease program also
200 included enzymes related to peroxisomal oxidation (e.g., *Acaa1a*), aldehyde processing (e.g.,
201 *Aldh5a1*), CoA biosynthesis (*Pank1*; rate-limiting), and acyl-CoA processing (e.g., *Acsl1*)[57–61].
202 Progressive decreases in *Hmgcs2*, the rate-limiting enzyme of ketogenesis, were notable given
203 that: 1) ketogenesis and cholesterol synthesis compete for the same starting metabolites[62]; and,
204 2) many cholesterol synthesis enzymes (including *Hmgcr*) followed opposing trajectories. Genes
205 capable of driving cell death and serving as tumor suppressors, like *Nlrp12* and *Dapk1*[63,64], were
206 also downregulated. Regulators of hepatocyte phenotype, like *Prlr* and *Nr1h4* (encoding FXR),
207 were likewise reduced with long-term metabolic stress. *Prlr* serves as a receptor for hormones

including prolactin and growth hormone, and is upregulated by estrogens – notable given MASLD's dramatically lower incidence in pre-menopause women[65,66]. The bile acid-activated nuclear receptor FXR, meanwhile, has been targeted in clinical trials via agonists including obeticholic acid, but the Phase III REVERSE trial recently failed to reach its primary endpoint[67]; we speculate that this may indicate opportunities for improved patient stratification based on expression and/or protein abundance of FXR.

Hepatocytes' Sustained Downregulation program also reflected reductions in core hepatocyte and liver functions including secreted complement (*C8a*; also *C6*, *C8b*), coagulation factors (e.g., *Fgg*, *F10*), urea cycle (e.g., *Cps1*; rate-limiting enzyme), and bile acid regulation (e.g., *Abcb11*) (Fig. 2K-M; Table S5)[68–70]. *Lsr* regulates circulating triglyceride levels, and its experimental knockout has been shown to increase weight gain[71], suggesting maladaptive repercussions of its diet-induced downregulation. We also observed downregulation of genes involved in proteostasis (e.g., *Pdia5*) and chromatin interactions (e.g., *Atxn1*)[72,73]. Suggestive of protection against hepatocyte self-destruction and complementing genes captured in other modules, downregulation of complement protein production (e.g., *C6*, *C8a*, *C8b*) harmonizes with the role of *Cd59a* (Sustained Upregulation program) in preventing complement-mediated hepatocyte death. *Esr1* encodes estrogen receptor alpha[74], and its downregulation aligns with progressive losses of *Prlr* (Longitudinal Decrease program). Likewise, *Lifr* knockout is associated with increases in *Lcn2* secretion (Longitudinal Increase program) and elevated tumorigenesis[75], potentially connecting metabolic stress-induced downregulation of an upstream receptor to subsequent pro-inflammatory, cancer-associated hepatocyte responses.

Towards validating dynamic shifts in hepatocyte phenotypes under chronic stress, we focused on HNF4A given its essential role in regulating hepatocyte differentiation and function. *In situ* mouse liver immunofluorescence demonstrated progressive decreases in HNF4A nuclear protein abundance with chronic metabolic stress, aligning with its Longitudinal Decrease transcriptional trajectory (Fig. 2N-O). We additionally cultured human liver HepG2 cells in lipid-rich media, and observed gene expression changes and functional metabolic and synthetic dysregulation concordant with sustained *in vivo* hepatocyte shifts (Fig. S5E-K, Methods)[76,77]

Thus, hepatocytes adapt to chronic metabolic stress by increasing expression of pro-survival responses, including anti-apoptotic effectors, WNT signaling, and intercellular interaction mediators. This pro-survival adaptation comes at the expense of genes linked to core hepatocyte identity and professional functions, including multiple classes of secreted molecules, transcription factors (including HNF4A and FXR), and metabolic enzymes.

Adaptations to metabolic stress extend to human MASLD progression, exhibit extreme manifestations in cancer, and are prognostic for cancer survival

We next examined whether chronic stress adaptations in hepatocytes connected to tumor phenotypes, predicted long-term disease outcomes, and extended to human patient cohorts. We performed non-invasive MRI on mice fed a high-fat diet for 15 months to identify potential tumors, then dissected lesions that corresponded to grossly abnormal areas (Fig. 3A). Histologic evaluation supported liver tumors' classification as HCC, with preserved liver architecture but also pleomorphism, prominent nucleoli, and hyperchromasia (Fig. 3B). Using snRNA-seq, we also found that tumors in our mouse model recapitulated human HCC markers, such as WNT and Notch pathway members (e.g., *Axin2*, *Tbx3*, *Ctnnb1, Lgr5*, *Jag1*), IGF signaling (e.g., *Igfbp1*, *Igf2r*), HCC diagnostic biomarkers (e.g., *Afp*, *Gpc3*), and proliferation (e.g., *Mki67*) (Fig. 3C, S6)[35,78–81]. More holistic analyses revealed enrichment for previously defined liver cancer marker sets with the expected directionality (Fig. S7A; Table S6-7)[82].

We next sought to understand whether and how chronic metabolic stress adaptations in non-transformed hepatocytes aligned with tumorigenesis-associated alterations. Tumor cells exhibited elevated expression of the Longitudinal Increase program relative to adjacent, non-transformed hepatocytes (Fig. 3D). External human cohorts exhibited a similar pattern, with the Longitudinal Increase program positively associated with human MASLD severity across microarray, bulk RNA-seq, and snRNA-seq readouts (Fig. 3E-F; Fig. S7B)[76,83–86]. Protein-level abundances of the Longitudinal Increase program likewise increased across MASLD progression in human liver tissue proteomics (Fig. 3G)[87]. Finally, we found that the Longitudinal Increase program was further heightened within human HCC and prognostically stratified HCC patient survival[88], with high expression predicting worsened outcomes (Fig. 3F,H). Thus, the same

pathways activated as adaptations to metabolic stress in non-transformed hepatocytes (e.g., pro-survival effectors, WNT activation) are additionally upregulated with tumorigenesis, so that tumor cells exhibit a heightened manifestation and extension of metabolic overload-induced adaptations.

The Longitudinal Decrease, Sustained Upregulation, and Sustained Downregulation programs also significantly associated with disease severity, tumor phenotypes, and patient survival outcomes (Fig. 3I-M; Fig. S7C-Q). As internal consistency, the Longitudinal Increase and Sustained Upregulation programs displayed opposite directionalities to the Longitudinal Decrease and Sustained Downregulation programs (i.e., worsened vs improved prognoses, respectively). Additionally, across studies, measurement modalities, and species, aggregate module scores were driven by similar underlying genes, with strong positive covariance between individual genes and overall module scores (Fig. S8A-D).

To further explore connections between early stress adaptations in non-transformed hepatocytes and later tumor phenotypes and outcomes, we investigated whether HCC-associated signatures were observed early in MASLD before overt tumorigenesis. Upon examining signatures of 1) cancer-linked chemical and genetic perturbations[82,89,90]; 2) HCC mutational subtypes[91]; and, 3) liver development and regeneration[92–96], we found that hepatocytes take on gene expression patterns reminiscent of earlier developmental stages as well as the HCC S1 subclass even early in MASLD: development-associated expression programs (driven by genes including *Krt8*, *Sox9*, and *Cd24a*, Fig. S8E) were increased not only in HCC, but even at early disease stages in the MASLD progression in non-transformed hepatocytes across species and measurement modalities (Fig. 4A-E); likewise, marker genes of the human HCC S1 subclass, characterized by aberrant WNT activation, were elevated in hepatocytes across species at early MASLD stages and further elevated with tumorigenesis (Fig. 4F-J, Fig. S8F)[97]. Notably, β-catenin regulates WNT signaling activation and is a commonly mutated HCC driver gene[15,16], suggesting pathway-level convergence between stress-induced transcriptional alterations and literature-established genomic drivers.

To better understand the broader spatial context of our observations, we inferred hepatocytes' spatial positions via periportal-vs-pericentral markers (Fig. 4K)[98]. Pre-malignant

pericentral hepatocytes exhibited larger increases in the WNT activation-associated HCC S1 signature than did periportal hepatocytes (Fig. 4L). Pericentral endothelial cells are a key source of secreted WNT ligands[99]. Upon similarly inferring endothelial cell zonation, we found that pericentral endothelial cells (but not periportal or mid-lobular endothelial cells) upregulate expression of *Rspo3*, *Wnt2*, and *Ctnnb1* with chronic metabolic stress, potentially aligning with spatially-structured signaling circuits (Fig. 4M, S9A-J; see Supplementary Note 2 for a discussion of the activation of hepatocyte chronic stress adaptation programs during acute regeneration and analyses of intercellular interactions potentially shaping hepatocyte adaptation programs). In examining whether these patterns were etiology-specific vs. generalizable, we found that hepatocytes' metabolic adaptation programs were consistently dysregulated across HCC tumors linked to metabolic disorders (i.e., MASLD, alcohol-related liver disease), viral infection (i.e., hepatitis B and hepatitis C virus), or no known risk factors (Fig. 4N). Internally-consistent directionalities across HCC etiologies support the generalizability of these hepatocyte functional adaptations in wide-ranging disease microenvironments.

Overall, these results help to link metabolic stress-induced adaptations in non-transformed hepatocytes to later tumor phenotypes: extensions to human cohorts, similarities to later tumor states, and predictive power for disease severity and HCC patient survival. The directionality of effects on human HCC survival further helps propose interpretations for these axes of hepatocyte adaptation: while elevated processes could plausibly be linked to improved survival of individual cells, they incur longer-term repercussions through reductions in the expression of genes related to hepatocyte identity-defining features and professional functions, as well as early activation of pathways that may eventually contribute to tumorigenesis and worsened survival at later disease stages.

Epigenetic dysregulation and WNT pathway priming under chronic metabolic stress

Having defined hepatocytes' longitudinal adaptations to metabolic stress, we sought to obtain mechanistic insights into the epigenetic landscape and cell-intrinsic regulatory factors shaping hepatocyte expression patterns under stressful disease microenvironments (Fig. S4).

Comparing hepatocyte metabolic adaptation programs across genomic regulatory layers, we found concordance between transcriptional expression and epigenetic accessibility (Fig. S10A).

To identify transcription factors (TFs) altered by metabolic stress in hepatocytes, we examined genome-wide accessibility of chromatin peaks containing TF binding motifs[100,101] in our snATAC-seq dataset (Fig. 5A). We observed progressive increases in motif accessibility for members of the AP-1 complex (e.g., FOS, JUN), which play roles in responses to cellular stressors and mediate long-lasting epigenetic rewiring and tissue memory of inflammation in other compartments[23,102]. Increased TEAD motif accessibility suggests involvement of the Hippo pathway, which regulates liver regeneration and development[103,104]. Reinforcing our inference of early activation of WNT pathway and development-associated programs, we observed increased motif accessibility for WNT-associated TFs (e.g., TCF7, LEF1) and TFs active during liver development (e.g., SOX4, SOX9)[105,106]. As TFs with increased motif accessibility across both early and long-term metabolic stress, we found PPAR members (regulating liver metabolic pathways and serving as a leading MASLD drug target, but also driving stemness and regeneration phenotypes in intestinal stem cells), RXRA-associated nuclear receptor TF complexes, and NFE2L1 (a key regulator of cellular responses to oxidative stress)[107–111].

Towards higher-resolution insights into temporally-varying chromatin states, we conducted pseudotemporal analyses to order hepatocytes according to smooth gradients in chromatin accessibility, prioritizing gene loci dynamically altered by chronic metabolic stress (Fig. 5B, S10B-E). Chromatin trajectories reinforced our multi-omic, cross-species analyses of hepatocyte adaptations: members of the Longitudinal Decrease program including *Hmgcs2*, *Acsl1*, *Scp2*, *Hnf4a*, and *Rxra* exhibited maximal chromatin accessibility early in the high fat diet pseudotime progression, whereas members of the Longitudinal Increase program like *Lcn2*, *Bcl2l1*, and *Il1r1* peaked towards the pseudotime terminus. Progressive decreases in chromatin accessibility at *Hnf4a*'s gene locus are notable given its role as a master regulator of hepatocyte identity[112], and aligns with its transcriptional and proteomic downregulation, decreases in professional hepatocyte functions, and increases in developmental marker expression. These (pseudo)temporal patterns were not observed in hepatocytes from control diet mice, supporting distinctive trajectories of stress-induced chromatin remodeling (Fig. S10B-F).

We additionally sought to identify genes where epigenetic alterations preceded and presaged transcriptomic shifts as these may indicate stress-induced epigenetic priming: chromatin remodeling establishing accessible epigenetic landscapes prior to transcriptional alterations, thereby priming cells for later activation and (dys)function[113]. We leveraged co-accessibility between intergenic chromatin peaks and promoters or gene bodies to create peak-gene linkages, capturing enhancer-gene regulatory interactions despite potentially large genomic distances[114]. The regeneration-upregulated, WNT/β-catenin target *Axin2* provides an example, with several distal chromatin regions co-accessible with *Axin2*'s promoter/gene body and elevated in accessibility with HFD across timepoints (Fig. 5C). More broadly, a variety of WNT-associated (e.g., *Tbx3*, *Axin2*, *Lgr5*, *Notum*) and HCC-linked (e.g., *Spp1*, *Igf2r*) genes exhibited increased epigenetic accessibility but only small changes in transcription at our earliest timepoint (6 months; Fig. 5D-E). However, these genes were strongly upregulated after tumorigenesis months later (15 months; Fig. 5E). Genes with the opposite directionality (i.e., early decreases in chromatin accessibility preceding more extreme transcriptional downregulation with longer-term metabolic stress) included: 1) metabolic enzymes and secreted proteins (e.g., *Aspg*, *Hao1*, *Hamp*); 2) suppressors of HCC and fetal hepatocyte-associated phenotypes (*Gls2* and *Zbtb20*, respectively); and, 3) *Esr1* as a receptor for MASLD-protective estrogen signaling (Fig. 5E,F)[115,116].

Thus, paralleling hepatocytes' transcriptomic and proteomic stress adaptations, discovery of early WNT- and HCC-linked chromatin changes that foreshadow transcriptional shifts suggests that even early exposure to metabolic stress may establish a permissive chromatin landscape that contributes to later activation of regeneration-, development-, and cancer-associated pathways.

MATCHA prioritizes causal transcription factors shaping MASLD-associated phenotypes

Computational nomination of driving TFs for functionally-important gene programs (e.g., disease-linked stress adaptation programs in hepatocytes) remains an unsolved, open problem (see Supplementary Note 3 for contextualization of prior work). To accomplish this, we developed MATCHA (Multiomic Ascertainment of Transcriptional Causality via Hierarchical Association), a computational framework to map user-specified gene programs (e.g., arbitrary

biological processes, disease (mal)adaptations, etc.) to distal enhancers and program-specific TF activities (see Methods). In brief, MATCHA links gene programs to cell-type-specific distal enhancers by identifying chromatin regions co-accessible with the gene program's promoters or gene bodies. MATCHA then prioritizes causal regulators by determining TF motifs whose accessibility at program-coaccessible enhancers likewise covaries with program transcriptional expression. MATCHA further optionally incorporates: 1) concordance across datasets towards robust regulatory inference (e.g., across species, single-cell vs. bulk measurements, etc.); and, 2) identification of TFs co-regulating multiple gene programs. MATCHA therefore enables prioritization of TFs driving arbitrary gene programs while also modeling context-dependent functions via cell type- and tissue-specific gene regulatory landscapes (see Data and Materials Availability).

As proof-of-concept, we examined two metabolic-stress-relevant processes with known driver TFs: ER stress response and beta-oxidation (defined externally through GO:BP)[82,89,90]. MATCHA recovered ground-truth causal TFs for these external test cases: the top two TFs prioritized for GO:BP ER stress response genes were XBP1 and ATF6 (i.e., two of three well-established master regulatory TFs) (Fig. S11A-G)[117]. The top 5 TFs for GO:BP beta-oxidation, meanwhile, included FXR/*NR1H4* and PPARA (whose agonists have advanced to Phase III clinical trials for MASLD)[110,118], and NR1I2, CEBPA, and NR1I3 (all with preclinical evidence for regulation of liver metabolism and lipid accumulation; Fig. S11H-N)[119,120].

To identify core TFs mediating hepatocyte longitudinal stress adaptations and early induction of development-associated and HCC-linked cell states, we applied MATCHA to construct a bipartite network of regulatory relationships between gene programs and TFs, supported by epigenetic and transcriptional evidence (Fig. 5F). Our network successfully recapitulated known ground truths, but also nominated targets for experimental validation. In addition to previously-discussed well-known drivers of ER stress (i.e., XBP1, ATF6) and beta-oxidation (i.e., PPARA, NR1H4/FXR), our network captured hepatocyte developmental regulation (i.e., SOX9 driving development states), hormonal influences on metabolism (i.e., androgen receptor driving beta-oxidation[121,122]), and stress response mediators (i.e., NFE2L1/NRF1 driving oxidative stress response[123]). We also note inferred regulatory links between: 1) THRB motifs; 2)

activation of hepatocyte identity-linked Longitudinal Decrease and Sustained Downregulation programs; and, 3) repression of development-associated and HCC S1 WNT programs. We note that the THRB agonist resmetirom is currently undergoing the Phase III MAESTRO-NASH trial[124]. We chose 15 TFs (20 isoforms) to prioritize for experimental validation (Fig. S11O).

RELB and SOX4 mediate wide-ranging MASLD adaptation and tumor-associated functional phenotypes

To validate MATCHA-nominated drivers of hepatocyte metabolic adaptation, we conducted arrayed human *in vitro* genetic perturbations[125]. We created HepG2 cells stably overexpressing TF isoforms and cultured them in lipid-rich media, followed by: 1) scRNA-seq to validate TF effects on hepatocyte transcriptomic phenotypes (n = 10,522 cells; median 417 cells per TF isoform); and, 2) live-cell imaging and immunofluorescence to validate TF effects on functional metabolic and cancer-associated phenotypes (Fig. 6A, S12; Table S8-10). Transcriptomic profiles of negative controls (non-transduced and BFP-transduced cells) clustered with each other by media condition but not transduction status, suggesting: 1) specificity of measured responses (via separation from TF-transduced cells); and, 2) preserved responses to lipid-induced metabolic stress following transduction (Fig. 6B). As a positive control, we confirmed that overexpression of PPARA in lipid-rich culture led to upregulation of known targets, including *HMGCS2*, *CPT1A*, *PLIN2*, *G6PC1*, and *PCK1* (Fig. 6B, S12C)[126].

With successful recovery of known TF-target relationships, we evaluated how MATCHA-nominated TFs regulated hepatocyte stress adaptations. Downstream targets of RELB, RORC, and MAFF overexpression included developmental markers (e.g., *EPCAM*), pro-survival effectors (e.g., *KLF6*, *JUN*), cholesterol synthesis and fatty acid oxidation enzymes (e.g., *SC5D*, *CPT1A*), and receptors involved in HCC-relevant immune interactions (e.g., *CD74*, *LGALS3BP*)[46,127]. Overexpression of SOX4 and ONECUT1 (active during liver maturation) drove downstream WNT pathway targets and regulators (e.g., *DKK1*, *DLK1*) and developmental markers (e.g., *CD24*, *KRT19*), among others[105,128]. Additionally, SOX4 (along with JUND, KLF4, and CUX2) upregulated metabolism-associated genes, such as lipid droplet-associated *PLIN2* and *CPT1A* (rate-limiting enzyme for beta-oxidation)[21,129]. With individual MATCHA-nominated TFs capable of regulating

wide-ranging aspects of hepatocytes' stress adaptations, we proceeded to focus on RELB and SOX4 given the strength of their effect on metabolic adaptation-associated transcriptional and functional states.

Overexpression of RELB (a member of the non-canonical NF-κB signaling complex) in lipid-rich media drove transcriptional shifts consistent with hepatocytes' *in vivo* long-term metabolic adaptations: elevation of the Longitudinal Increase, Sustained Upregulation, and HCC S1 signature programs, and decreases in the Longitudinal Decrease and Sustained Downregulation programs (Fig. 6C). Specific RELB targets included increases in development-associated and regeneration-linked markers (e.g., *CD24*, *CDKN1A*, trend towards *LGR5*), decreases in hepatocyte secreted protein products (e.g., *FGB*, *FABP1*), and upregulation of HCC-associated genes (e.g., p62/*SQSTM1*, *CD151*, *LGALS1, CD74*) (Fig. S13A). Contextualizing against *in vivo* shifts, we observed that RELB drove effect sizes that were smaller than, but of comparable magnitude to, tumor-vs-healthy differences (Fig. 6C), suggesting both the breadth and strength of RELB on hepatocyte stress adaptations. Towards *in vivo* MASLD relevance, RELB expression increased with fibrosis stage across multiple human cohorts, and RELB as a single marker stratified patient survival, associating with worsened outcomes (Fig. 6D-E). For further human *in vivo* evidence of RELB regulation of hepatocytes' stress adaptations, we examined how human HCC patients' copy number variations (CNVs) at the RELB locus altered expression of downstream gene programs, analogous to a "natural genetic perturbation experiment" in human tumors[130]. RELB CNVs drove significant changes in *RELB* transcription and predicted: 1) significant increases in the development-associated, HCC S1 WNT Activation, and Sustained Upregulation programs; 2) significant decreases in the Sustained Downregulation and Longitudinal Decrease programs; and, 3) non-significant elevation of the Longitudinal Increase program (Fig. S13B-H). Thus, human *in vitro* genetic perturbations and human *in vivo* HCC CNVs support RELB as a driver of hepatocyte stress adaptation programs, loss of cell identity, and development-associated and cancer-linked states.

Overexpression of SOX4 (SRY-box transcription factor 4, active during liver development) in lipid-rich media also depleted genes associated with hepatocyte cellular identity and function including TFs (e.g., *NR1H4*, *PPARA*, trend towards *HNF4A*), metabolic enzymes (e.g., *GPD1*,

AKR1D1, *DPYD*, *EHHADH*), and secreted protein products (e.g., *SERPINF2*, *PLG*, *FABP1*, *FGG*) (Fig. 6F, S13I). In contrast, SOX4 increased expression of developmental markers (e.g., *CD24*, WNT targets *LGR5* and *NKD1*), as well as genes with functional effects in HCC (e.g., *CD151*, *MDK*, *AIFM2*, *AKR1C2*, *ROBO1*). As a broader, orthogonal examination of SOX4 driving loss of hepatocyte identity, we examined how SOX4 overexpression altered genes enriched in hepatocytes relative to all other cell types[131], as a proxy for the distinguishing features and cell identity of hepatocytes. 89% of statistically-significant hepatocyte identity-associated genes were downregulated by SOX4 expression, supporting SOX4 as driving loss of hepatocyte identity (Fig. S13J). Towards human *in vivo* relevance, we found that *SOX4* expression increased with MASLD severity and stratified HCC patient survival, associating with worsened survival (Fig. 6G-H).

Functionally, overexpression of RELB and SOX4 decreased cellular lipid accumulation, but also caused accompanying dose-dependent increases in ROS accumulation (Fig. 6I-J, M). These effects may be mediated by shared downstream target genes linked to reduced lipid accumulation (e.g., increased *ABCC1* and *CPT1A*, decreased *FABP1*) and elevated oxidative stress (e.g., decreased *GSTA1*, decreased *MT1E* and *MT2A*) (Fig. S13I)[132–137]. TF-mediated tradeoffs between lipid and ROS accumulation may align with prior work proposing a potentially protective role for lipid droplets, which can sequester lipid species that might otherwise drive lipotoxicity or increase ROS levels upon processing and oxidation[138,139]. To connect these factors to longer-term disease outcomes, we further examined the effect of RELB and SOX4 on proliferation. Overexpression of RELB or SOX4 each increased nuclear accumulation of p53 protein (Fig. 6K,M). SOX4 overexpression drove increased proliferation as measured by nuclear KI67; RELB overexpression did not drive significant changes in nuclear KI67, which could be interpreted as preserved proliferative capacity despite cellular stress (Fig. 6L-M; see Supplementary Note 4 for discussion of p53 regulation of hepatocyte phenotypes and proliferation, and especially connections to metabolism, development-associated states, WNT, and AP-1 signaling).

Through human *in vitro* genetic perturbation experiments and extensions to human cohorts, we validated RELB and SOX4 as causal regulators of stress-induced transcriptional and

functional tradeoffs between hepatocyte identity and dysfunction-associated phenotypes, unifying hepatocytes' metabolic adaptations around specific regulatory nodes.

HMGCS2 is a metabolic mediator of hepatocytes' adaptation and shifts towards cancer-associated phenotypes

We finally sought to demonstrate the *in vivo* importance of dynamic shifts in hepatocyte stress adaptation programs. We chose to focus on HMGCS2 (3-hydroxy-3-methylglutaryl-CoA synthase 2) and the regulatory effects of ketogenesis-cholesterol metabolic rewiring, given HMGCS2's: 1) role as the rate-limiting enzyme of ketogenesis; 2) strong upregulation with PPARA activation (Fig. S12C); 3) cross-species expression decreases with longitudinal metabolic stress; and, 4) association with low expression and worsened HCC survival (Fig. S14A-D). We generated a mouse model with hepatocyte-specific knockout of HMGCS2, analogous to *Hmgcs2* expression decreases with the natural stress adaptation progression (*Hmgcs2*$^{fl/fl}$; *Alb-Cre*) (Fig. S14A-D). After 6 months on either HFD or CD, steatosis occurred in mice on HFD with either wildtype (WT) or hepatocyte-specific HMGCS2 knockout (HepKO), but was comparatively less severe in CD mice of either genotype (Fig. 7A). Liver damage (measured by circulating transaminases) and cholesterol increased with the combination of HFD and HMGCS2 HepKO (Fig. S14E-F). Validating the HMGCS2 knockout, HMGCS2 abundance was decreased throughout the liver lobule via immunohistochemistry, and we observed reduced circulating ketone bodies following a 24-hour fast (Fig. 7B-C). However, HMGCS2 HepKO did not alter circulating glucose concentrations or HFD-driven weight gains, indicating specificity of effects on ketogenesis (Fig. 7D-E).

To investigate the effects of HMGCS2 loss on liver adaptation under chronic metabolic stress, we conducted snRNA-seq at the 6-month timepoint across HMGCS2 genotypes and diet conditions (N = 8 mice, n = 27,119 cells) (Fig. S15; Table S11). Neighborhood-based analyses of our snRNA-seq data indicated compositional shifts specific to the metabolic stress response of HepKO mice, but not that of WT mice (Fig. 7F). To validate our compositional analyses, we focused on Kupffer cells, where immunohistochemistry against the Kupffer cell-predominant marker CD68 supported their compositional enrichment with the interaction of metabolic stress and hepatocyte-specific HMGCS2 loss (Fig. S14G-I).

To understand the causal role of HMGCS2 and ketogenesis on hepatocyte phenotypes, we mapped gene expression profiles of HFD.WT and HFD.HepKO hepatocytes to the natural progression of chronic metabolic stress (Fig. 7G). Approximately equal proportions of HFD.WT hepatocytes mapped to the natural progression 12-month (44%) and 6-month (52%) conditions. However, with HFD.HepKO, over twice as many hepatocytes mapped to the 12-month HFD condition (63%) as to the 6-month HFD condition (30%), indicative that HMGCS2 loss causes accelerated dysfunctional phenotypes associated with later stages of hepatocyte stress adaptation.

More directly, we examined how HMGCS2 HepKO under metabolic stress altered hepatocyte stress adaptation programs. HMGCS2 HepKO on HFD led to extreme expression states, even relative to diet-induced shifts in WT mice: larger elevation of Sustained Upregulation, Longitudinal Increase, Development-Associated, and HCC S1 WNT Activation programs, but also further reductions of Sustained Downregulation and Longitudinal Decrease programs (Fig. 7H). To uncover specific targets, we prioritized genes exhibiting emergent transcriptional shifts with metabolic stress and HMGCS2 HepKO. Diverse cholesterol synthesis-related genes were upregulated with the combination of HMGCS2 HepKO and HFD (e.g., *Hmgcs1*, *Srebf2*, *Hmgcl*, *Mvk*), in line with compensation at the metabolic branch point between ketogenesis and cholesterol synthesis (Fig. 7I,J). Additionally, HMGCS2 HepKO on HFD induced cell states directly associated with long-term dysfunction, including HCC-linked intercellular signaling proteins (e.g., *Spp1*, *Lcn2*, *Lgals1*) and development-associated markers (e.g., *Afp*, *Axin2*, *Hes1*, *Cd24a*, *Sox9*). Loss of HMGCS2 on the background of HFD also accentuated decreases in expression associated with hepatocyte identity (e.g., *Hnf4a*) and function (e.g., *Cps1*, *Pck1*, *Hamp*, *C4a*, *Pdia5*). Finally, connecting HMGCS2 to MATCHA TF inference and experimental validation, HMGCS2 loss under chronic metabolic stress significantly upregulated *Sox4*.

Thus, *in vivo* genetic perturbation validated wide-ranging regulatory effects of HMGCS2 itself, but also more broadly the functional importance and interpretation of our stress adaptation programs. Premature HMGCS2 loss induced accelerated damage and extreme manifestations of transcriptional programs, including compensatory cholesterol synthesis, HCC phenotypes, development-associated states, and loss of hepatocyte canonical functionality.

552 Experimental *in vivo* modeling of accelerated stress adaptation progression (through genetic
553 manipulation) further supported adaptation programs derived in this work as fundamental,
554 functionally-important axes of hepatocyte response to environmental stressors.

Discussion

During chronic stress, cells must balance survival against performing their professional functions. We investigated how the liver manages longitudinal tradeoffs under environmental stressors through the paradigm of chronic metabolic overload, which precipitates progressive steatosis, inflammation, fibrosis, cirrhosis, and malignant transformation. In addition to its pressing clinical need, the biological context of MASLD offers opportunities to understand how initial cellular adaptations connect to longer-term tissue dysfunction, disease pathogenesis, and patient prognosis.

We developed a diet-only mouse model that exhibits functional, histologic, and cellular phenotypes paralleling human MASLD. Our longitudinal single-cell multi-omics datasets, ranging from early steatosis to late spontaneous tumorigenesis (with matched diet controls) along with harmonized human MASLD/HCC transcriptomic and proteomic cohorts, provide rich computational resources on the intersection of aging and metabolic stress. Importantly, as our mouse model does not require genetic manipulation or exogenous chemical insults, it may be a broadly-translatable experimental resource for further investigations. We acknowledge that HFD mice do not develop significant bridging fibrosis or nodule formation, possibly due to lifespan differences between mice and humans, thereby creating opportunities for further model development for investigations focused on stellate cells and fibrosis.

We demonstrated that chronic metabolic stress induces development-linked and cancer-associated adaptations in hepatocytes to the detriment of cellular identity and professional tissue functions. Hepatocytes increased genes related to: 1) early developmental stages; 2) pro-survival and anti-apoptotic effectors; and, 3) intercellular signaling (including WNT). In contrast, hepatocytes downregulated genes underpinning homeostatic roles, including diverse secreted proteins and metabolic enzymes. These findings were corroborated across multiple human cohorts, and recapitulated across epigenetic, transcriptomic, and proteomic readouts.

Suggesting discovery of generalizable stress adaptation responses, gene programs uncovered in this work exhibited extreme manifestations in acute regeneration and across HCC risk factors (consistent across etiologies including MASLD, alcohol, and viral infections). These connections support further investigations of uncovered gene programs' generalizability as core,

584 conserved axes of hepatocyte adaptation to diverse stressors. Such conserved axes in cross-
585 disease cohorts could uncover broadly applicable vs. disease-specific therapeutic vulnerabilities
586 and patient stratification hierarchies. As potential mechanisms for conserved cross-disease stress
587 adaptation programs, distinct molecular changes associated with each etiology may converge on
588 similar intracellular mediators (e.g., analogous to diverse PAMPS or DAMPs converging on TLR
589 activation)[140]. Alternatively, each etiology may drive recurrent microenvironmental or immune
590 signals that in turn produce consistent hepatocyte phenotypes (e.g., regeneration contributions
591 of both pro-inflammatory IL-6 and pro-fibrotic TGFB1)[141,142].

592 Progressive decreases in lineage-determining HNF4A, decreases in genes mediating
593 canonical hepatocyte functions, and increases in developmental markers raise the question of
594 how stress adaptations overlap or contrast with dedifferentiation. Partial hepatocyte
595 dedifferentiation occurs following acute stressors like experimentally-induced partial
596 hepatectomy and acetaminophen overdose, where subsets of hepatocytes downregulate
597 canonical functions and assume fetal-like phenotypes to enable regeneration and restoration of
598 liver mass[4–6,99,143]. Genetically-induced priming of dedifferentiation improves hepatocyte survival
599 during subsequent acute stressors, but long-term dedifferentiated states predispose to liver
600 failure and worsened survival in HCC[92,144–146]. In cancer, dedifferentiation is acknowledged as a
601 recurrent hallmark, where cancer cells unlock fetal, plastic cell states associated with elevated
602 proliferation and tumor progression[147,148]. One potential model to unify our findings involves
603 clinically-relevant chronic stresses driving partial dedifferentiation-associated states even in non-
604 transformed hepatocytes. In this model, the trajectory of hepatocytes' progressive stress
605 adaptations would increasingly draw upon the liver's regenerative capacity for improved survival
606 of individual cells, but with deleterious repercussions: 1) maladaptive tradeoffs with the liver's
607 tissue-level functions and homeostatic setpoints; and, 2) early induction and priming of similar
608 transcriptional programs as occur with tumorigenesis and worsened cancer outcomes. Future
609 demonstrations functionally linking stress adaptation and de-differentiation (especially early in
610 disease progression in non-transformed hepatocytes) may further delineate transcriptional and
611 epigenetic contributions to liver failure and elevated cancer incidence.

Extending hepatocytes' immediate adaptations, we also found WNT signaling members and HCC markers exhibited increased chromatin accessibility months prior to transcriptional upregulation with tumorigenesis. These findings suggest stress adaptations incurring not just direct changes, but also epigenetic priming for later activation of cancer-associated states. In other organs and diseases, AP-1 synergizes with context-specific TFs to maintain poised epigenetic landscapes and memory of past inflammation[23]; future work could investigate the minimal requirements and timescales needed for epigenetic dysregulation and eventual phenotypic manifestations. For instance, in the skin and pancreas, inflammatory bouts lasting only a handful of days are sufficient to drive effects upon triggers administered months later (e.g., improved wound healing, increased cancer risk)[19,20,24,149], which would suggest even short stressors can instill persistent tissue memory and altered tissue function. In contrast, in the nasal epithelium, modulation of cytokine signaling can partially revert dysfunctional progenitor memory of allergic memory, towards therapeutic restoration of aspects of homeostatic baseline[102,150]. Hepatocytes' primed accessibility at AP-1 binding motifs and WNT-related loci occurred at relatively early disease stages which precede when many patients exhibit symptoms[7]. Thus, future investigations into the timescales of stress adaptations' initiation and persistence are motivated by not only fundamental biological understanding but also the longer-term clinical relevance of whether hepatocyte tissue memory and lasting maladaptive phenotypes may be seeded prior to disease diagnosis. With weight loss capable of stabilizing or even reversing histologic features of MASLD, whether and to what degree hepatocytes' memory of chronic stress is reversible or establishes lasting (mal)adaptive responses even after return to normal weights could shape long-term implications of ongoing trials and novel therapeutic avenues for MASLD[111,151,152].

To extend from discovering axes of cellular stress adaptations to uncovering their causal regulators, we developed MATCHA, a computational framework to prioritize TFs driving arbitrary, user-specified gene programs. MATCHA infers gene program – co-accessible enhancer – causal TF triads by leveraging: 1) multimodal -omics data across layers of genome regulation, temporal trajectories, and species; and, 2) context-dependent regulatory relationships ensuring output predictions capture disease- and cell type-specific TF activity. When applied to externally-defined

biological processes, MATCHA recovered their well-established ground truth drivers. When applied to uncover "central hub" TFs with strong co-regulatory connections across multiple stress adaptation programs derived in this work, MATCHA re-discovered targets of multiple Phase III clinical trials, but also highlighted comparatively less-explored TFs that we experimentally validated. A strength of MATCHA is its emphasis on incorporating tissue and cell type-specific data as the basis for computational inference of regulatory relationships, capturing context dependence of chromatin landscapes and TF activity. Future improvements of MATCHA could include prioritizing sets of TFs needed to drive the breadth of a gene program (e.g., beyond top TFs prioritized in isolation that may have overlapping, noncomprehensive target genes). Alternatively, in addition to predicting TF regulatory effects on a gene program, we could seek to elucidate differential gene regulatory network structures for each TF under different contexts (e.g., longitudinal stress adaptation stages) to reveal context-specific alterations in enhancer binding and downstream target genes.

Validating MATCHA predictions, RELB and SOX4 indeed regulated hepatocytes' stress adaptation transcriptional programs and functional metabolic phenotypes in human *in vitro* genetic perturbation experiments. RELB constitutes the transcriptional effector of the non-canonical NF-κB signaling pathway[153]. In the liver, significant prior work investigated cytoplasmic regulation of NF-κB kinase complexes, as well as canonical NF-κB signaling's roles in tumorigenesis and hepatocyte survival under inflammatory conditions[26,154–160]. However, canonical and non-canonical NF-κB differ significantly in terms of input signaling mechanisms and output phenotypic effects: they are activated by distinct ligands, involve separate intracellular interactions, and culminate in different TFs[153,161,162]. Thus, the role of RELB in hepatocytes' progressive remodeling under metabolic stress has received less attention[163]. Our work points towards novel contributions of RELB to driving hepatocytes' (mal)adaptations to chronic stress, across disease associations, transcriptional targets, and functional effects. Likewise, SOX4 is active during fetal liver development and epithelial progenitor fate specification towards cholangiocytes[93,105]. In HCC, SOX4 predicts worsened survival, and ectopic Sox4 overexpression *in vivo* decreases hepatocyte identity features and drives metaplasia[164–166]. However, SOX4's involvement in adult hepatocytes' metabolic regulation and adaptations has been less explored.

Our work supports SOX4's involvement in mature hepatocytes' stress responses, with especially strong links to downregulation of canonical hepatocyte functions that align with the posited de-differentiation interpretation of hepatocytes' stress adaptations. These results additionally support that even individual regulatory nodes can be sufficient to co-regulate and couple phenotypes with opposing functional associations and temporal trajectories during progressive stress adaptations.

While our human *in vitro* model allowed us to validate the cell-intrinsic effects of different TFs, future work could define their upstream activating signals (e.g., through co-culture or organoid systems enabling dissection of intercellular interactions, biochemical cues, mechanical microenvironments, etc.)[3,167–173]. Our intercellular signaling analyses nominated ligands associated with both MASLD severity and activation of RELB or SOX4 (Supplementary Note 2): LTβ activates both canonical and non-canonical NF-κB signaling, supports successful acute regeneration and mouse survival after partial hepatectomy, and was predicted to drive hepatocytes' Sustained Upregulation program via production by compositionally-enriched T cells[159,161,174,175]. Likewise, TGFB1 activates SOX4[176] and was predicted to drive hepatocytes' Longitudinal Increase and development-associated programs via production from compositionally-enriched macrophages. SOX4 can additionally be activated by WNT signaling[177,178], whose activation was supported at epigenetic, transcriptomic, and proteomic levels.

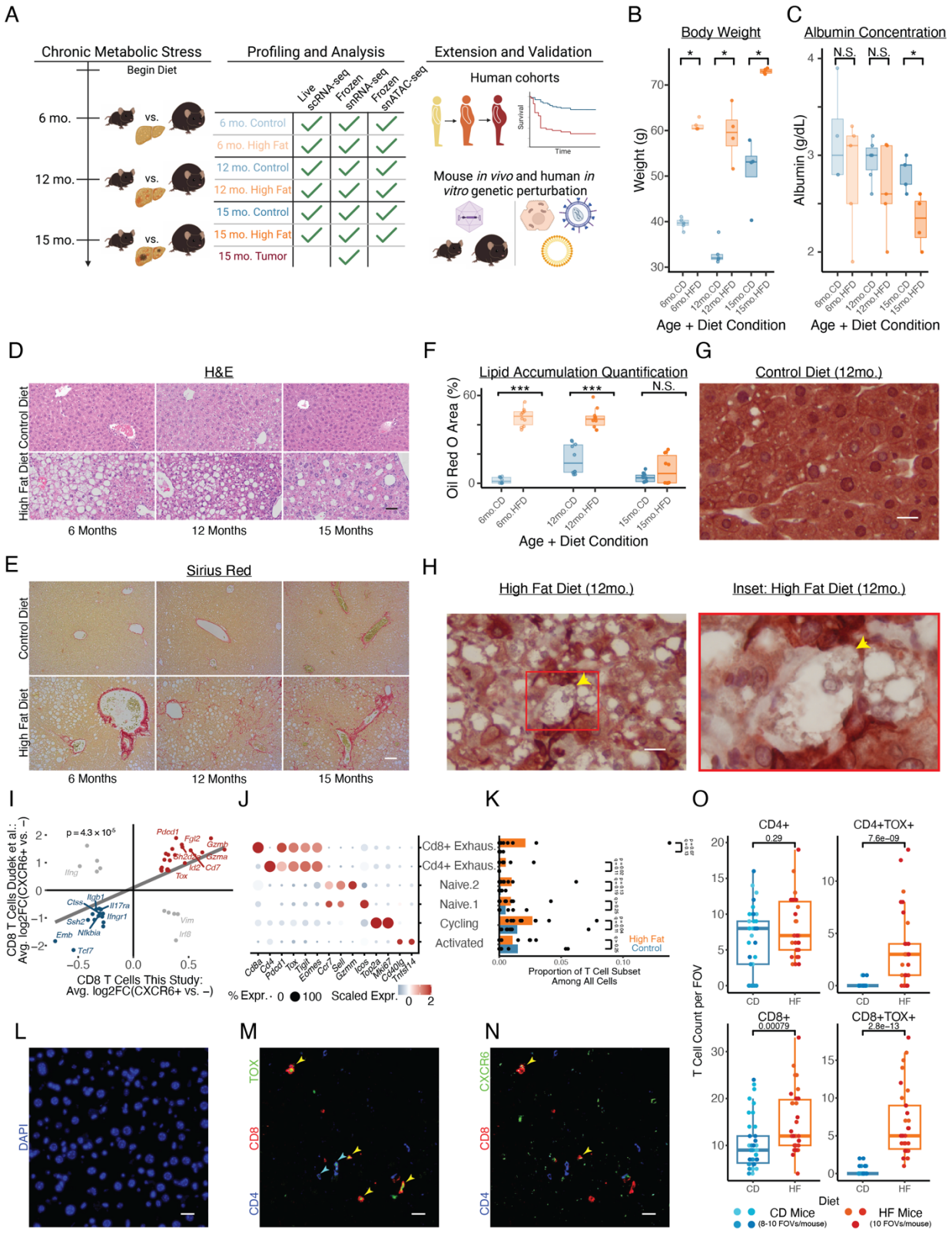
Towards demonstrating beneficial-vs-maladaptative repercussions of dynamic shifts in stress adaptation programs, our analyses highlighted opposing temporal trajectories of ketogenesis and cholesterol synthesis. These pathways compete for the same starting precursor metabolite of acetoacetyl-CoA (product of fatty acid oxidation)[62]. Recent work demonstrated that acetoacetyl-CoA accumulation may increase liver tumorigenesis risk via histone acetylation modifications that facilitate accessible, permissive chromatin landscapes[179]. However, as hepatocytes allocate flux through these metabolic pathways for processing acetoacetyl-CoA, their downstream products have opposing associations with hepatocyte health: free cholesterol can directly cause lipotoxicity[8], while hepatocytes lack the necessary enzyme to catabolize ketone bodies and therefore export them to other organs[180]. Demonstrating maladaptive effects

of ketogenesis decreases during naturally-occurring stress adaptations, metabolically-stressed *Hmgcs2*$^{-/-}$ hepatocytes exhibited accelerated shifts towards phenotypes characteristic of later stages of chronic stress exposure: 1) compensatory increases in cholesterol synthesis enzymes; 2) extreme manifestations of adaptation gene programs; 3) upregulation of HCC markers and development-associated genes including *Sox4*; and 4) downregulation of lineage-determining *Hnf4a* and canonical hepatocyte functions. We previously showed that ketone bodies regulate intestinal stem cell regenerative capacity via epigenetic interactions[181]; future work could further elucidate precise molecular mechanisms and interactions by which ketogenesis, cholesterol synthesis, and associated metabolites regulate the hepatocyte response to metabolic stress.

One final important question is whether stress adaptation programs uncovered in this work are necessarily co-regulated or can be disentangled for selective modulation. Pareto analyses have investigated core archetypes of cellular contributions to collective functions, with hepatocytes accomplishing tissue-scale tasks through division of labor along the periportal-pericentral axis[182–184]. However, these analyses focused on steady-state healthy hepatocytes and did not consider dynamic axes of disease progression. Thus, disease contexts present additional complexity (but also opportunities) for understanding how cells and tissues navigate the state space of potential phenotypes and maintain essential functions despite external stressors. Our work sought to demonstrate the existence and functional implications of chronic stress adaptation programs by focusing on validation of transcriptional (RELB, SOX4) and metabolic (HMGCS2) mediators of wide-ranging dysfunction. However, future work could attempt to decouple these programs, therapeutically activating only beneficial features while mitigating otherwise-linked deleterious phenotypes. For instance, engineered transcriptional activators may enable support of both cellular survival and tissue function without priming of phenotypes associated with worsened cancer outcomes. Alternatively, targeting key enzymes to tune relative metabolic fluxes could provide greater control of tissue homeostatic setpoints, towards buffering of healthy function against environmental stressors.

In this work, we demonstrated how long-term stress drives adaptations that balance immediate cellular survival, homeostatic tissue functions, and long-term dysfunction. Through dissection of hepatocytes' temporal adaptation trajectories, computational methods

development to nominate cell-extrinsic and cell-intrinsic drivers, and experimental validation via human *in vitro* and mouse *in vivo* genetic perturbations, we coalesced diverse axes of hepatocyte adaptation around their specific causal factors. Ultimately, our work provides a foundation for revealing the principles behind cellular and tissue decision-making during stress, translating complex descriptions of disease dysfunction into unifying core mechanisms, and deriving fundamental connections of how even early stress can precipitate cellular adaptations and tradeoffs which lead to long-term dysfunction.



735

Figure 1: Diet-only mouse model of liver adaptations to chronic metabolic stress. (A) Study design schematic. (B) Mouse body weight. (C) Mouse blood albumin concentrations. (D-H) Histology of tissue morphology (D; scalebar=50µm), fibrosis (E; scalebar=50µm), lipid accumulation (F), and hepatocyte ballooning (CK8/18 staining; G-H; scalebar=15µm). (I) $CXCR6^{+}CD8^{+}$ T cell markers between human MASLD patients and our mouse model. (J) Annotated T cell subcluster markers (this study). (K) T cell composition with diet (this study). (L-N) Immunofluorescence of TOX^{+} and $CXCR6^{+}$ T cells in mouse liver tissue (6mo HFD); (L) DAPI, (M) CD4/CD8/TOX, (N) CD4/CD8/CXCR6 (scalebar=50µm). (O) *In situ* T cell enrichment based on TOX status (6mo.HFD, this study). P-value in (I) calculated through Spearman's correlation; all other p-values calculated with Mann-Whitney U test. * indicates $p < 0.05$; ** indicates $p < 0.01$; *** indicates $p < 0.001$.

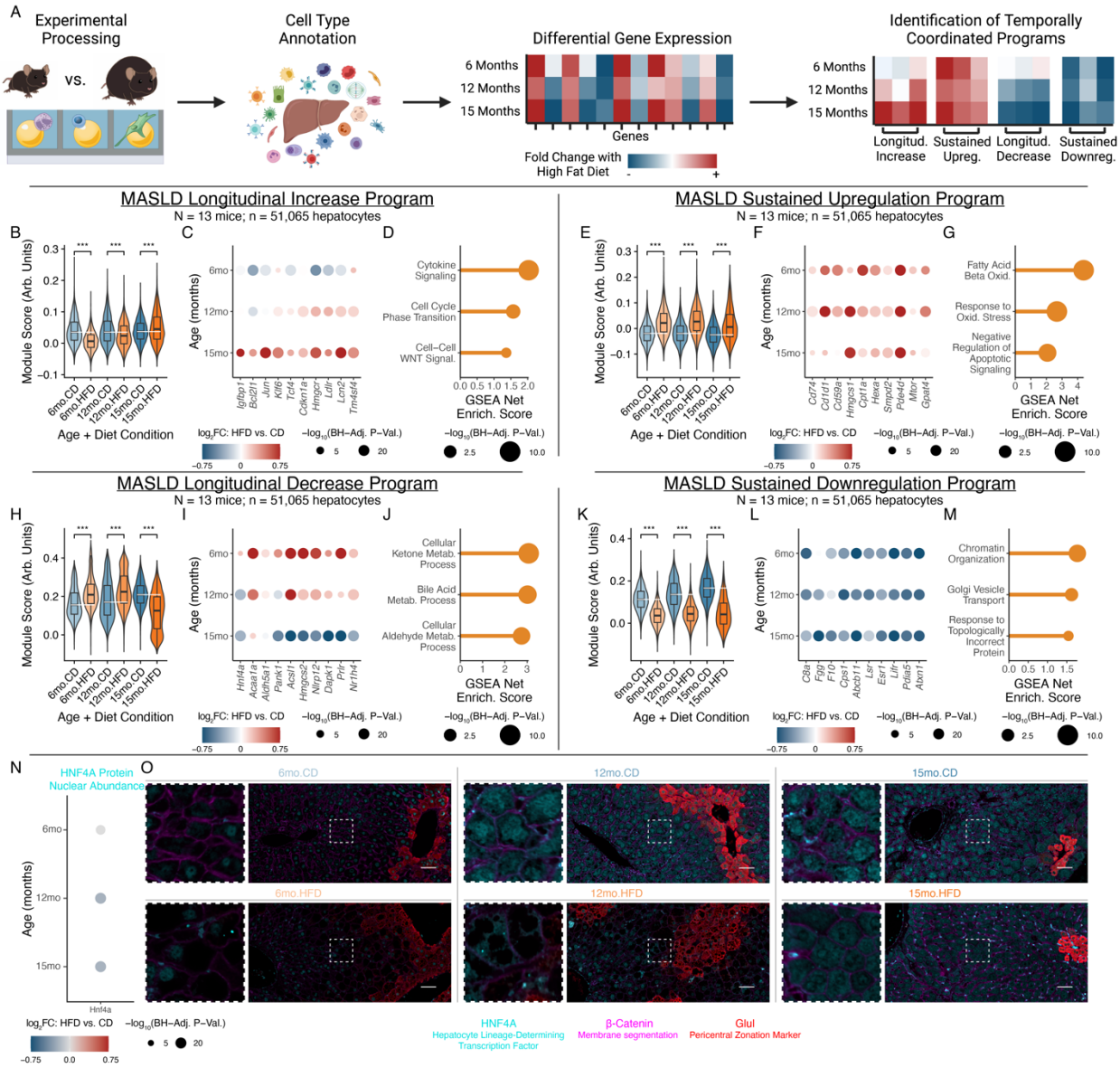


Figure 2: Dynamic adaptations of hepatocytes undergoing chronic metabolic stress. (A) Analysis schematic. (B-D) Longitudinal Increase program, with aggregate expression (B), expression changes of representative genes (C), or enriched GO:BP genesets (D). (E-G) Sustained Upregulation program, following (B-D). (H-J) Longitudinal Decrease program, following (B-D). (K-M) Sustained Downregulation program, following (B-D). (N) HNF4A protein nuclear abundance $\log_2$(fold-change) in hepatocytes with chronic metabolic stress, quantified through *in situ* tissue multiplexed immunofluorescence (N = 18 mice, 3 per age×diet condition; n = 350,842 nuclei). (O) Representative images of HNF4A protein nuclear abundance (scalebar=50μm). GSEA statistical testing through fgsea package; Mann-Whitney U-test used for all other tests; Benjamini-

757 Hochberg correction applied for multiple testing. * indicates $p < 0.05$; ** indicates $p < 0.01$; ***

758 indicates $p < 0.001$.

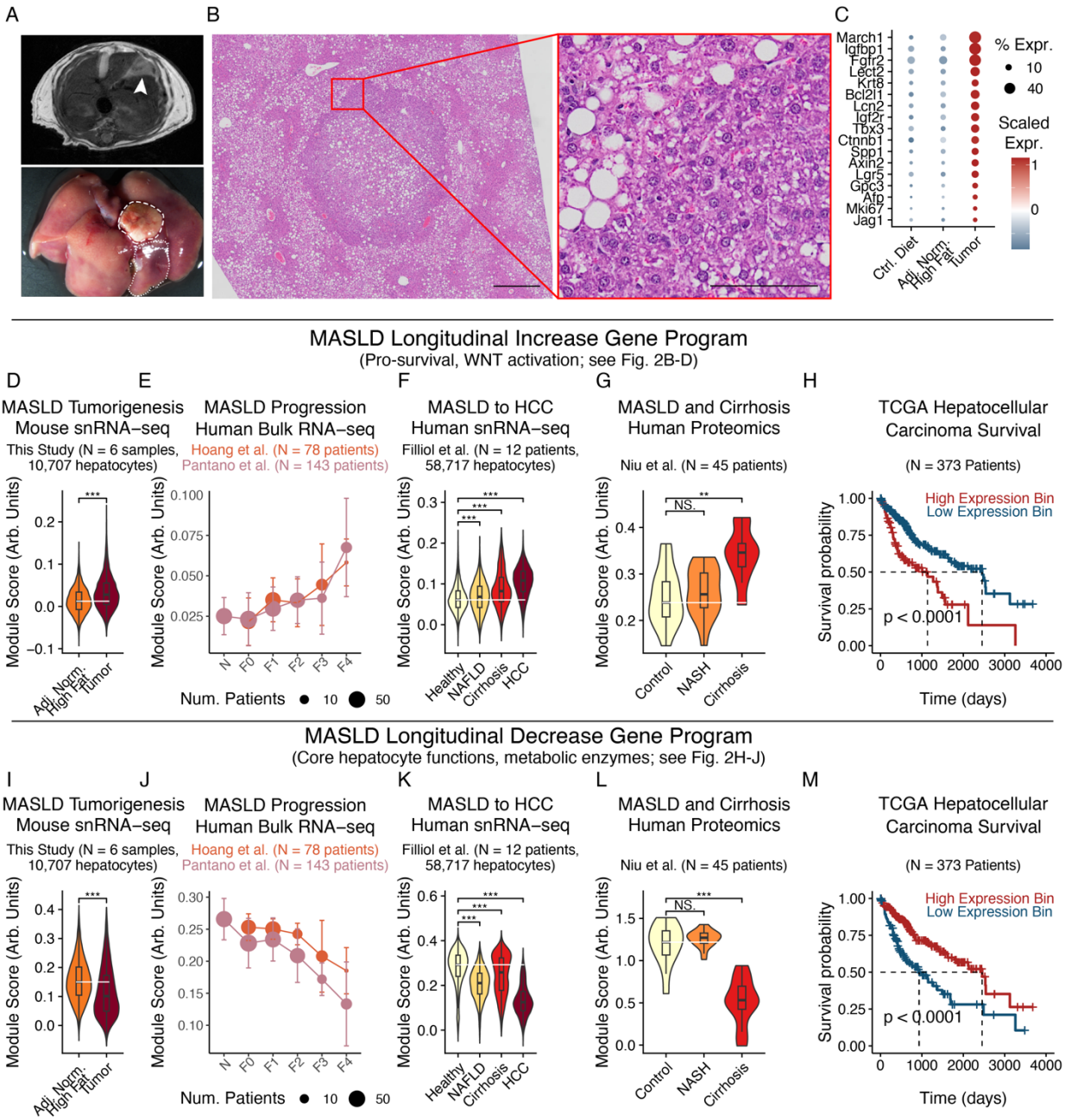


Figure 3: Chronic stress adaptations extend to human cohorts and connect to cancer phenotypes and outcomes. (A) MRI (top) and gross imaging (bottom) of mouse model spontaneous HCC. Dashed line indicates tumor; dotted indicates adjacent normal. (B) H&E of mouse model spontaneous HCC (left scalebar=500µm; inset scalebar=100µm). (C) HCC marker expression in mouse model spontaneous tumors. (D-H) Longitudinal Increase program expression in mouse snRNA-seq tumor cells and adjacent normal hepatocytes (D), human bulk liver RNA-seq (E), human snRNA-seq hepatocytes and tumor cells (F), human bulk liver

767 proteomics (G), or human HCC survival outcomes (H). (I-M) Longitudinal Decrease program
768 expression, following (D-H). Survival outcome p-values calculated with log-rank test; all other p-
769 values calculated using Mann-Whitney U test with Benjamini-Hochberg correction. * indicates p
770 < 0.05; ** indicates p < 0.01; *** indicates p < 0.001.

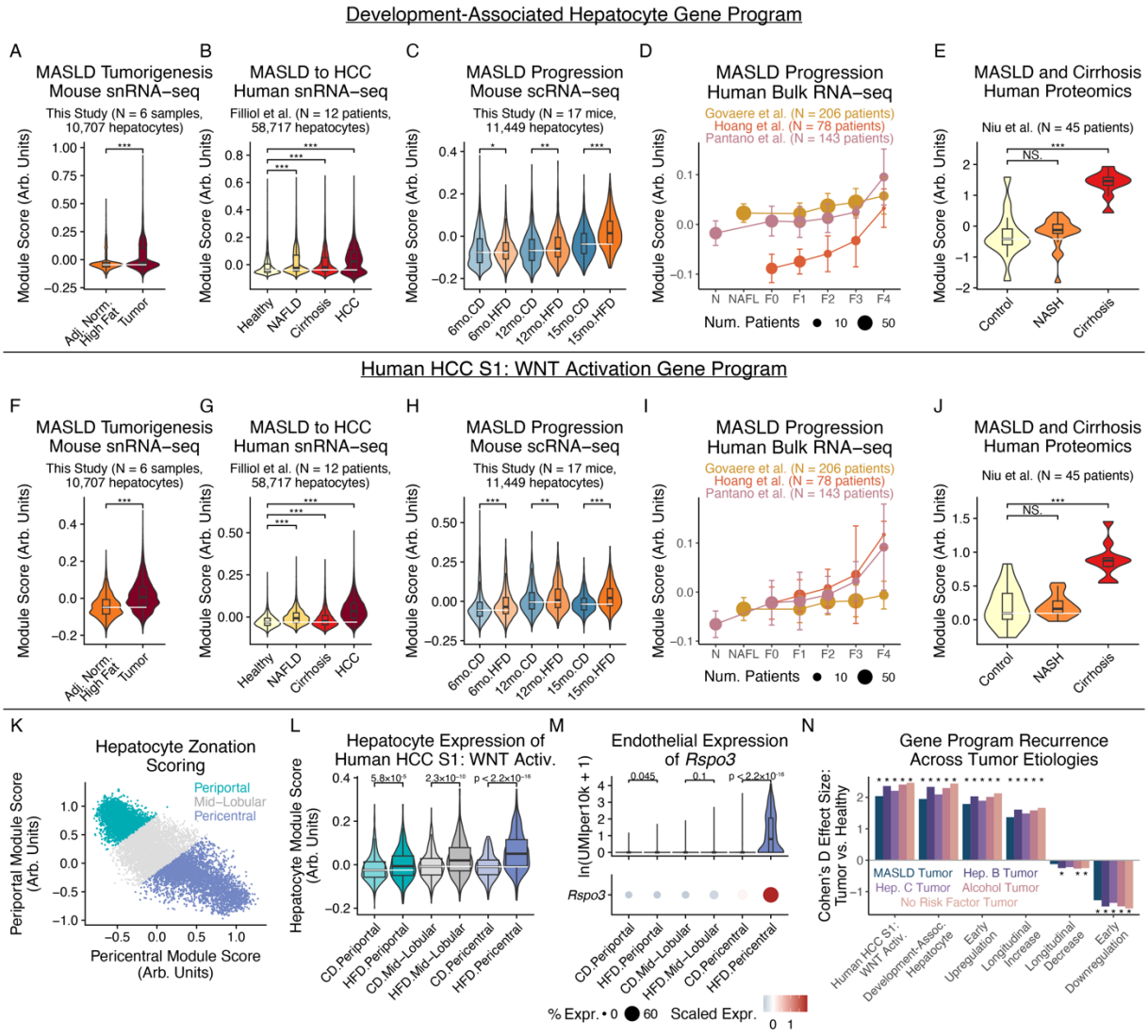


Figure 4: Chronic metabolic stress induces cancer signatures at early stages of MASLD progression. (A-E) Development-associated program expression in mouse snRNA-seq tumor cells and adjacent normal hepatocytes (A), human snRNA-seq hepatocytes and tumor cells (B), mouse snRNA-seq hepatocytes (C), human bulk liver RNA-seq (D), or human bulk liver proteomics (E). (F-J) Human HCC S1 WNT Activation program expression, following (A-E). (K) Mouse hepatocyte zonation scoring and annotation. (L) HCC S1 WNT Activation program expression by hepatocyte zonation and diet. (M) Endothelial *Rspo3* expression by zonation and diet. (N) Tumor-vs-healthy expression differences of this work's hepatocyte adaptation programs, split by tumor etiology. All p-values calculated using Mann-Whitney U test with Benjamini-Hochberg correction. * indicates $p < 0.05$; ** indicates $p < 0.01$; *** indicates $p < 0.001$.

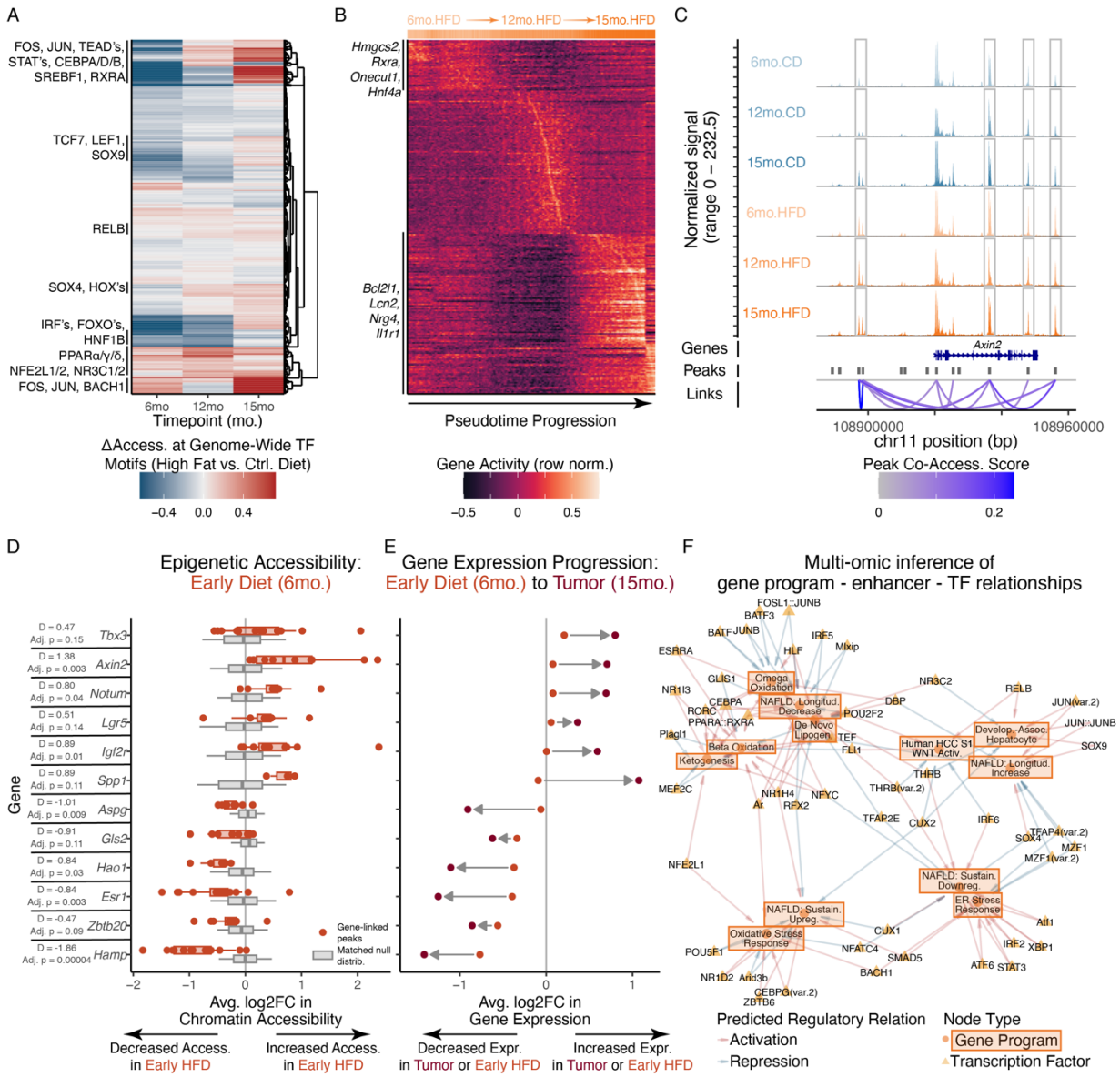


Figure 5: Chronic metabolic stress drives altered epigenetic trajectories and WNT pathway priming. (A) Chromatin accessibility deviation of TF motifs across age and diet conditions. (B) Epigenetic gene activity trajectories of HFD hepatocytes ordered by pseudotime progression. (C) Coverage plot of epigenetic accessibility at *Axin2* gene locus across age and diet. (D) Accessibility changes of gene-linked peaks at 6-month timepoint, between actual observed HFD-vs-CD fold-changes (red) and GC-and-background-matched null distribution (gray). (E) Expression changes of genes at 6-month HFD-vs-CD (red) or 15-month tumor-vs-adjacent-normal (brown) comparisons. (F) Computationally-inferred network of TFs predicted to regulate MASLD-relevant

791 gene programs. All p-values calculated using Mann-Whitney U test with Benjamini-Hochberg
792 correction; effect size quantified through Cohen's D.

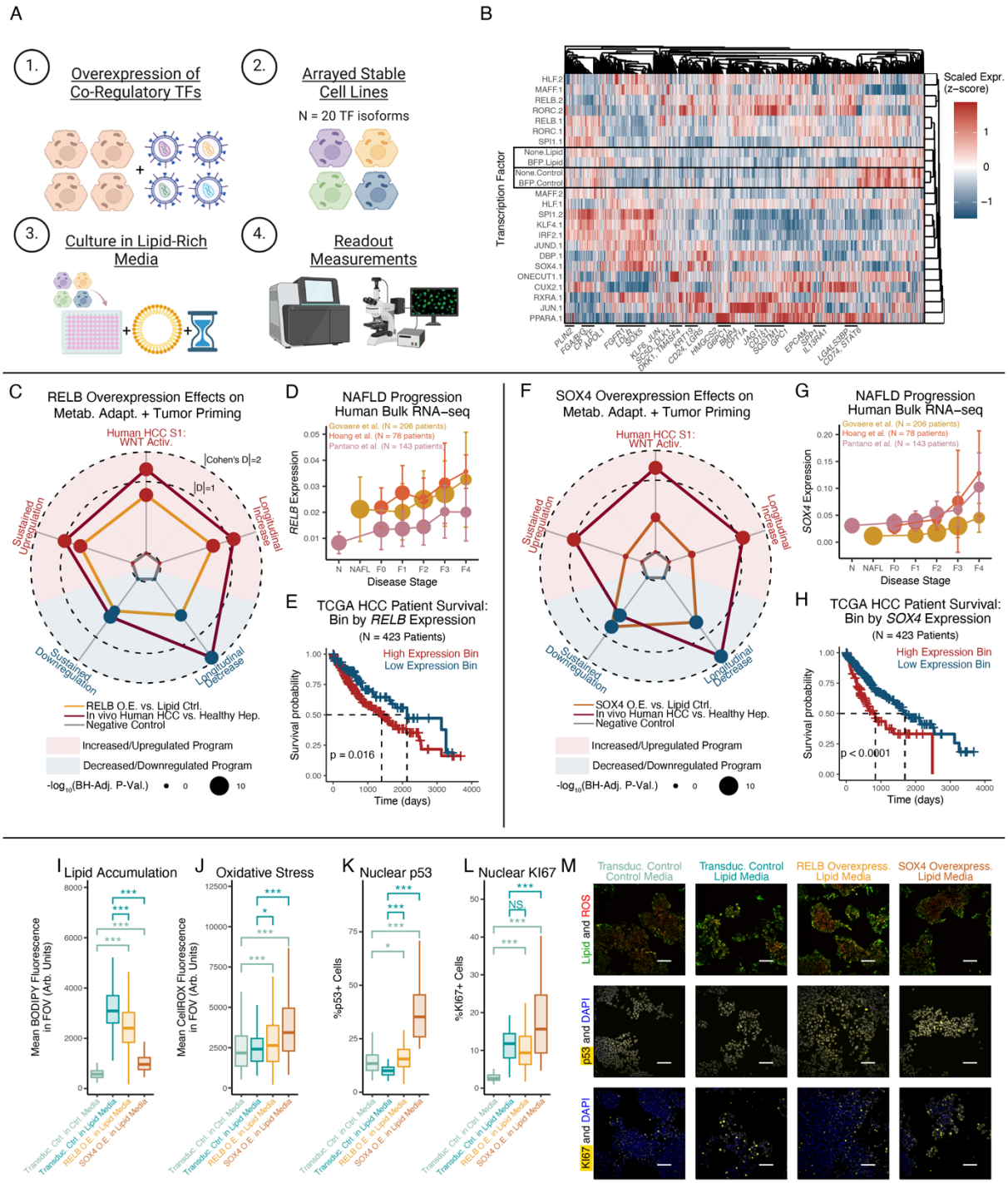


Figure 6: Human *in vitro* validation of RELB and SOX4 as regulators of hepatocyte metabolic (mal)adaptation. (A) Experimental design schematic. (B) Pseudobulked TF expression profiles. Where applicable, ".1" and ".2" indicate isoforms of the same TF gene. (C) RELB's effect sizes on gene program expression. Radius equals |Cohen's D| if concordant directionality with human

MASLD/HCC, and 0 if discordant. (D) RELB expression across human MASLD progression. (E) HCC human survival stratified RELB by expression. (F-H) Transcriptomic regulatory effects of SOX4, following (C-E). (I-L) Lipid accumulation (I; BODIPY 493), ROS accumulation (J; CellROX), nuclear p53 (K), and nuclear KI67 (L). (M) Representative microscopy images supporting (I-L) (scalebar=100µm). Survival outcome p-values calculated with log-rank test; all other p-values calculated using Mann-Whitney U test with Benjamini-Hochberg correction. * indicates $p < 0.05$; ** indicates $p < 0.01$; *** indicates $p < 0.001$.

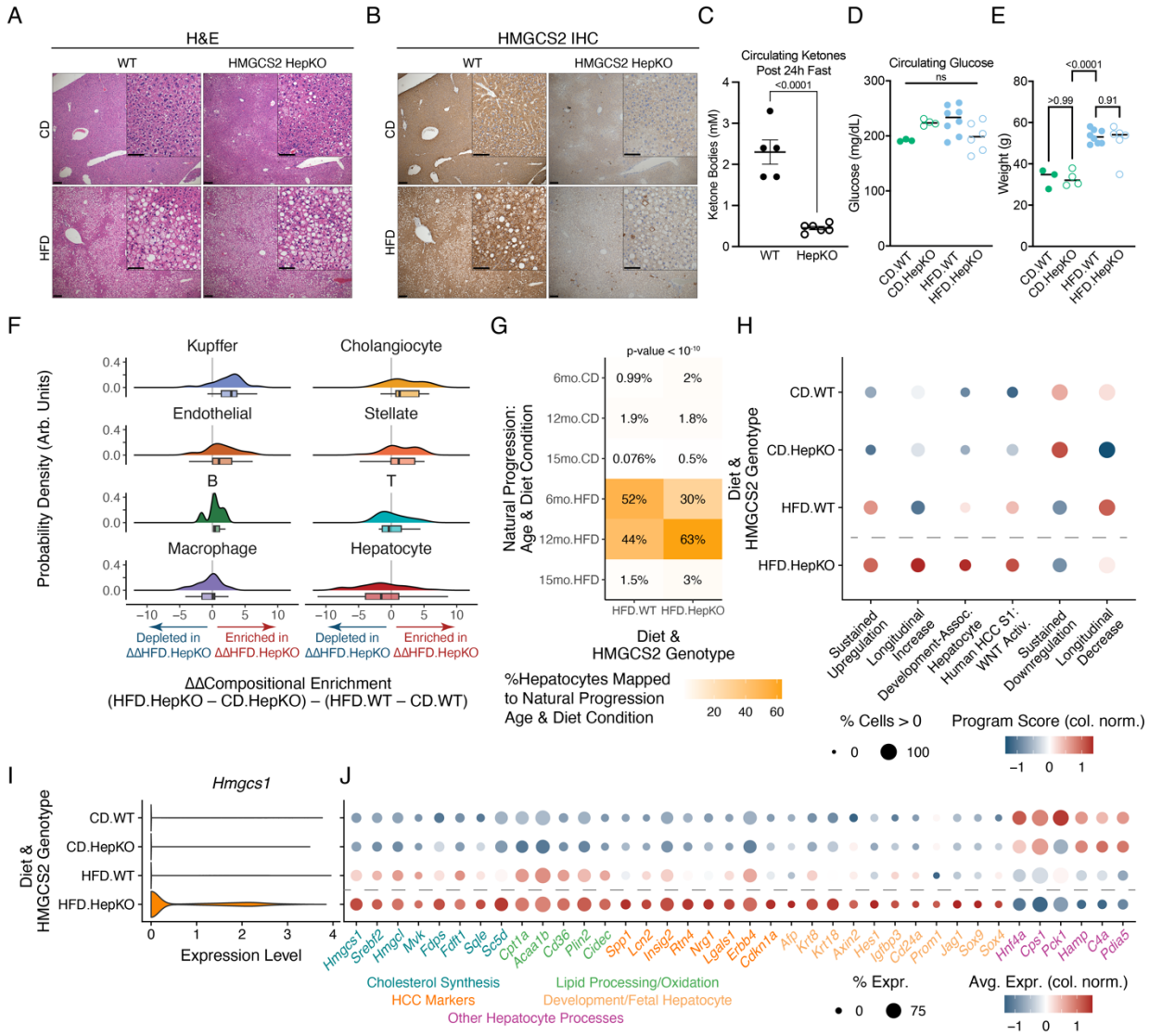


Figure 7: *In vivo* validation of HMGCS2 as a regulator of hepatocyte metabolic (mal)adaptation. (A) H&E staining of wildtype (WT) or *Hmgcs2*$^{fl/fl}$; *Alb-Cre* (HMGCS2 HepKO) mice on CD or HFD (main image scalebar=200µm; inset scalebar=100µm). (B) HMGCS2 immunohistochemistry. (C) Circulating ketone body concentrations after 24-hour fast. (D-E) Circulating glucose concentrations (D) or body weights (E). (F) HMGCS2 HepKO and diet-induced compositional shifts via Milo-derived neighborhoods. (G) Reference mapping of hepatocytes from HMGCS2 HepKO cohort mice to this study's natural stress adaptation progression. (H) Gene program expression with diet and HMGCS2 genotype. (I) Expression of *Hmgcs1* with diet and HMGCS2 genotype. (J) Gene expression across diet and HMGCS2 genotype. P-value in (G) calculated using Fisher's exact test; all other p-values calculated using Student's t-test.

Materials and Methods

Mouse Husbandry and High Fat Diet

C57BL/6 mice in the MIT cohort were housed and cared for in accordance with the American Association for the Accreditation of Laboratory Animal Care and approved by MIT's Committee on Animal Care. A long-term high-fat diet containing 60% kcal from fats (Research Diets D12492) was provided ad libitum to male mice starting at the age of 8-12 weeks for 6-15 months. Sex- and age-matched control mice were provided a purified Control diet containing 10% kcal fats and matched sucrose (Research Diets D12450J). C57BL/6 mice in the second mouse cohort at Brigham and Women's Hospital were housed and cared for in accordance with the American Association for the Accreditation of Laboratory Animal Care and approved by Brigham and Women's Hospital Committee on Animal Care, with diets started at the age of 4 weeks in these mice (Research Diets D12492 and D12450J). *Alb-Cre* mice (#003574) were purchased from Jackson Laboratory and described previously[185]. *Hmgcs2*$^{fl/fl}$ were generated in-house and described previously[181]. The following mice were bred in-house: *Hmgcs2*$^{fl/fl}$; *Alb-Cre*. Studies involving hepatocyte-specific loss of HMGCS2 were performed using littermates, whereby Cre-negative mice served as controls. Live animal imaging was arranged through the Koch Institute at MIT Animal Imaging and Preclinical Testing Core and performed using a Varian 7T MRI imaging system.

Mouse Functional and Histological Characterization

- Cholesterol, ALT, Albumin: For cholesterol, ALT, and albumin measurements, blood was collected through cheek bleed or cardiac puncture in Microvette 200 Z-Gel containers (Sarstedt 20.1291). Briefly, samples were centrifuged immediately after collection, and serum was frozen at -80C before being submitted for IDEXX lab testing as coordinated by the Division of Comparative Medicine at MIT.
- HMGCS2 and CD68 IHC: Tissue was fixed in 10% normal buffered formalin prior to paraffin-embedding. 4-5 micron sections underwent deparaffinization and rehydration prior to antigen retrieval using Borg Decloacker RTU solution (Biocare Medical, BD1000G1) and a pressurized Decloaking Chamber (Biocare Medical, NxGen). Antibodies and respective dilutions used for immunohistochemistry are as follows: rabbit

monoclonal anti-HMGCS2 (1:2000, Abcam, ab137043) and rabbit monoclonal anti-CD68 (1:500, Cell Signaling Technology, #97778) with dilutions performed in Signalstain Antibody Diluent (Cell Signaling Technology, #8112). Biotin conjugated secondary donkey anti-rabbit antibodies (1:500, Jackson ImmunoResearch) were used prior to Vectastain Elite ABC immunoperoxidase detection kit (Vector Laboratories, PK6100). Visualization was performed using Signalstain DAB substrate kit (Cell Signaling Technology, #8049). Counterstaining was performed with 50% Gill No. 1 hematoxylin solution (Sigma-Aldrich, GHS116). Images were acquired using an Olympus BX43 microscope with 4x and 10x objectives (Olympus UPlanSApo). Aperio Digital Slide Scanning at 20x magnification was performed for CD68 slides prior to analysis and quantification using QuPath software.

- Histologic Evaluation: H&E stained sections of liver tissue from 6, 12, and 15-month diet cohorts as well as identified tumor samples were reviewed in a blinded-manner by Dr. Vikram Deshpande and Dr. Ömer Yilmaz, both of whom serve as clinical hepatopathologists in the Department of Pathology at MGH.
- Sirius Red: For collagen staining, 5µm paraffin-deparaffinized sections were hematoxylin stained, washed in running tap water, and then stained with Picro Sirius Red for 1 hour (0.1% Direct Red 80 in picric acid).
- Oil Red O: 10µm frozen sections were briefly fixed, rinsed in isopropranol, and incubated in Oil Red O for 15min with hematoxylin counterstain.
- Glucose tolerance: To measure systemic insulin resistance, intraperitoneal glucose tolerance testing was performed in 6-hours fasted mice as described[186].
- IHC (CK8/18): Staining was performed as described[187] with the following modifications. Antigen retrieval was in buffer TE (pH=9) in a pressure cooker (Biocare Medical Decloaking Chamber, DC2002). A primary antibody already targeting K8/18 was used instead of mixing K8 & K18 1:1 (Progen, GP11; 1:400), detected with ImmPACT AMEC Red per manufacturer instructions (Vector Laboratories, SK-4285), and counterstained with hematoxylin.

874 Multiplexed Immunofluorescence

875 Iterative, multiplexed immunofluorescence was conducted as previously described[188] with the
876 following modifications. Briefly, samples were deparaffinized and antigen retrieved in TE (pH=9)
877 in a pressure cooker (Biocare Medical Decloaking Chamber, DC2002). The following antibodies
878 were used: anti-HNF4A (Abcam, ab201460; 1:400), anti-CXCR6 (Thermo Fisher, PA5-79117;
879 1:125), anti-TOX (Abcam, ab237009; 1:50), anti-CD4 (Abcam, ab183685; 1:500), and anti-CD8A
880 (Abcam, ab217344; 1:500). Imaging was conducted on a confocal Nikon Ti2 microscope with a
881 Yokogawa CSU-W1. Primaries were incubated overnight at 4C in a humidified slide box. For
882 quantitative measurement of nuclear localized HNF4A, Cellpose[189] was used to segment nuclei
883 based on DAPI signal with mean targeted diameter of 9µm (Cellpose optimized for hepatocyte
884 nuclei). Nuclear masks from Cellpose were imported into CellProfiler[190] along with HNF4a
885 channels. Nuclei with area < 20µm^2 were discarded to enrich for hepatocyte nuclei, and then
886 mean fluorescence of nuclear segmentations in the HNF4A channel were computed using the
887 MeasureObjectIntensity module.
888
889 Live-Tissue Dissociation for scRNA-seq
890 Two-step collagenase liver perfusion was performed as previously described[99]. 15,000 cells were
891 loaded onto a Seq-Well array for experimental processing. Samples were sequenced on an
892 Illumina NextSeq 500/500, NextSeq 2000, or NovaSeq 6000.
893
894 Frozen-Tissue Nuclei Isolation for Tandem snRNA-seq and snATAC-seq
895 For each sample, the following solutions and volumes were prepared for isolation of nuclei
896 for tandem Seq-Well snRNA-seq and 10x v1.1 snATAC-seq. All solutions were pre-chilled and kept
897 on ice except when actively handling the tissue or nuclei solution, and then immediately returned
898 to ice.
899 • Base buffer: 100µL of 1M Tris-HCl pH 7.4, 20µL of 5M NaCl, 30µL of 1M $MgCl_2$, 1000µL of
900 10% bovine serum albumin
901 • Wash buffer + RNase inhibitor: 230µL of base buffer, 20µL of 10% Tween-20, 1.7mL of
902 nuclease-free water, 50µL of Sigma-Aldrich Protector

- Wash buffer + digitonin: 230µL of base buffer, 20µL of 10% Tween-20, 1.746mL of nuclease-free water, 4µL of 5% digitonin
- 1X lysis buffer: 230µL of base buffer, 20µL of 10% Tween-20, 20µL of 10% Nonidet P40 substitute, 1.736mL nuclear-free water
- Lysis dilution buffer: 230µL of base buffer, 1.77mL of nuclease-free water
- 0.1X lysis buffer: 200µL of 1X lysis buffer, 1.75mL of lysis dilution buffer, 50µL of Sigma-Aldrich Protector
- Diluted nuclei buffer: 48.75µL of 20X Nuclei Buffer (from 10x v1.1 scATAC-seq reagent kit), 926.25µL of nuclease-free water, 25µL of Sigma-Aldrich Protector
- PBS + 1% BSA + RNase inhibitor: 875µL PBS, 100µL of 10% BSA, 25µL of Sigma-Aldrich Protector

Flash-frozen pieces of liver tissue were kept on dry ice; if needed, a smaller piece of tissue (~2-3mm diameter) was cut using a scalpel on a petri dish on dry ice. The tissue piece was placed into a Miltenyi C tube containing 2 mL of 0.1X lysis buffer, then homogenized using 2 iterations of the m_spleen_01 program on the gentleMACS tissue dissociator. Half of the solution was passed through a 40µm filter pre-wet with 1mL wash buffer + RNase inhibitor, and the other half of the solution was passed through a separate 40µm filter prewet with 1mL wash buffer + digitonin. The C tube was washed with 1mL of wash buffer + RNase inhibitor to capture remnant nuclei stuck to the tube side or lid, and 500µL was transferred to each separate tube. Each solution was transferred to a separate 15mL Falcon tube and centrifuged for 10min at 500g and 4^{oC}, with brake set to 5 (out of a maximum of 10). The supernatant was aspirated, and each pellet was resuspended in 100µL of diluted nuclei buffer. Two 35µm filters were pre-wet with diluted nuclei buffer, and flow-through from the pre-wetting solution was removed to leave the tubes empty; each nuclei solution was then passed through the separate pre-wet 35µm filters using a P200 pipette. Each set of nuclei was counted using a hemocytometer, followed by snRNA-seq and snATAC-seq as previously described:

- 15,000 nuclei in 200µL of PBS + 1% BSA + RNase inhibitor as input to Seq-Well S^3, as described in Hughes*, Wadsworth II*, Gierahn*, et al., *Immunity* (2020)[191]

- 7,000*1.53 nuclei in 5µL of diluted nuclei buffer as input to 10x snATAC-seq, as described in Chromium Next GEM Single Cell ATAC Reagent Kits v1.1 User Guide, CG000209 Rev F

A similar protocol was followed for nuclei isolation from: 1) the frozen tumor and adjacent normal tissue cohort; and, 2) the HMGCS2 HepKO cohort. The following modifications were made:

- As snATAC-seq was not conducted, the splitting of nuclei into a tube containing wash buffer + digitonin and subsequent parallel processing steps were omitted; all nuclei were passed through a single 40um filter pre-wet with 1mL wash buffer + RNase inhibitor.
- At the 35µm filter step, instead of diluted nuclei buffer, the filter was pre-wet with PBS + 1% BSA + RNase inhibitor solution.

Samples were sequenced on an Illumina NextSeq 500/500, NextSeq 2000, or NovaSeq 6000.

Lentiviral Production

Plasmids encoding TF ORFs were obtained as generated in Joung et al., *Cell* (2023) (also deposited in Addgene MORF Collection).

600,000 Lenti-X 293T cells were plated in 2mL of media (DMEM + 10% FBS + 1% penicillin-streptomycin) in a 6-well plate and incubated overnight at 37^{oC} and 5% CO_2 (Day 1). The following afternoon (Day 2), 1µg of psPAX2, 0.33µg of pMD2.G, 1.33µg of ORF plasmid, 5.3µL of P3000 reagent, and 125µL of Opti-MEM were mixed, added to a solution of 6.7µL of L3000 and 125µL of Opti-MEM (not mixed), and incubated for 15min at room temperature. The mixture was added to the Lenti-X 293T cells and incubated overnight. The following morning (Day 3), Lenti-X 293T media was replaced. The following afternoon (Day 4), Lenti-X 293T media was harvested and replaced; the overnight media was passed through a 0.45µm low protein binding filter, combined with Lenti-X Concentrator at a 3:1 media:Lenti-X ratio, and stored overnight at 4^{oC}. The following afternoon (Day 5), Lenti-X 293T media was again harvested and added to the previous day's media and Lenti-X Concentrator solution at 4^{oC} for 30min. The media was spun at 1500g for 45min at 4^{oC}. Supernatant was aspirated, and the pellet was resuspended in 160µL for storage at -80^{oC} in 20µL aliquots.

960 Lentiviral Transduction

961 For each TF ORF, 500,000 HepG2 cells were plated in 2mL of media (Advanced DMEM/F12 +
962 10% FBS + 1% penicillin-streptomycin) in a 6-well plate and incubated overnight at 37^{oC} and 5%
963 CO_2 (Day 1). The following morning (Day 2), lentiviral stocks of each TF ORF and polybrene were
964 thawed to room temperature. Cells' media was replaced with a 10µg/mL solution of polybrene
965 in 2mL of HepG2 media, followed by addition of 8µL of lentivirus to separate HepG2 wells (i.e.,
966 arrayed format; 1 TF ORF per well). Cells were incubated with lentivirus overnight, and media
967 was replaced the following morning (Day 3). The following afternoon (Day 4), media was replaced
968 with a 1ug/mL solution of puromycin in HepG2 media.

969 To produce stable, arrayed HepG2 lines overexpressing each TF, cells were maintained in
970 puromycin-containing media to select for successful transduction, with puromycin-containing
971 media being changed every 2-3 days. Upon expanding to reach confluency in a 6-well plate, each
972 TF ORF-overexpressing HepG2 line was passaged and replated in a 10cm dish. Upon reaching
973 confluency in a 10cm dish, each TF ORF-overexpressing HepG2 line was passaged to freeze down
974 cell stocks (1,000,000 cells in 1mL of 90% FBS + 10% DMSO at -80oC in a Mr. Frosty Freezing
975 Container), followed by subsequent functional and transcriptomic assays in the metabolic stress
976 of lipid-rich media.

977

978 Liver Cell Lipid Culture, scRNA-seq, Functional Imaging, and Immunofluorescence

979 HepG2 cells overexpressing each TF ORF were seeded in a 96-well plate at a density of 4,000
980 cells/well. As lipid-rich, metabolically-stressful media, TF ORF-overexpressing cells were cultured
981 in Advanced DMEM/F12 media containing 10% FBS, 500µM palmitic acid, 100µM oleic acid, and
982 1% penicillin-streptomycin. As controls, BFP-transduced HepG2 cells and non-transduced HepG2
983 cells were each cultured in lipid-rich media or control media (Advanced DMEM/F12 media
984 containing 10% FBS, 600mM BSA control, and 1% penicillin-streptomycin). Each well received
985 200µL of its respective media condition (i.e., TF ORF-overexpressing HepG2's in lipid-rich
986 puromycin-containing media; BFP-transduced HepG2's in lipid-rich or control puromycin-
987 containing media; non-transduced HepG2's in lipid-rich or control puromycin-free media). Media
988 was changed 2 and 4 days after seeding, with assays occurring 7 days after seeding.

For scRNA-seq, HepG2 cells for each condition were seeded across triplicate wells. 7 days after seeding, cells were passaged, and triplicate wells for each condition were pooled and used as input for scRNA-seq using Seq-Well S^3.

For functional imaging, cells were stained with BODIPY (2µM final concentration), CellROX Deep Red (1:500 final dilution), and Hoechst 33342 (5µg/mL final concentration) in PBS at 37^{oC} and 5% CO_2. After 30min incubation, cells were washed 3 times with PBS, followed by imaging on an Opera Phenix at 37^{oC} and 5% CO_2.

After functional imaging, cells were fixed in a solution of 4% paraformaldehyde in PBS for 10 minutes, followed by a 3X wash with PBS. Cells were then permeabilized with 0.1% Tween (KI67 imaging wells) or 0.3% Triton-X (total p53 imaging wells) in PBS with 1% BSA for 10min, followed by overnight incubation at room temperature with the primary antibody (ThermoFisher SolA15 for KI67, Cell Signaling Technology 7F5 for total p53). The following morning, cells were incubated with secondary antibody and imaged on an Opera Phenix.

scRNA-seq and snRNA-seq QC, Filtering, and Annotation

Bcl2fastq was used to convert sequencing reads into bcl files for alignment with either the DropSeq pipeline (live-tissue scRNA-seq samples) or STARsolo (frozen-tissue snRNA-seq and HepG2 scRNA-seq samples). Mouse samples were aligned to the mm10 reference genome, and human samples were aligned to the Hg38 reference genome with the BFP sequence appended (to validate successful transduction via BFP positive control samples).

Cells were first filtered based on number of detected genes, detected UMIs, and percent of mitochondrial counts (Fig. S2-3, S6, S12, S15). The number of principal components was chosen based on an automated elbow-based selection criterion, followed by construction of nearest-neighbor and shared nearest-neighbor graphs and UMAP visualization. Clustering was implemented using the Leiden algorithm, and resolution was chosen based on a parameter scan and maximization of silhouette coefficient. Clusters solely distinguished by quality-associated or doublet-associated features were removed: 1) high expression of mitochondrial genes or well-established markers of low-quality cells (e.g., MALAT1, NEAT); or, 2) co-expression of markers for mutually-exclusive lineages (e.g., immune and epithelial cells). Cells were annotated based on

canonical markers established in prior liver atlases, and clusters corresponding to broad lineages were merged for further subclustering (i.e., epithelial, immune, structural/stromal). Within each lineage, variable gene selection, PCA, nearest-neighbor graph construction, clustering, and UMAP visualization were re-performed; clusters distinguished by markers of other lineages were removed as doublets. Lineage-specific subclustering and doublet removal was repeated 2-3 times within each lineage to ensure robust heterotypic doublet identification.

snATAC-seq QC, Filtering, and Annotation

CellRanger was used for processing and alignment to mm10-2020-A_arc_v2.0.0. Cells were filtered based on number of detected peaks, percent of reads in peaks, nucleosome signal, and TSS enrichment. The number of latent semantic indexing components was chosen based on an automated elbow-based selection, followed by construction of nearest-neighbor and shared nearest-neighbor graphs and UMAP visualization. Iterative subclustering and doublet removal were implemented as described in "scRNA-seq and snRNA-seq QC, Filtering, and Annotation", using chromatin-based gene activity scores instead of canonical marker genes.

Metabolic Adaptation Gene Program Derivation and Driver Gene Identification

At each timepoint for each of the scRNA-seq and snRNA-seq hepatocyte datasets, the average $\log_2$(fold-change) was calculated across all genes detected in at least 25% of cells (chosen based on differential expression benchmarking analyses examining robust parameter estimates and statistical outputs as a function of average expression and detection rate[192]). To promote prioritization of generalizable gene programs across species, we additionally incorporated differential expression information from Govaere et al.'s human bulk RNA-seq cohort, using limma-trend to model gene expression as a function of MASLD stage. See below for qualitative descriptions of the temporal patterns and trajectories captured by each gene program (defined quantitatively further below):

- Sustained Upregulation program: Genes whose elevation is maintained over time
- Sustained Downregulation program: Genes whose lessening is maintained over time

- Longitudinal Increase program: Genes progressively elevated with long-term chronic stress exposure
- Longitudinal Decrease program: Genes progressively lessened with long-term chronic stress exposure

Different genes are preferentially retained and measured in live-tissue scRNA-seq (nuclear + cytoplasmic mRNA) vs. frozen-tissue scRNA-seq (nuclear mRNA only). As a result, a gene may have a large fold-change in one dataset, but non-detection or sparse detection in the other dataset (in turn causing fold-changes that are inestimable or equal to 0). To maximize insights from both our live-tissue scRNA-seq dataset (higher molecular capture given the retention of cytoplasmic mRNA) and frozen-tissue snRNA-seq dataset (higher hepatocyte abundance), we sought to leverage complementary information from each dataset during gene program derivation. We followed the principle that genes should follow a given temporal trajectory (e.g., Sustained Upregulation) in at least one dataset, while allowing for non-detection/sparse detection in the other dataset (e.g., positive fold-changes in one dataset and non-negative fold-changes in the other dataset). As a result, gene programs were derived by filtering to retain genes matching the general structure shown below:

$$(Condition_{Live\_scRNA}\ OR\ Condition_{Frozen\_snRNA})\ AND\ Condition_{Govaere}$$

See below for each gene program's definitions of $Condition_{Live\_scRNA}$, $Condition_{Frozen\_snRNA}$, and $Condition_{Govaere}$, where log_2FC represents average log_2(fold-change) at the noted timepoint between hepatocytes from high fat vs. control diet mice in live-tissue scRNA-seq or frozen-tissue snRNA-seq datasets, and $\beta_{Govaere\_limmatrend}$ represents the regression coefficient of gene expression as a function of disease stage in Govaere et al.:

- Sustained Upregulation:

  $Condition_{Live\_scRNA}$: Across all timepoints, consistent upregulation in live-tissue scRNA-seq and non-negative fold-changes in frozen-tissue snRNA-seq

  $$log_2FC_{Live\_6mo} > 0\ AND\ log_2FC_{Live\_12mo} > 0\ AND\ log_2FC_{Live\_15mo} > 0\ AND\ log_2FC_{Frozen\_6mo} \geq 0\ AND\ log_2FC_{Frozen\_12mo} \geq 0\ AND\ log_2FC_{Frozen\_15mo} \geq 0$$

  $Condition_{Frozen\_snRNA}$: Across all timepoints, consistent upregulation in frozen-tissue snRNA-seq and non-negative fold-changes in live-tissue scRNA-seq

$$log_2FC_{Frozen\_6mo} > 0\ AND\ log_2FC_{Frozen\_12mo} > 0\ AND\ log_2FC_{Frozen\_15mo} > 0\ AND$$
$$log_2FC_{Live\_6mo} \geq 0\ AND\ log_2FC_{Live\_12mo} \geq 0\ AND\ log_2FC_{Live\_15mo} \geq 0$$

$Condition_{Govaere}$: Increased expression with disease stage

$$\beta_{Govaere\_limmatrend} > 0$$

- Sustained Downregulation program:

  $Condition_{Live\_scRNA}$: Across all timepoints, consistent downregulation in live-tissue scRNA-seq and non-positive fold-changes in frozen-tissue snRNA-seq

$$log_2FC_{Live\_6mo} < 0\ AND\ log_2FC_{Live\_12mo} < 0\ AND\ log_2FC_{Live\_15mo} < 0\ AND$$
$$log_2FC_{Frozen\_6mo} \leq 0\ AND\ log_2FC_{Frozen\_12mo} \leq 0\ AND\ log_2FC_{Frozen\_15mo} \leq 0$$

  $Condition_{Frozen\_snRNA}$: Across all timepoints, consistent downregulation in frozen-tissue snRNA-seq and non-positive fold-changes in live-tissue scRNA-seq

$$log_2FC_{Frozen\_6mo} < 0\ AND\ log_2FC_{Frozen\_12mo} < 0\ AND\ log_2FC_{Frozen\_15mo} < 0\ AND$$
$$log_2FC_{Live\_6mo} \leq 0\ AND\ log_2FC_{Live\_12mo} \leq 0\ AND\ log_2FC_{Live\_15mo} \leq 0$$

  $Condition_{Govaere}$: Decreased expression with disease stage

$$\beta_{Govaere\_limmatrend} < 0$$

- Longitudinal Increase program:

  $Condition_{Live\_scRNA}$: Progressive increases from 6-month to 15-month timepoints in live-tissue scRNA-seq and non-decreases in frozen-tissue snRNA-seq

$$log_2FC_{Live\_15mo} > log_2FC_{Live\_6mo}\ AND\ log_2FC_{Frozen\_15mo} \geq log_2FC_{Frozen\_6mo}$$

  $Condition_{Frozen\_snRNA}$: Progressive increases from 6-month to 15-month timepoints in frozen-tissue snRNA-seq and non-decreases in live-tissue scRNA-seq

$$log_2FC_{Frozen\_15mo} > log_2FC_{Frozen\_6mo}\ AND\ log_2FC_{Live\_15mo} \geq log_2FC_{Live\_6mo}$$

  $Condition_{Govaere}$: Increased expression with disease stage

$$\beta_{Govaere\_limmatrend} > 0$$

- Longitudinal Decrease program:

Condition$_{Live_scRNA}$: Progressive decreases from 6-month to 15-month timepoints in live-tissue scRNA-seq and non-increases in frozen-tissue snRNA-seq

$$log_2FC_{Live\_15mo} < log_2FC_{Live\_6mo}\ AND\ log_2FC_{Frozen\_15mo} \leq log_2FC_{Frozen\_6mo}$$

Condition$_{Frozen_snRNA}$: Progressive decreases from 6-month to 15-month timepoints in frozen-tissue snRNA-seq and non-increases in live-tissue scRNA-seq

$$log_2FC_{Frozen\_15mo} < log_2FC_{Frozen\_6mo}\ AND\ log_2FC_{Live\_15mo} \leq log_2FC_{Live\_6mo}$$

Condition$_{Govaere}$: Decreased expression with disease stage

$$\beta_{Govaere\_limmatrend} < 0$$

We highlight that only the mouse live-tissue scRNA-seq dataset, mouse frozen-tissue snRNA-seq dataset, and Govaere et al. bulk RNA-seq dataset were used for stress adaptation gene program derivation. Importantly, none of the other datasets presented elsewhere in the paper (e.g., Fig. 2P-Q, 3D-M, 6B-H, 7G-J, S3, S6, S10A, S7B-Q, or S12-14) were used as part of gene program derivation, ensuring that these figures' analyses were conducted on "test sets" of entirely external, independent datasets. Module scores for each program and dataset were calculated using AddModuleScore in Seurat, and TCGA patient stratification and visualization were conducted with the survival and survminer packages in R.

Comparison and Identification of Cancer Signature Induction during Stress Adaptation

To evaluate whether and which cancer and developmental phenotypes are mirrored in the spontaneous HCC tumors in our chronic metabolic stress mouse model, we began by conducting differential gene expression testing between tumor cells and hepatocytes in matched adjacent normal tissue. As a broad view of potential cancer phenotypes, we conducted gene set enrichment analysis using the fgsea R package, comparing our model's tumor differential expression against gene sets from: 1) MSigDB's "Chemical and Genetic Perturbations" (3,405 genesets spanning diverse prior studies' mutational and signaling perturbations); 2) mutation-specific mouse liver tumor models[91]; and, 3) liver development and regenerative states[92–96]. Thus, all gene sets considered during gene set enrichment analysis were defined externally to

this study, providing a complementary framework (to the previously-described adaptation program derivation) for uncovering phenotypes accentuated with tumorigenesis in our model.

We then sought to understand which tumorigenesis-linked gene sets were also differentially regulated during hepatocytes' progressive stress adaptations (i.e., before tumorigenesis). Seurat's AddModuleScore was used to score each dataset for leading-edge genes from statistically-significantly enriched cancer and developmental gene sets. Thus, statistically-significant module score differences in mouse tumor-vs-adjacent normal comparisons (Fig. 4A, 4F) reflect the statistically-significant gene set enrichment results on which the gene program is based; all significant differences in other datasets (Fig. 4B-E, 4G-J, S7D-E) represent "test set" results in entirely external datasets and independent extensions and connections to hepatocytes' progressive adaptations to chronic stress across species and cohorts.

Intercellular Signaling Analyses

We inferred potential intercellular signaling proteins that may drive hepatocytes' stress adaptation gene programs by applying NicheNet to our live-tissue scRNA-seq data (which enables elevated proportions of immune cells as compared to frozen-tissue nuclei isolation-based datasets). Ligands were considered if detected in at least 10% of cells across any cell type. For each gene program, potential regulatory ligands were prioritized using NicheNet's predict_ligand_activities. To control for ligands that broadly regulate hepatocyte functions or non-specific computational inference, we derived 100 random gene sets with matched average expression levels to the input gene program of interest; ligands' regulatory potential scores were z-scored against a null background of their regulatory potentials for these random gene sets. Ligands were ranked by these regulatory potential z-scores, and only considered if their z-score was positive; a maximum of 8 ligands were retained for each gene program.

Computational Methods for Longitudinal Epigenetic Alterations and Priming

Differential chromatin accessibility was calculated by pseudobulking snATAC-seq hepatocytes from each mouse, then calculating differential expression within each timepoint between high fat vs. control diet pseudobulked hepatocyte profiles using edgeR. chromVAR

scores were calculated using the JASPAR2020 core motif collection and the Signac wrapper function RunChromVAR. Chromatin peak co-accessibility links were calculated using the Signac wrapper function run_cicero.

For higher-resolution inference of hepatocytes' epigenetic adaptation trajectories, we conducted pseudotime analyses. To avoid confounding effects of diet and time, we derived separate pseudotime trajectories for hepatocytes from high fat and control diet hepatocytes, following concepts in previous analyses on allergic inflammation and stem cell differentiation[150]. Pseudotime trajectories were calculated for each diet condition using chromatin peaks that were differentially accessible with age within each diet condition. To control for varying cell counts across samples, each timepoint was downsampled to equal cell counts (matching the timepoint with the fewest cells). Latent semantic indexing and UMAP visualization (on LSI components 2 to 30) were implemented for each diet condition, and Slingshot was used to find pseudotime trajectories.

For epigenetic priming analyses, we began by linking distal peaks to genes based on co-accessibility with peaks in a gene's promoter or gene body (only considering peaks within 100 kilobases of the gene and with Cicero co-accessibility score greater than 0.1). To identify genes whose putative distal chromatin regulators collectively exhibit differential accessibility, we then compared: 1) the observed distribution of differential accessibility fold-changes of actual co-accessible peaks; against, 2) 50 random background peaksets with matched average accessibility and GC bias. Finally, towards genes that may exhibit evidence of epigenetic priming in hepatocytes, we identified genes where:

1. Early chromatin accessibility changes exhibited the same directionality as transcriptional changes across longitudinal stress adaptation and tumorigenesis
   Epigenetic comparison: 6-month high fat vs. control diet snATAC-seq
   Transcriptional comparison: [15-month tumor vs. adjacent normal snRNA-seq] – [6-month high fat vs. control diet snRNA-seq]
2. Late chromatin accessibility changes (15-month high fat vs. control diet snATAC-seq) exhibited the same directionality as tumorigenesis transcriptional changes (15-month tumor vs. adjacent normal snRNA-seq)

Epigenetic comparison: 15-month high fat vs. control diet snATAC-seq

Transcriptional comparison: 15-month tumor vs. adjacent normal snRNA-seq

3. Tumorigenesis incurred a large shift in gene expression (15-month tumor vs. adjacent normal snRNA-seq)

   Transcriptional comparison: 15-month tumor vs. adjacent normal snRNA-seq

4. Early stress drove large changes in chromatin accessibility (6-month high fat vs. control diet snATAC-seq)

   Epigenetic comparison: 6-month high fat vs. control diet snATAC-seq differential accessibility fold-changes at gene-linked co-accessible peaks, relative to matched random background control

MATCHA: Multiomic Ascertainment of Transcriptional Causality via Hierarchical Association

MATCHA seeks to prioritize transcription factors regulating arbitrary, user-specified gene programs, while leveraging multi-omic information on context-specific regulatory relationships (e.g., cell type- or tissue-specific gene-enhancer regulation).

As inputs, MATCHA accepts: 1) one or more arbitrary gene programs; 2) a TF motif database (e.g., JASPAR 2020); 3) multi-omic sc/snRNA-seq data; and optionally, 4) external datasets with relevance to the user's context (e.g., prior bulk or sc/snRNA-seq atlases).

As outputs, MATCHA provides: 1) prioritization scores of the predicted strength and directionality of a TF's regulatory effect on each arbitrary gene program (along with contributions of each input dataset to the overall prioritization score); 2) a bipartite network of which TFs regulate which gene programs (and in what direction, along with TF and gene program network centrality metrics); and, 3) rankings of which TFs may co-regulate multiple gene programs simultaneously (if multiple gene programs were provided).

Towards inference of gene program – co-accessible enhancer – causal TF triads, MATCHA is based on the principle that robust, strong regulatory relationships should be reflected across multiple -omic layers and across datasets. Therefore, MATCHA follows the following steps:

1. If a particular TF regulates the user-specified gene program(s), then it is plausible that the regulatory relationship should be reflected in accessibility changes at program-associated

chromatin regions containing the TF's motif. Therefore, MATCHA begins by identifying chromatin regions that are co-accessible with a gene's promoter or gene body (based on Cicero), then filters peaks based on genomic distance and co-accessibility strength towards plausible regulatory relationships. For each TF motif, MATCHA creates a program-specific motif score by evaluating accessibility at program-coaccessible peaks containing each TF motif (based on chromVAR). Finally, MATCHA calculates the correlation between the program-specific motif score and transcriptional expression of the gene program itself. In this way, MATCHA connects epigenetic alterations at program-linked, TF motif-containing peaks to transcriptional levels of the gene program.

2. If a particular TF regulates the user-specified gene program(s), then it is plausible that the regulatory relationship should be reflected in concordant changes between TF abundance and gene program transcriptions. Therefore, MATCHA calculates the correlation between expression level of each TF in the user-input TF motif database and transcriptional expression of the gene program, thereby connecting transcriptional TF abundance to transcriptional levels of the gene program.
3. If a particular TF regulates the user-specified gene program(s) robustly, then it is plausible that the regulatory relationship should be reflected not only in a single dataset, but across studies. Therefore, if the user provided multiple input datasets (whether bulk or single-cell), MATCHA calculates similar correlations as described above between gene program transcriptional level and either TF motif accessibility at program-linked peaks or TF transcriptional abundance. In this way, MATCHA connects each TF (motif) to transcriptional levels of the gene program across wide-ranging contexts (e.g., spanning species, disease severities, experimental designs, measurement technologies, etc.).
4. To create a single prioritization score for each TF that incorporates information across -omic measurements and studies, MATCHA aggregates each correlation (as calculated in preceding steps). To account for different correlation distributions across -omic measurement modalities and studies, correlations calculated in each previous step are scaled from -1 (most negative association between TF and gene program, indicative of repression) to 1 (most positive association between TF and gene program, indicative of

activation). Scaled correlations are averaged, and TFs are ranked by the average of scaled correlations.

5. In addition to identifying which TFs regulate a particular gene program, it is important to understand whether TFs may co-regulate multiple gene programs, towards essential couplings of cellular phenotypes encapsulated by different programs or potential opportunities to disentangle otherwise-associated phenotypes. Therefore, if the user provided multiple input gene programs, MATCHA will create rankings of each TF's cross-dataset association with each gene program (as described in preceding steps), and calculate network centrality metrics for TFs and gene programs (e.g., out-degree for TFs to identify the number of gene programs that they strongly regulate, in-degree for gene programs to identify the number of TFs potentially regulating them). For concise summarization, MATCHA will filter to retain TFs that are ranked within the top TFs in at least a minimum number of gene programs (Fig. 5F generated with the top 10 TFs for each gene program, showing TFs linked to at least 2 gene programs). Optionally, users can further filter to retain only TFs whose regulatory relationships match observed transcriptional correlations between gene programs (e.g., if a given TF is linked to $program_1$ and $program_2$ which are in turn negatively correlated with each other at the transcriptional level, the TF must activate one of the programs but repress the other). In this way, MATCHA enables nomination and prioritization of TFs with strong regulatory effects on wide-ranging, but specific cellular phenotypes of particular interest to the user (e.g., this work's stress adaptation gene programs, with distinct temporal trends, functional enrichments, and prognostic stratification of human HCC survival).

1274 References

1275 1. Meizlish, M.L., Franklin, R.A., Zhou, X., and Medzhitov, R. (2021). Tissue Homeostasis and
1276 Inflammation. Annu. Rev. Immunol. *39*, 557–581. 10.1146/annurev-immunol-061020-
1277 053734.

1278 2. Ben-Moshe, S., and Itzkovitz, S. (2019). Spatial heterogeneity in the mammalian liver. Nat.
1279 Rev. Gastroenterol. Hepatol. *16*, 395–410. 10.1038/s41575-019-0134-x.

1280 3. Bhatia, S.N., Underhill, G.H., Zaret, K.S., and Fox, I.J. (2014). Cell and tissue engineering for
1281 liver disease. Sci. Transl. Med. *6*, 245sr2-245sr2. 10.1126/scitranslmed.3005975.

1282 4. Michalopoulos, G.K. (2017). Hepatostat: Liver regeneration and normal liver tissue
1283 maintenance. Hepatology. 10.1002/hep.28988.

1284 5. Campana, L., Esser, H., Huch, M., and Forbes, S. (2021). Liver regeneration and inflammation:
1285 from fundamental science to clinical applications. Nat. Rev. Mol. Cell Biol. *22*, 608–624.
1286 10.1038/s41580-021-00373-7.

1287 6. Michalopoulos, G.K., and Bhushan, B. (2021). Liver regeneration: biological and pathological
1288 mechanisms and implications. Nat. Rev. Gastroenterol. Hepatol. *18*, 40–55. 10.1038/s41575-
1289 020-0342-4.

1290 7. Brunt, E.M., Wong, V.W.S., Nobili, V., Day, C.P., Sookoian, S., Maher, J.J., Bugianesi, E., Sirlin,
1291 C.B., Neuschwander-Tetri, B.A., and Rinella, M.E. (2015). Nonalcoholic fatty liver disease.
1292 Nat. Rev. Dis. Primer *1*, 1–22. 10.1038/nrdp.2015.80.

1293 8. Hardy, T., Oakley, F., Anstee, Q.M., and Day, C.P. (2016). Nonalcoholic Fatty Liver Disease:
1294 Pathogenesis and Disease Spectrum. 10.1146/annurev-pathol-012615-044224.

1295 9. Anstee, Q.M., Reeves, H.L., Kotsiliti, E., Govaere, O., and Heikenwalder, M. (2019). From
1296 NASH to HCC: current concepts and future challenges. Nat. Rev. Gastroenterol. Hepatol. *16*,
1297 411–428. 10.1038/s41575-019-0145-7.

1298 10. Younossi, Z., Anstee, Q.M., Marietti, M., Hardy, T., Henry, L., Eslam, M., George, J., and
1299 Bugianesi, E. (2018). Global burden of NAFLD and NASH: trends, predictions, risk factors and
1300 prevention. Nat. Rev. Gastroenterol. Hepatol. *15*, 11–20. 10.1038/nrgastro.2017.109.

1301 11. Rinella, M.E., Lazarus, J.V., Ratziu, V., Francque, S.M., Sanyal, A.J., Kanwal, F., Romero,
1302 D., Abdelmalek, M.F., Anstee, Q.M., Arab, J.P., et al. (2023). A multisociety Delphi consensus
1303 statement on new fatty liver disease nomenclature. Hepatology.
1304 10.1097/HEP.0000000000000520.

1305 12. Mu, X., Español-Suñer, R., Mederacke, I., Affò, S., Manco, R., Sempoux, C., Lemaigre,
1306 F.P., Adili, A., Yuan, D., Weber, A., et al. (2015). Hepatocellular carcinoma originates from

hepatocytes and not from the progenitor/biliary compartment. J. Clin. Invest. *125*, 3891–3903. 10.1172/JCI77995.

13. Sia, D., Villanueva, A., Friedman, S.L., and Llovet, J.M. (2017). Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis. Gastroenterology *152*, 745–761. 10.1053/j.gastro.2016.11.048.

14. Simon, T.G., Roelstraete, B., Sharma, R., Khalili, H., Hagström, H., and Ludvigsson, J.F. (2021). Cancer Risk in Patients With Biopsy-Confirmed Nonalcoholic Fatty Liver Disease: A Population-Based Cohort Study. Hepatology *74*, 2410–2423. 10.1002/hep.31845.

15. Lee, J.-S. (2015). The mutational landscape of hepatocellular carcinoma. Clin. Mol. Hepatol. *21*, 220. 10.3350/cmh.2015.21.3.220.

16. Müller, M., Bird, T.G., and Nault, J.-C. (2020). The landscape of gene mutations in cirrhosis and hepatocellular carcinoma. J. Hepatol. *72*, 990–1002. 10.1016/j.jhep.2020.01.019.

17. Zhu, M., Lu, T., Jia, Y., Luo, X., Gopal, P., Li, L., Odewole, M., Renteria, V., Singal, A.G., Jang, Y., et al. (2019). Somatic Mutations Increase Hepatic Clonal Fitness and Regeneration in Chronic Liver Disease. Cell *177*, 608-621.e12. 10.1016/j.cell.2019.03.026.

18. Ng, S.W.K., Rouhani, F.J., Brunner, S.F., Brzozowska, N., Aitken, S.J., Yang, M., Abascal, F., Moore, L., Nikitopoulou, E., Chappell, L., et al. (2021). Convergent somatic mutations in metabolism genes in chronic liver disease 10.1038/s41586-021-03974-6.

19. Poggetto, E.D., Ho, I., Balestrieri, C., Yen, E., Zhang, S., Citron, F., Hayakawa, Y., Valenti, G., Jiang, S., Sapio, L., et al. (2021). Epithelial memory of inflammation limits tissue damage while promoting pancreatic tumorigenesis. *0486*. 10.1126/science.abj0486.

20. Falvo, D.J., Grimont, A., Zumbo, P., Yang, J.L., Osterhoudt, A., Pan, G., Andre, F., Wilkinson, J.E., Dündar, F., Elemento, O., et al. (2022). An epigenetic memory of inflammation controls context-dependent lineage plasticity in the pancreas. bioRxiv. 10.1101/2021.11.01.466807.

21. Mana, M.D., Hussey, A.M., Tzouanas, C.N., Imada, S., Barrera Millan, Y., Bahceci, D., Saiz, D.R., Webb, A.T., Lewis, C.A., Carmeliet, P., et al. (2021). High-fat diet-activated fatty acid oxidation mediates intestinal stemness and tumorigenicity. Cell Rep. *35*, 109212. 10.1016/j.celrep.2021.109212.

22. Beyaz, S., Mana, M.D., Roper, J., Kedrin, D., Saadatpour, A., Hong, S.J., Bauer-Rowe, K.E., Xifaras, M.E., Akkad, A., Arias, E., et al. (2016). High-fat diet enhances stemness and tumorigenicity of intestinal progenitors. Nature *531*, 53–58. 10.1038/nature17173.

23. Naik, S., and Fuchs, E. (2022). Inflammatory memory and tissue adaptation in sickness and in health. Nature *607*, 249–255. 10.1038/s41586-022-04919-3.

24. Naik, S., Larsen, S.B., Gomez, N.C., Alaverdyan, K., Sendoel, A., Yuan, S., Polak, L., Kulukian, A., Chai, S., and Fuchs, E. (2017). Inflammatory memory sensitizes skin epithelial stem cells to tissue damage. Nature *550*, 475–480. 10.1038/nature24271.

25. Gallage, S., Avila, J.E.B., Ramadori, P., Focaccia, E., Rahbari, M., Ali, A., Malek, N.P., Anstee, Q.M., and Heikenwalder, M. (2022). A researcher's guide to preclinical mouse NASH models. Nat. Metab. *4*, 1632–1649. 10.1038/s42255-022-00700-y.

26. Wolf, M.J., Adili, A., Piotrowitz, K., Abdullah, Z., Boege, Y., Stemmer, K., Ringelhan, M., Simonavicius, N., Egger, M., Wohlleber, D., et al. (2014). Metabolic activation of intrahepatic CD8+ T cells and NKT cells causes nonalcoholic steatohepatitis and liver cancer via cross-talk with hepatocytes. Cancer Cell *26*, 549–564. 10.1016/j.ccell.2014.09.003.

27. Li, X., Ramadori, P., Pfister, D., Seehawer, M., Zender, L., and Heikenwalder, M. (2021). The immunological and metabolic landscape in primary and metastatic liver cancer. Nat. Rev. Cancer *21*, 541–557. 10.1038/s41568-021-00383-9.

28. Sharma, S., Le Guillou, D., and Chen, J.Y. (2023). Cellular stress in the pathogenesis of nonalcoholic steatohepatitis and liver fibrosis. Nat. Rev. Gastroenterol. Hepatol. 10.1038/s41575-023-00832-w.

29. Llovet, J.M., Willoughby, C.E., Singal, A.G., Greten, T.F., Heikenwälder, M., El-Serag, H.B., Finn, R.S., and Friedman, S.L. (2023). Nonalcoholic steatohepatitis-related hepatocellular carcinoma: pathogenesis and treatment. Nat. Rev. Gastroenterol. Hepatol. 10.1038/s41575-023-00754-7.

30. Ratziu, V., and Friedman, S.L. (2023). Why Do So Many Nonalcoholic Steatohepatitis Trials Fail? Gastroenterology *165*, 5–10. 10.1053/j.gastro.2020.05.046.

31. Im, Y.R., Hunter, H., de Gracia Hahn, D., Duret, Cheah, Q., Dong, J., Fairey, M., Hjalmarsson, C., Li, A., Lim, H.K., et al. (2021). A Systematic Review of Animal Models of NAFLD Finds High-Fat, High-Fructose Diets Most Closely Resemble Human NAFLD. Hepatology *74*. 10.1002/hep.31897.

32. Dudek, M., Pfister, D., Donakonda, S., Filpe, P., Schneider, A., Laschinger, M., Hartmann, D., Hüser, N., Meiser, P., Bayerl, F., et al. (2021). Auto-aggressive CXCR6+ CD8 T cells cause liver immune pathology in NASH. Nature. 10.1038/s41586-021-03233-8.

33. Pfister, D., Núñez, N.G., Pinyol, R., Govaere, O., Pinter, M., Szydlowska, M., Gupta, R., Qiu, M., Deczkowska, A., Weiner, A., et al. (2021). NASH limits anti-tumour surveillance in immunotherapy-treated HCC. Nature *592*, 450–456. 10.1038/s41586-021-03362-0.

34. Ramachandran, P., Dobie, R., Wilson-Kanamori, J.R., Dora, E.F., Henderson, B.E.P., Luu, N.T., Portman, J.R., Matchett, K.P., Brice, M., Marwick, J.A., et al. (2019). Resolving the fibrotic niche of human liver cirrhosis at single-cell level. Nature *575*, 512–518. 10.1038/s41586-019-1631-3.

35. Leu, J.I.-J., and George, D.L. (2007). Hepatic IGFBP1 is a prosurvival factor that binds to BAK, protects the liver from apoptosis, and antagonizes the proapoptotic actions of p53 at mitochondria. Genes Dev. *21*, 3095–3109. 10.1101/gad.1567107.

36. Warren, C.F.A., Wong-Brown, M.W., and Bowden, N.A. (2019). BCL-2 family isoforms in apoptosis and cancer. Cell Death Dis. *10*, 177. 10.1038/s41419-019-1407-6.

37. Schulien, I., Hockenjos, B., Schmitt-Graeff, A., Perdekamp, M.G., Follo, M., Thimme, R., and Hasselblatt, P. (2019). The transcription factor c-Jun/AP-1 promotes liver fibrosis during non-alcoholic steatohepatitis by regulating Osteopontin expression. Cell Death Differ. *26*, 1688–1699. 10.1038/s41418-018-0239-8.

38. Sirach, E., Bureau, C., Péron, J.M., Pradayrol, L., Vinel, J.P., Buscail, L., and Cordelier, P. (2007). KLF6 transcription factor protects hepatocellular carcinoma-derived cells from apoptosis. Cell Death Differ. *14*, 1202–1210. 10.1038/sj.cdd.4402114.

39. Albrecht, J.H., Meyer, A.H., and Hu, M.Y. (1997). Regulation of cyclin-dependent kinase inhibitor p21WAF1/Cip1/Sdi1 gene expression in hepatic regeneration. Hepatology *25*, 557–563. 10.1002/hep.510250311.

40. Russell, J.O., and Monga, S.P. (2018). Wnt/β-Catenin Signaling in Liver Development, Homeostasis, and Pathobiology. Annu. Rev. Pathol. Mech. Dis. *13*, 351–378. 10.1146/annurev-pathol-020117-044010.

41. Goldstein, J.L., and Brown, M.S. (1984). Progress in understanding the LDL receptor and HMG-CoA reductase, two membrane proteins that regulate the plasma cholesterol. J. Lipid Res. *25*, 1450–1461. 10.1016/S0022-2275(20)34418-7.

42. Schumacher, M.M., and DeBose-Boyd, R.A. (2021). Posttranslational Regulation of HMG CoA Reductase, the Rate-Limiting Enzyme in Synthesis of Cholesterol. Annu. Rev. Biochem. *90*, 659–679. 10.1146/annurev-biochem-081820-101010.

43. Qiu, J., Liu, Z., Da, L., Li, Y., Xuan, H., Lin, Q., Li, F., Wang, Y., Li, Z., and Zhao, M. (2007). Overexpression of the gene for transmembrane 4 superfamily member 4 accelerates liver damage in rats treated with CCl4. J. Hepatol. *46*, 266–275. 10.1016/j.jhep.2006.08.011.

44. Xu, M., Feng, D., Wu, H., Wang, H., Chan, Y., Kolls, J., Borregaard, N., Porse, B., Berger, T., Mak, T.W., et al. (2015). Liver is the major source of elevated serum lipocalin-2 levels after bacterial infection or partial hepatectomy: A critical role for IL-6/STAT3. Hepatology *61*, 692–702. 10.1002/hep.27447.

45. Wirtz, T.H., Saal, A., Bergmann, I., Fischer, P., Heinrichs, D., Brandt, E.F., Koenen, M.T., Djudjaj, S., Schneider, K.M., Boor, P., et al. (2021). Macrophage migration inhibitory factor exerts pro-proliferative and anti-apoptotic effects via CD74 in murine hepatocellular carcinoma. Br. J. Pharmacol. *178*, 4452–4467. 10.1111/bph.15622.

46. Heinrichs, D., Berres, M., Coeuru, M., Knauel, M., Nellen, A., Fischer, P., Philippeit, C., Bucala, R., Trautwein, C., Wasmuth, H.E., et al. (2014). Protective role of macrophage migration inhibitory factor in nonalcoholic steatohepatitis. FASEB J. *28*, 5136–5147. 10.1096/fj.14-256776.

47. Zeissig, S., Peuker, K., Iyer, S., Gensollen, T., Dougan, S.K., Olszak, T., Kaser, A., and Blumberg, R.S. (2017). CD1d-Restricted pathways in hepatocytes control local natural killer T cell homeostasis and hepatic inflammation. Proc. Natl. Acad. Sci. *114*, 10449–10454. 10.1073/pnas.1701428114.

48. Rollins, S.A., and Sims, P.J. (1990). The complement-inhibitory activity of CD59 resides in its capacity to block incorporation of C9 into membrane C5b-9. J. Immunol. *144*, 3478–3483. 10.4049/jimmunol.144.9.3478.

49. Liang, K. (2023). Mitochondrial CPT1A: Insights into structure, function, and basis for drug development. Front. Pharmacol. *14*, 1160440. 10.3389/fphar.2023.1160440.

50. Sapir, A., Tsur, A., Koorman, T., Ching, K., Mishra, P., Bardenheier, A., Podolsky, L., Bening-Abu-Shach, U., Boxem, M., Chou, T.-F., et al. (2014). Controlled sumoylation of the mevalonate pathway enzyme HMGS-1 regulates metabolism during aging. Proc. Natl. Acad. Sci. *111*. 10.1073/pnas.1414748111.

51. Tao, X., He, H., Peng, J., Xu, R., Fu, J., Hu, Y., Li, L., Yang, X., Feng, X., Zhang, C., et al. (2022). Overexpression of PDE4D in mouse liver is sufficient to trigger NAFLD and hypertension in a CD36-TGF-β1 pathway: therapeutic role of roflumilast. Pharmacol. Res. *175*, 106004. 10.1016/j.phrs.2021.106004.

52. Montgomery, M.K., Taddese, A.Z., Bayliss, J., Nie, S., Williamson, N.A., and Watt, M.J. (2021). Hexosaminidase A (HEXA) regulates hepatic sphingolipid and lipoprotein metabolism in mice. FASEB J *35*. 10.1096/fj.202101186R.

53. Matter, M.S., Decaens, T., Andersen, J.B., and Thorgeirsson, S.S. (2014). Targeting the mTOR pathway in hepatocellular carcinoma: Current state and future trends. J. Hepatol. *60*, 855–865. 10.1016/j.jhep.2013.11.031.

54. Wang, K., Li, C., Lin, X., Sun, H., Xu, R., Li, Q., Wei, Y., Li, Y., Qian, J., Liu, C., et al. (2020). Targeting alkaline ceramidase 3 alleviates the severity of nonalcoholic steatohepatitis by reducing oxidative stress. Cell Death Dis. *11*, 28. 10.1038/s41419-019-2214-9.

55. Zhang, C., Cooper, D.E., Grevengoed, T.J., Li, L.O., Klett, E.L., Eaton, J.M., Harris, T.E., and Coleman, R.A. (2014). Glycerol-3-phosphate acyltransferase-4-deficient mice are protected from diet-induced insulin resistance by the enhanced association of mTOR and rictor. Am. J. Physiol.-Endocrinol. Metab. *307*, E305–E315. 10.1152/ajpendo.00034.2014.

1447 56. Li, J., Ning, G., and Duncan, S.A. (2000). Mammalian hepatocyte differentiation requires
1448 the transcription factor HNF-4␣. Genes Dev. *14*, 464–474. 10.1101/gad.14.4.464.

1449 57. Young, P.A., Senkal, C.E., Suchanek, A.L., Grevengoed, T.J., Lin, D.D., Zhao, L., Crunk, A.E.,
1450 Klett, E.L., Füllekrug, J., Obeid, L.M., et al. (2018). Long-chain acyl-CoA synthetase 1 interacts
1451 with key proteins that activate and direct fatty acids into niche hepatic pathways. J. Biol.
1452 Chem. *293*, 16724–16740. 10.1074/jbc.RA118.004049.

1453 58. Baes, M., and Van Veldhoven, P.P. (2016). Hepatic dysfunction in peroxisomal disorders.
1454 Biochim. Biophys. Acta BBA - Mol. Cell Res. *1863*, 956–970. 10.1016/j.bbamcr.2015.09.035.

1455 59. Leonardi, R., Rehg, J.E., Rock, C.O., and Jackowski, S. (2010). Pantothenate Kinase 1 Is
1456 Required to Support the Metabolic Transition from the Fed to the Fasted State. PLoS ONE *5*,
1457 e11107. 10.1371/journal.pone.0011107.

1458 60. Grinberg, M., Stöber, R.M., Edlund, K., Rempel, E., Godoy, P., Reif, R., Widera, A.,
1459 Madjar, K., Schmidt-Heck, W., Marchan, R., et al. (2014). Toxicogenomics directory of
1460 chemically exposed human hepatocytes. Arch. Toxicol. *88*, 2261–2287. 10.1007/s00204-014-
1461 1400-x.

1462 61. Li, H., Toth, E., and Cherrington, N.J. (2018). Alcohol Metabolism in the Progression of
1463 Human Nonalcoholic Steatohepatitis. Toxicol. Sci. *164*, 428–438. 10.1093/toxsci/kfy106.

1464 62. McGarry, J.D., and Foster, D.W. (1969). Ketogenesis and Cholesterol Synthesis in Normal
1465 and Neoplastic Tissues of the Rat. J. Biol. Chem. *244*, 4251–4256. 10.1016/S0021-
1466 9258(17)36409-8.

1467 63. Udden, S.N., Kwak, Y.-T., Godfrey, V., Khan, M.A.W., Khan, S., Loof, N., Peng, L., Zhu, H.,
1468 and Zaki, H. (2019). NLRP12 suppresses hepatocellular carcinoma via downregulation of cJun
1469 N-terminal kinase activation in the hepatocyte. eLife *8*, e40396. 10.7554/eLife.40396.

1470 64. Wu, Y.-H., Chou, T.-F., Young, L., Hsieh, F.-Y., Pan, H.-Y., Mo, S.-T., Brown, S.B., Chen, R.-
1471 H., Kimchi, A., and Lai, M.-Z. (2020). Tumor suppressor death-associated protein kinase 1
1472 inhibits necroptosis by p38 MAPK activation. Cell Death Dis. *11*, 305. 10.1038/s41419-020-
1473 2534-9.

1474 65. DiStefano, J.K. (2020). NAFLD and NASH in Postmenopausal Women: Implications for
1475 Diagnosis and Treatment. Endocrinology *161*, bqaa134. 10.1210/endocr/bqaa134.

1476 66. Jolicoeur, C., Boutin, J.-M., Okamura, H., Raguet, S., Djiane, J., and Kelly, P.A. (1989).
1477 Multiple Regulation of Prolactin Receptor Gene Expression in Rat Liver. Mol. Endocrinol. *3*,
1478 895–900. 10.1210/mend-3-6-895.

1479 67. Intercept Pharmaceuticals Announces REVERSE Phase 3 Study of Obeticholic Acid (OCA)
1480 in Compensated Cirrhosis due to NASH Did Not Meet its Primary Endpoint (2022). Intercept

Pharm. https://ir.interceptpharma.com/news-releases/news-release-details/intercept-pharmaceuticals-announces-reverse-phase-3-study.

68. Zhang, Y., Li, F., Patterson, A.D., Wang, Y., Krausz, K.W., Neale, G., Thomas, S., Nachagari, D., Vogel, P., Vore, M., et al. (2012). Abcb11 Deficiency Induces Cholestasis Coupled to Impaired β-Fatty Acid Oxidation in Mice. J. Biol. Chem. *287*, 24784–24794. 10.1074/jbc.M111.329318.

69. Park, M.-J., D'Alecy, L.G., Anderson, M.A., Basrur, V., Feng, Y., Brady, G.F., Kim, D., Wu, J., Nesvizhskii, A.I., Lahann, J., et al. (2019). Constitutive release of CPS1 in bile and its role as a protective cytokine during acute liver injury. Proc. Natl. Acad. Sci. *116*, 9125–9134. 10.1073/pnas.1822173116.

70. Zhou, Z., Xu, M.-J., and Gao, B. (2016). Hepatocytes: a key cell type for innate immunity. Cell. Mol. Immunol. *13*, 301–315. 10.1038/cmi.2015.97.

71. Yen, F.T., Roitel, O., Bonnard, L., Notet, V., Pratte, D., Stenger, C., Magueur, E., and Bihain, B.E. (2008). Lipolysis Stimulated Lipoprotein Receptor A Novel Molecular Link Between Hyperlipidemia, Weight Gain, and Atherosclerosis in Mice. J. Biol. Chem. *283*, 25650–25659. 10.1074/jbc.M801027200.

72. Tong, X., Gui, H., Jin, F., Heck, B.W., Lin, P., Ma, J., Fondell, J.D., and Tsai, C. (2011). Ataxin-1 and Brother of ataxin-1 are components of the Notch signalling pathway. EMBO Rep. *12*, 428–435. 10.1038/embor.2011.49.

73. Higa, A., Taouji, S., Lhomond, S., Jensen, D., Fernandez-Zapico, M.E., Simpson, J.C., Pasquet, J.-M., Schekman, R., and Chevet, E. (2014). Endoplasmic Reticulum Stress-Activated Transcription Factor ATF6α Requires the Disulfide Isomerase PDIA5 To Modulate Chemoresistance. Mol. Cell. Biol. *34*, 1839–1849. 10.1128/MCB.01484-13.

74. Qiu, S., Vazquez, J.T., Boulger, E., Liu, H., Xue, P., Hussain, M.A., and Wolfe, A. (2017). Hepatic estrogen receptor α is critical for regulation of gluconeogenesis and lipid metabolism in males. Sci. Rep. *7*, 1661. 10.1038/s41598-017-01937-4.

75. Yao, F., Deng, Y., Zhao, Y., Mei, Y., Zhang, Y., Liu, X., Martinez, C., Su, X., Rosato, R.R., Teng, H., et al. (2021). A targetable LIFR−NF-κB−LCN2 axis controls liver tumorigenesis and vulnerability to ferroptosis. Nat. Commun. *12*, 7333. 10.1038/s41467-021-27452-9.

76. Govaere, O., Cockell, S., Tiniakos, D., Queen, R., Younes, R., Vacca, M., Alexander, L., Ravaioli, F., Palmer, J., Petta, S., et al. (2020). Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis. Sci. Transl. Med. *12*, 1–18. 10.1126/scitranslmed.aba4448.

77. Ramos, M.J., Bandiera, L., Menolascina, F., and Andrew, J. (2022). In vitro models for non-alcoholic fatty liver disease: emerging platforms and their applications.

78. Tschaharganeh, D.F., Chen, X., Latzko, P., Malz, M., Gaida, M.M., Felix, K., Ladu, S., Singer, S., Pinna, F., Gretz, N., et al. (2013). Yes-Associated Protein Up-regulates Jagged-1 and Activates the NOTCH Pathway in Human Hepatocellular Carcinoma. Gastroenterology *144*, 1530-1542.e12. 10.1053/j.gastro.2013.02.009.

79. Xu, C., Xu, Z., Zhang, Y., Evert, M., Calvisi, D.F., and Chen, X. (2022). β-Catenin signaling in hepatocellular carcinoma. J. Clin. Invest. *132*, e154515. 10.1172/JCI154515.

80. Lee, C.-W., Tsai, H.-I., Lee, W.-C., Huang, S.-W., Lin, C.-Y., Hsieh, Y.-C., Kuo, T., Chen, C.-W., and Yu, M.-C. (2019). Normal Alpha-Fetoprotein Hepatocellular Carcinoma: Are They Really Normal? J. Clin. Med. *8*, 1736. 10.3390/jcm8101736.

81. Wu, S., Liao, P., Yan, L., Zhao, Q., Xie, Z., Dong, J., and Sun, H. (2021). Correlation of MKI67 with prognosis, immune infiltration, and T cell exhaustion in hepatocellular carcinoma. BMC Gastroenterol. *21*, 416. 10.1186/s12876-021-01984-2.

82. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739–1740. 10.1093/bioinformatics/btr260.

83. Filliol, A., Saito, Y., Nair, A., Dapito, D.H., Yu, L.-X., Ravichandra, A., Bhattacharjee, S., Affo, S., Fujiwara, N., Su, H., et al. (2022). Opposing roles of hepatic stellate cell subpopulations in hepatocarcinogenesis. Nature. 10.1038/s41586-022-05289-6.

84. Lefebvre, P., Lalloyer, F., Baugé, E., Pawlak, M., Gheeraert, C., Dehondt, H., Vanhoutte, J., Woitrain, E., Hennuyer, N., Mazuy, C., et al. (2017). Interspecies NASH disease activity whole-genome profiling identifies a fibrogenic role of PPARα-regulated dermatopontin. JCI Insight *2*, e92264. 10.1172/jci.insight.92264.

85. Hoang, S.A., Oseini, A., Feaver, R.E., Cole, B.K., Asgharpour, A., Vincent, R., Siddiqui, M., Lawson, M.J., Day, N.C., Taylor, J.M., et al. (2019). Gene Expression Predicts Histological Severity and Reveals Distinct Molecular Profiles of Nonalcoholic Fatty Liver Disease. Sci. Rep. *9*, 1–14. 10.1038/s41598-019-48746-5.

86. Pantano, L., Agyapong, G., Shen, Y., Zhuo, Z., Fernandez-Albert, F., Rust, W., Knebel, D., Hill, J., Boustany-Kari, C.M., Doerner, J.F., et al. (2021). Molecular characterization and cell type composition deconvolution of fibrosis in NAFLD. Sci. Rep. *11*, 1–14. 10.1038/s41598-021-96966-5.

87. Niu, L., Geyer, P.E., Gupta, R., Santos, A., Meier, F., Doll, S., Albrechtsen, N.J.W., Klein, S., Ortiz, C., Uschner, F.E., et al. (2022). Dynamic human liver proteome atlas reveals functional insights into disease pathways. Mol. Syst. Biol. *18*, 2022.01.28.478194. 10.15252/msb.202210947.

88. Ally, A., Balasundaram, M., Carlsen, R., Chuah, E., Clarke, A., Dhalla, N., Holt, R.A., Jones, S.J.M., Lee, D., Ma, Y., et al. (2017). Comprehensive and Integrative Genomic

1552 Characterization of Hepatocellular Carcinoma. Cell *169*, 1327-1341.e23.
1553 10.1016/j.cell.2017.05.046.

1554 89. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P.,
1555 Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: Tool for the unification of
1556 biology. Nat. Genet. *25*, 25–29. 10.1038/75556.

1557 90. Ebert, D., Feuermann, M., Gaudet, P., Harris, N.L., Hill, D.P., Lee, R., Mi, H., Moxon, S.,
1558 Mungall, C.J., Muruganugan, A., et al. (2023). The Gene Ontology knowledgebase in 2023.
1559 Genetics *224*. 10.1093/genetics/iyad031.

1560 91. Tang, M., Zhao, Y., Zhao, J., Wei, S., Liu, M., Zheng, N., Geng, D., Han, S., Zhang, Y.,
1561 Zhong, G., et al. (2022). Liver cancer heterogeneity modeled by in situ genome editing of
1562 hepatocytes. Sci. Adv. *8*, 1–21. 10.1126/sciadv.abn5683.

1563 92. Hishida, T., Yamamoto, M., Hishida-Nozaki, Y., Shao, C., Huang, L., Wang, C., Shojima, K.,
1564 Xue, Y., Hang, Y., Shokhirev, M., et al. (2022). In vivo partial cellular reprogramming enhances
1565 liver plasticity and regeneration. Cell Rep. *39*, 110730. 10.1016/j.celrep.2022.110730.

1566 93. Yang, L., Wang, W., Qiu, W., Guo, Z., Bi, E., and Xu, C. (2017). A single-cell transcriptomic
1567 analysis reveals precise pathways and regulatory mechanisms underlying hepatoblast
1568 differentiation. Hepatology *66*, 1387–1401. 10.1002/hep.29353.

1569 94. Brazovskaja, A., Gomes, T., Körner, C., He, Z., Schaffer, T., Eckel, J.C., Hänsel, R., Santel,
1570 M., Denecke, T., Dannemann, M., et al. (2021). Cell atlas of the regenerating human liver
1571 after portal vein embolization. bioRxiv, 1–16. 10.1101/2023.06.03.444016.

1572 95. Ma, S., Wang, S., Ye, Y., Ren, J., Chen, R., Li, W., Li, J., Zhao, L., Zhao, Q., Sun, G., et al.
1573 (2022). Heterochronic parabiosis induces stem cell revitalization and systemic rejuvenation
1574 across aged tissues. Cell Stem Cell *29*, 990-1005.e10. 10.1016/j.stem.2022.04.017.

1575 96. Wesley, B.T., Ross, A.D.B., Muraro, D., Miao, Z., Saxton, S., Tomaz, R.A., Morell, C.M.,
1576 Ridley, K., Zacharis, E.D., Petrus-Reurer, S., et al. (2022). Single-cell atlas of human liver
1577 development reveals pathways directing hepatic cell fates. Nat. Cell Biol. *24*, 1487–1498.
1578 10.1038/s41556-022-00989-7.

1579 97. Hoshida, Y., Nijman, S.M.B., Kobayashi, M., Chan, J.A., Brunet, J.-P., Chiang, D.Y.,
1580 Villanueva, A., Newell, P., Ikeda, K., Hashimoto, M., et al. (2009). Integrative Transcriptome
1581 Analysis Reveals Common Molecular Subclasses of Human Hepatocellular Carcinoma. Cancer
1582 Res. *69*, 7385–7392. 10.1158/0008-5472.CAN-09-1089.

1583 98. Ben-Moshe, S., Shapira, Y., Moor, A.E., Manco, R., Veg, T., Bahar Halpern, K., and
1584 Itzkovitz, S. (2019). Spatial sorting enables comprehensive characterization of liver zonation.
1585 Nat. Metab. *1*, 899–911. 10.1038/s42255-019-0109-9.

1586 99. Walesky, C.M., Kolb, K.E., Winston, C.L., Henderson, J., Kruft, B., Fleming, I., Ko, S.,
1587 Monga, S.P., Mueller, F., and Apte, U. (2020). Functional compensation precedes recovery of
1588 tissue mass following acute liver injury. Nat. Commun. 10.1038/s41467-020-19558-3.

1589 100. Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR : inferring
1590 associated accessibility from single-cell epigenomic data. Nat. Methods *14*, 975–978.
1591 10.1038/nmeth.4401.

1592 101. Saelens, W., Cannoodt, R., Todorov, H., and Saeys, Y. (2019). A comparison of single-cell
1593 trajectory inference methods. Nat. Biotechnol. *37*, 547–554. 10.1038/s41587-019-0071-9.

1594 102. Ordovas-Montanes, J., Beyaz, S., Rakoff-Nahoum, S., and Shalek, A.K. (2020).
1595 Distribution and storage of inflammatory memory in barrier tissues. Nat. Rev. Immunol. *20*,
1596 308–320. 10.1038/s41577-019-0263-z.

1597 103. Kowalczyk, W., Romanelli, L., Atkins, M., Hillen, H., Bravo González-Blas, C., Jacobs, J.,
1598 Xie, J., Soheily, S., Verboven, E., Moya, I.M., et al. (2022). Hippo signaling instructs ectopic
1599 but not normal organ growth. Science *378*, eabg3679. 10.1126/science.abg3679.

1600 104. Cox, A.G., Tsomides, A., Yimlamai, D., Hwang, K.L., Miesfeld, J., Galli, G.G., Fowl, B.H.,
1601 Fort, M., Ma, K.Y., Sullivan, M.R., et al. (2018). Yap regulates glucose utilization and sustains
1602 nucleotide synthesis to enable organ growth. EMBO J. *37*. 10.15252/embj.2018100294.

1603 105. Poncy, A., Antoniou, A., Cordi, S., Pierreux, C.E., Jacquemin, P., and Lemaigre, F.P.
1604 (2015). Transcription factors SOX4 and SOX9 cooperatively control development of bile
1605 ducts. Dev. Biol. *404*, 136–148. 10.1016/j.ydbio.2015.05.012.

1606 106. Perugorria, M.J., Olaizola, P., Labiano, I., Esparza-Baquer, A., Marzioni, M., Marin, J.J.G.,
1607 Bujanda, L., and Banales, J.M. (2019). Wnt–β-catenin signalling in liver development, health
1608 and disease. Nat. Rev. Gastroenterol. Hepatol. *16*, 121–136. 10.1038/s41575-018-0075-9.

1609 107. Qiu, L., Yang, Q., Zhao, W., Xing, Y., Li, P., Zhou, X., Ning, H., Shi, R., Gou, S., Chen, Y., et
1610 al. (2022). Dysfunction of the energy sensor NFE2L1 triggers uncontrollable AMPK signaling
1611 and glucose metabolism reprogramming. Cell Death Dis. *13*, 501. 10.1038/s41419-022-
1612 04917-3.

1613 108. Berthier, A., Johanns, M., Zummo, F.P., Lefebvre, P., and Staels, B. (2021). PPARs in liver
1614 physiology. Biochim. Biophys. Acta BBA - Mol. Basis Dis. *1867*, 166097.
1615 10.1016/j.bbadis.2021.166097.

1616 109. Li, B., Cai, S.-Y., and Boyer, J.L. (2021). The role of the retinoid receptor, RAR/RXR
1617 heterodimer, in liver physiology. Biochim. Biophys. Acta BBA - Mol. Basis Dis. *1867*, 166085.
1618 10.1016/j.bbadis.2021.166085.

1619 110. Francque, S.M., Bedossa, P., Ratziu, V., Anstee, Q.M., Bugianesi, E., Sanyal, A.J., Loomba,
1620 R., Harrison, S.A., Balabanska, R., Mateva, L., et al. (2021). A Randomized, Controlled Trial of

1621 the Pan-PPAR Agonist Lanifibranor in NASH. N. Engl. J. Med. *385*, 1547–1558.
1622 10.1056/NEJMoa2036205.

1623 111. Musso, G., Cassader, M., and Gambino, R. (2016). Non-alcoholic steatohepatitis:
1624 Emerging molecular targets and therapeutic strategies. Nat. Rev. Drug Discov. *15*, 249–274.
1625 10.1038/nrd.2015.3.

1626 112. Dubois, V., Staels, B., Lefebvre, P., Verzi, M., and Eeckhoute, J. (2020). Control of Cell
1627 Identity by the Nuclear Receptor HNF4 in Organ Pathophysiology. Cells *9*, 2185.
1628 10.3390/cells9102185.

1629 113. Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the
1630 regulatory epigenome. Nat. Rev. Genet. *20*, 29–35. 10.1038/s41576-018-0089-8.

1631 114. Pliner, H.A., Packer, J.S., McFaline-Figueroa, J.L., Cusanovich, D.A., Daza, R.M.,
1632 Aghamirzaie, D., Srivatsan, S., Qiu, X., Jackson, D., Minkina, A., et al. (2018). Cicero Predicts
1633 cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data. Mol. Cell *71*,
1634 858-871.e8. 10.1016/j.molcel.2018.06.044.

1635 115. Suzuki, S., Venkatesh, D., Kanda, H., Nakayama, A., Hosokawa, H., Lee, E., Miki, T.,
1636 Stockwell, B.R., Yokote, K., Tanaka, T., et al. (2022). GLS2 Is a Tumor Suppressor and a
1637 Regulator of Ferroptosis in Hepatocellular Carcinoma. Cancer Res. *82*, 3209–3222.
1638 10.1158/0008-5472.CAN-21-3914.

1639 116. Xie, Z., Zhang, H., Tsai, W., Zhang, Y., Du, Y., Zhong, J., Szpirer, C., Zhu, M., Cao, X.,
1640 Barton, M.C., et al. (2008). Zinc finger protein ZBTB20 is a key repressor of alpha-fetoprotein
1641 gene transcription in liver. Proc. Natl. Acad. Sci. *105*, 10859–10864.
1642 10.1073/pnas.0800647105.

1643 117. Navid, F., and Colbert, R.A. (2017). Causes and consequences of endoplasmic reticulum
1644 stress in rheumatic disease. Nat. Rev. Rheumatol. *13*, 25–40. 10.1038/nrrheum.2016.192.

1645 118. Younossi, Z.M., Ratziu, V., Loomba, R., Rinella, M., Anstee, Q.M., Goodman, Z., Bedossa,
1646 P., Geier, A., Beckebaum, S., Newsome, P.N., et al. (2019). Obeticholic acid for the treatment
1647 of non-alcoholic steatohepatitis: interim analysis from a multicentre, randomised, placebo-
1648 controlled phase 3 trial. The Lancet *394*, 2184–2196. 10.1016/S0140-6736(19)33041-7.

1649 119. Jin, J., Iakova, P., Breaux, M., Sullivan, E., Jawanmardi, N., Chen, D., Jiang, Y., Medrano,
1650 E.M., and Timchenko, N.A. (2013). Increased Expression of Enzymes of Triglyceride Synthesis
1651 Is Essential for the Development of Hepatic Steatosis. Cell Rep. *3*, 831–843.
1652 10.1016/j.celrep.2013.02.009.

1653 120. Xiao, Y., Kim, M., and Lazar, M.A. (2021). Nuclear receptors and transcriptional
1654 regulation in non-alcoholic fatty liver disease. Mol. Metab. *50*, 101119.
1655 10.1016/j.molmet.2020.101119.

121. Zhang, H., Liu, Y., Wang, L., Li, Z., Zhang, H., Wu, J., Rahman, N., Guo, Y., Li, D., Li, N., et al. (2013). Differential effects of estrogen/androgen on the prevention of nonalcoholic fatty liver disease in the male rat. J. Lipid Res. *54*, 345–357. 10.1194/jlr.M028969.

122. Hall, Z., Chiarugi, D., Charidemou, E., Leslie, J., Scott, E., Pellegrinet, L., Allison, M., Mocciaro, G., Anstee, Q.M., Evan, G.I., et al. (2021). Lipid Remodeling in Hepatocyte Proliferation and Hepatocellular Carcinoma. Hepatology *73*, 1028–1044. 10.1002/hep.31391.

123. Xu, Z., Chen, L., Leung, L., Yen, T.S.B., Lee, C., and Chan, J.Y. (2005). Liver-specific inactivation of the Nrf1 gene in adult mouse leads to nonalcoholic steatohepatitis and hepatic neoplasia. Proc. Natl. Acad. Sci. *102*, 4120–4125. 10.1073/pnas.0500660102.

124. Harrison, S.A., Bashir, M.R., Guy, C.D., Zhou, R., Moylan, C.A., Frias, J.P., Alkhouri, N., Bansal, M.B., Baum, S., Neuschwander-Tetri, B.A., et al. (2019). Resmetirom (MGL-3196) for the treatment of non-alcoholic steatohepatitis: a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. The Lancet *394*, 2012–2024. 10.1016/S0140-6736(19)32517-6.

125. Joung, J., Ma, S., Tay, T., Geiger-Schuller, K.R., Kirchgatterer, P.C., Verdine, V.K., Guo, B., Arias-Garcia, M.A., Allen, W.E., Singh, A., et al. (2023). A transcription factor atlas of directed differentiation. Cell *186*, 209-229.e26. 10.1016/j.cell.2022.11.026.

126. Rakhshandehroo, M., Knoch, B., Müller, M., and Kersten, S. (2010). Peroxisome Proliferator-Activated Receptor Alpha Target Genes. PPAR Res. *2010*, 1–20. 10.1155/2010/612089.

127. Capone, E., Iacobelli, S., and Sala, G. (2021). Role of galectin 3 binding protein in cancer progression: a potential novel therapeutic target. J. Transl. Med. *19*, 405. 10.1186/s12967-021-03085-w.

128. Pierreux, C.E., Vanhorenbeeck, V., Jacquemin, P., Lemaigre, F.P., and Rousseau, G.G. (2004). The Transcription Factor Hepatocyte Nuclear Factor-6/Onecut-1 Controls the Expression of Its Paralog Onecut-3 in Developing Mouse Endoderm. J. Biol. Chem. *279*, 51298–51304. 10.1074/jbc.M409038200.

129. Orlicky, D.J., Libby, A.E., Bales, E.S., McMahan, R.H., Monks, J., La Rosa, F.G., and McManaman, J.L. (2019). Perilipin-2 promotes obesity and progressive fatty liver disease in mice through mechanistically distinct hepatocyte and extra-hepatocyte actions. J. Physiol. *597*, 1565–1584. 10.1113/JP277140.

130. Ji, A.L., Rubin, A.J., Thrane, K., Jiang, S., Reynolds, D.L., Meyers, R.M., Guo, M.G., George, B.M., Mollbrink, A., Bergenstråhle, J., et al. (2020). Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma. Cell *182*, 497-514.e22. 10.1016/j.cell.2020.05.039.

131. Karlsson, M., Zhang, C., Mear, L., Zhong, W., Digre, A., Katona, B., Sjostedt, E., Butler, L., Odeberg, J., Dusart, P., et al. (2021). A single–cell type transcriptomics map of human tissues. Sci. Adv. *7*. 10.1126/sciadv.abh2169.

132. Ruttkay-Nedecky, B., Nejdl, L., Gumulec, J., Zitka, O., Masarik, M., Eckschlager, T., Stiborova, M., Adam, V., and Kizek, R. (2013). The Role of Metallothionein in Oxidative Stress. Int. J. Mol. Sci. *14*, 6044–6066. 10.3390/ijms14036044.

133. Wang, G., Gong, Y., Anderson, J., Sun, D., Minuk, G., Roberts, M.S., and Burczynski, F.J. (2005). Antioxidative function of L-FABP in L-FABP stably transfected Chang liver cells. Hepatology *42*, 871–879. 10.1002/hep.20857.

134. Wang, G., Bonkovsky, H.L., De Lemos, A., and Burczynski, F.J. (2015). Recent insights into the biological functions of liver fatty acid binding protein 1. J. Lipid Res. *56*, 2238–2247. 10.1194/jlr.R056705.

135. Chang, Y.-C., Liu, F.-P., Ma, X., Li, M.-M., Li, R., Li, C.-W., Shi, C.-X., He, J.-S., Li, Z., Lin, Y.-X., et al. (2017). Glutathione S-transferase A1 – a sensitive marker of alcoholic injury on primary hepatocytes. Hum. Exp. Toxicol. *36*, 386–394. 10.1177/0960327116650013.

136. Newberry, E.P., Xie, Y., Lodeiro, C., Solis, R., Moritz, W., Kennedy, S., Barron, L., Onufer, E., Alpini, G., Zhou, T., et al. (2019). Hepatocyte and stellate cell deletion of liver fatty acid binding protein reveals distinct roles in fibrogenic injury. FASEB J. *33*, 4610–4625. 10.1096/fj.201801976R.

137. Cole, S.P.C. (2014). Targeting Multidrug Resistance Protein 1 (MRP1, *ABCC1*): Past, Present, and Future. Annu. Rev. Pharmacol. Toxicol. *54*, 95–117. 10.1146/annurev-pharmtox-011613-135959.

138. Zadoorian, A., Du, X., and Yang, H. (2023). Lipid droplet biogenesis and functions in health and disease. Nat. Rev. Endocrinol. *19*, 443–459. 10.1038/s41574-023-00845-0.

139. Gluchowski, N.L., Becuwe, M., Walther, T.C., and Farese, R.V. (2017). Lipid droplets and liver disease: from basic biology to clinical implications. Nat. Rev. Gastroenterol. Hepatol. *14*, 343–355. 10.1038/nrgastro.2017.32.

140. Kiziltas, S. (2016). Toll-like receptors in pathophysiology of liver diseases. World J. Hepatol. *8*, 1354. 10.4254/wjh.v8.i32.1354.

141. Li, L., Cui, L., Lin, P., Liu, Z., Bao, S., Ma, X., Nan, H., Zhu, W., Cen, J., Mao, Y., et al. (2023). Kupffer-cell-derived IL-6 is repurposed for hepatocyte dedifferentiation via activating progenitor genes from injury-specific enhancers. Cell Stem Cell *30*, 1–17. 10.1016/j.stem.2023.01.009.

142. Oh, S.-H., Swiderska-Syn, M., Jewell, M.L., Premont, R.T., and Diehl, A.M. (2018). Liver regeneration requires Yap1-TGFβ-dependent epithelial-mesenchymal transition in hepatocytes. J. Hepatol. *69*, 359–367. 10.1016/j.jhep.2018.05.008.

143. Lotto, J., Stephan, T.L., and Hoodless, P.A. (2023). Fetal liver development and implications for liver disease pathogenesis. Nat. Rev. Gastroenterol. Hepatol. *20*, 561–581. 10.1038/s41575-023-00775-2.

144. Bai, D.-S., Zhang, C., Chen, P., Jin, S.-J., and Jiang, G.-Q. (2017). The prognostic correlation of AFP level at diagnosis with pathological grade, progression, and survival of patients with hepatocellular carcinoma. Sci. Rep. *7*, 12870. 10.1038/s41598-017-12834-1.

145. Li, P., Wang, S.-S., Liu, H., Li, N., McNutt, M.A., Li, G., and Ding, H.-G. (2011). Elevated serum alpha fetoprotein levels promote pathological progression of hepatocellular carcinoma. World J. Gastroenterol. *17*, 4563. 10.3748/wjg.v17.i41.4563.

146. Yamashita, T., and Wang, X.W. (2013). Cancer stem cells in the development of liver cancer. J. Clin. Invest. *123*, 1911–1918. 10.1172/JCI66024.

147. Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. Cancer Discov. *12*, 31–46. 10.1158/2159-8290.CD-21-1059.

148. Flavahan, W.A., Gaskell, E., and Bernstein, B.E. (2017). Epigenetic plasticity and the hallmarks of cancer. Science *357*, eaal2380. 10.1126/science.aal2380.

149. Larsen, S.B., Cowley, C.J., Sajjath, S.M., Barrows, D., Yang, Y., Carroll, T.S., Larsen, S.B., Cowley, C.J., Sajjath, S.M., Barrows, D., et al. (2021). Establishment, maintenance, and recall of inflammatory memory. Cell Stem Cell, 1–17. 10.1016/j.stem.2021.07.001.

150. Ordovas-Montanes, J., Dwyer, D.F., Nyquist, S.K., Buchheit, K.M., Vukovic, M., Deb, C., Wadsworth, M.H., Hughes, T.K., Kazer, S.W., Yoshimoto, E., et al. (2018). Allergic inflammatory memory in human respiratory epithelial progenitor cells. Nature *560*, 649–654. 10.1038/s41586-018-0449-8.

151. Campbell, J.E., Müller, T.D., Finan, B., DiMarchi, R.D., Tschöp, M.H., and D'Alessio, D.A. (2023). GIPR/GLP-1R dual agonist therapies for diabetes and weight loss—chemistry, physiology, and clinical applications. Cell Metab., S1550413123002693. 10.1016/j.cmet.2023.07.010.

152. Chalasani, N., Younossi, Z., Lavine, J.E., Charlton, M., Cusi, K., Rinella, M., Harrison, S.A., Brunt, E.M., and Sanyal, A.J. (2018). The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. Hepatology *67*, 328–357. 10.1002/hep.29367.

153. Sun, S.-C. (2017). The non-canonical NF-κB pathway in immunity and inflammation. Nat. Rev. Immunol. *17*, 545–558. 10.1038/nri.2017.52.

154. Luedde, T., and Schwabe, R.F. (2011). NF-κB in the liver—linking injury, fibrosis and hepatocellular carcinoma. Nat. Rev. Gastroenterol. Hepatol. *8*, 108–118. 10.1038/nrgastro.2010.213.

155. He, G., Yu, G.-Y., Temkin, V., Ogata, H., Kuntzen, C., Sakurai, T., Sieghart, W., Peck-Radosavljevic, M., Leffert, H.L., and Karin, M. (2010). Hepatocyte IKKβ/NF-κB Inhibits Tumor Promotion and Progression by Preventing Oxidative Stress-Driven STAT3 Activation. Cancer Cell *17*, 286–297. 10.1016/j.ccr.2009.12.048.

156. Vucur, M., Ghallab, A., Schneider, A.T., Adili, A., Cheng, M., Castoldi, M., Singer, M.T., Büttner, V., Keysberg, L.S., Küsgens, L., et al. (2023). Sublethal necroptosis signaling promotes inflammation and liver cancer. Immunity *56*, 1578-1595.e8. 10.1016/j.immuni.2023.05.017.

157. He, G., and Karin, M. (2011). NF-κB and STAT3 – key players in liver inflammation and cancer. Cell Res. *21*, 159–168. 10.1038/cr.2010.183.

158. Kondylis, V., Polykratis, A., Ehlken, H., Ochoa-Callejero, L., Straub, B.K., Krishna-Subramanian, S., Van, T.-M., Curth, H.-M., Heise, N., Weih, F., et al. (2015). NEMO Prevents Steatohepatitis and Hepatocellular Carcinoma by Inhibiting RIPK1 Kinase Activity-Mediated Hepatocyte Apoptosis. Cancer Cell *28*, 582–598. 10.1016/j.ccell.2015.10.001.

159. Haybaeck, J., Zeller, N., Wolf, M.J., Weber, A., Wagner, U., Kurrer, M.O., Bremer, J., Iezzi, G., Graf, R., Clavien, P.-A., et al. (2009). A Lymphotoxin-Driven Pathway to Hepatocellular Carcinoma. Cancer Cell *16*, 295–308. 10.1016/j.ccr.2009.08.021.

160. Maeda, S., Kamata, H., Luo, J.-L., Leffert, H., and Karin, M. (2005). IKKβ Couples Hepatocyte Death to Cytokine-Driven Compensatory Proliferation that Promotes Chemical Hepatocarcinogenesis. Cell *121*, 977–990. 10.1016/j.cell.2005.04.014.

161. Cildir, G., Low, K.C., and Tergaonkar, V. (2016). Noncanonical NF-κB Signaling in Health and Disease. Trends Mol. Med. *22*, 414–429. 10.1016/j.molmed.2016.03.002.

162. Yu, H., Lin, L., Zhang, Z., Zhang, H., and Hu, H. (2020). Targeting NF-κB pathway for the therapy of diseases: mechanism and clinical study. Signal Transduct. Target. Ther. *5*, 209. 10.1038/s41392-020-00312-6.

163. Chen, Q., Lu, X., and Zhang, X. (2020). Noncanonical NF-κB Signaling Pathway in Liver Diseases. J. Clin. Transl. Hepatol. *000*, 1–9. 10.14218/JCTH.2020.00063.

164. Katsuda, T., Sussman, J., Ito, K., Katznelson, A., Yuan, S., Li, J., Merrell, A.J., Takenaka, N., Cure, H., Li, Q., et al. (2023). Physiological reprogramming *in vivo* mediated by Sox4 pioneer factor activity (Cell Biology) 10.1101/2023.02.14.528556.

165. Lee, T.K.-W., Guan, X.-Y., and Ma, S. (2022). Cancer stem cells in hepatocellular carcinoma — from origin to clinical implications. Nat. Rev. Gastroenterol. Hepatol. *19*, 26–44. 10.1038/s41575-021-00508-3.

166. Chen, Z., Huang, L., Wu, Y., Zhai, W., Zhu, P., and Gao, Y. (2016). LncSox4 promotes the self-renewal of liver tumour-initiating cells through Stat3-mediated Sox4 expression. Nat. Commun. *7*, 12598. 10.1038/ncomms12598.

167. Bhatia, S.N., and Ingber, D.E. (2014). Microfluidic organs-on-chips. Nat. Biotechnol. *32*, 760–772. 10.1038/nbt.2989.

168. Stevens, K.R., Scull, M.A., Ramanan, V., Fortin, C.L., Chaturvedi, R.R., Knouse, K.A., Xiao, J.W., Fung, C., Mirabella, T., Chen, A.X., et al. (2017). In situ expansion of engineered human liver tissue in a mouse model of chronic liver disease. Sci. Transl. Med. *9*. 10.1126/scitranslmed.aah5505.

169. Huch, M., Knoblich, J.A., Lutolf, M.P., and Martinez-Arias, A. (2017). The hope and the hype of organoid research. Dev. Camb. *144*, 938–941. 10.1242/dev.150201.

170. Hofer, M., and Lutolf, M.P. (2021). Engineering organoids. Nat. Rev. Mater. *0123456789*. 10.1038/s41578-021-00279-y.

171. Nikolaev, M., Mitrofanova, O., Broguiere, N., Geraldo, S., Dutta, D., Tabata, Y., Elci, B., Brandenberg, N., Kolotuev, I., Gjorevski, N., et al. (2020). Homeostatic mini-intestines through scaffold-guided organoid morphogenesis. Nature *585*, 574–578. 10.1038/s41586-020-2724-8.

172. Mead, B.E., Ordovas-Montanes, J., Braun, A.P., Levy, L.E., Bhargava, P., Szucs, M.J., Ammendolia, D.A., MacMullan, M.A., Yin, X., Hughes, T.K., et al. (2018). Harnessing single-cell genomics to improve the physiological fidelity of organoid-derived cell types. BMC Biol. *16*, 62. 10.1186/s12915-018-0527-2.

173. Mead, B.E., Hattori, K., Levy, L., Imada, S., Goto, N., Vukovic, M., Sze, D., Kummerlowe, C., Matute, J.D., Duan, J., et al. (2022). Screening for modulators of the cellular composition of gut epithelia via organoid models of intestinal stem cell differentiation. Nat. Biomed. Eng. *6*, 476–494. 10.1038/s41551-022-00863-9.

174. Anders, R.A., Subudhi, S.K., Wang, J., Pfeffer, K., and Fu, Y.-X. (2005). Contribution of the Lymphotoxin β Receptor to Liver Regeneration. J. Immunol. *175*, 1295–1300. 10.4049/jimmunol.175.2.1295.

175. Ganeff, C., Remouchamps, C., Boutaffala, L., Benezech, C., Galopin, G., Vandepaer, S., Bouillenne, F., Ormenese, S., Chariot, A., Schneider, P., et al. (2011). Induction of the Alternative NF-κB Pathway by Lymphotoxin αβ (LTαβ) Relies on Internalization of LTβ Receptor. Mol. Cell. Biol. *31*, 4319–4334. 10.1128/MCB.05033-11.

176. Tiwari, N., Tiwari, V.K., Waldmeier, L., Balwierz, P.J., Arnold, P., Pachkov, M., Meyer-Schaller, N., Schübeler, D., van Nimwegen, E., and Christofori, G. (2013). Sox4 Is a Master Regulator of Epithelial-Mesenchymal Transition by Controlling Ezh2 Expression and Epigenetic Reprogramming. Cancer Cell *23*, 768–783. 10.1016/j.ccr.2013.04.020.

177. Sinner, D., Kordich, J.J., Spence, J.R., Opoka, R., Rankin, S., Lin, S.-C.J., Jonatan, D., Zorn, A.M., and Wells, J.M. (2007). Sox17 and Sox4 Differentially Regulate β-Catenin/T-Cell Factor Activity and Proliferation of Colon Carcinoma Cells. Mol. Cell. Biol. *27*, 7802–7815. 10.1128/MCB.02179-06.

178. Saegusa, M., Hashimura, M., and Kuwata, T. (2012). Sox4 functions as a positive regulator of β-catenin signaling through upregulation of TCF4 during morular differentiation of endometrial carcinomas. Lab. Invest. *92*, 511–521. 10.1038/labinvest.2011.196.

179. Assante, G., Chandrasekaran, S., Ng, S., Tourna, A., Chung, C.H., Isse, K.A., Banks, J.L., Soffientini, U., Filippi, C., Dhawan, A., et al. (2022). Acetyl-CoA metabolism drives epigenome change and contributes to carcinogenesis risk in fatty liver disease. Genome Med. *14*, 67. 10.1186/s13073-022-01071-5.

180. Puchalska, P., and Crawford, P.A. (2017). Multi-dimensional Roles of Ketone Bodies in Fuel Metabolism, Signaling, and Therapeutics. Cell Metab. *25*, 262–284. 10.1016/j.cmet.2016.12.022.

181. Cheng, C.W., Biton, M., Haber, A.L., Gunduz, N., Eng, G., Gaynor, L.T., Tripathi, S., Calibasi-Kocal, G., Rickelt, S., Butty, V.L., et al. (2019). Ketone Body Signaling Mediates Intestinal Stem Cell Homeostasis and Adaptation to Diet. Cell *178*, 1115–1131. 10.1016/j.cell.2019.07.048.

182. Adler, M., Korem Kohanim, Y., Tendler, A., Mayo, A., and Alon, U. (2019). Continuum of Gene-Expression Profiles Provides Spatial Division of Labor within a Differentiated Cell Type. Cell Syst. *0*, 43-52.e5. 10.1016/j.cels.2018.12.008.

183. Hart, Y., Sheftel, H., Hausser, J., Szekely, P., Ben-moshe, N.B., Korem, Y., Tendler, A., Mayo, A.E., and Alon, U. (2015). Inferring biological tasks using Pareto analysis of high-dimensional data. *12*. 10.1038/nmeth.3254.

184. Ye, J., and Medzhitov, R. (2019). Control strategies in systemic metabolism. Nat. Metab. *1*, 947–957. 10.1038/s42255-019-0118-8.

185. Postic, C., Shiota, M., Niswender, K.D., Jetton, T.L., Chen, Y., Moates, J.M., Shelton, K.D., Lindner, J., Cherrington, A.D., and Magnuson, M.A. (1999). Dual Roles for Glucokinase in Glucose Homeostasis as Determined by Liver and Pancreatic β Cell-specific Gene Knock-outs Using Cre Recombinase. J. Biol. Chem. *274*, 305–315. 10.1074/jbc.274.1.305.

186. Dinger, K., Mohr, J., Vohlen, C., Hirani, D., Hucklenbruch-Rother, E., Ensenauer, R., Dötsch, J., and Alejandre Alcazar, M.A. (2018). Intraperitoneal Glucose Tolerance Test,

Measurement of Lung Function, and Fixation of the Lung to Study the Impact of Obesity and Impaired Metabolism on Pulmonary Outcomes. J. Vis. Exp., 56685. 10.3791/56685.

187. Lackner, C., Gogg-Kamerer, M., Zatloukal, K., Stumptner, C., Brunt, E.M., and Denk, H. (2008). Ballooned hepatocytes in steatohepatitis: The value of keratin immunohistochemistry for diagnosis. J. Hepatol. *48*, 821–828. 10.1016/j.jhep.2008.01.026.

188. Bolognesi, M.M., Manzoni, M., Scalia, C.R., Zannella, S., Bosisio, F.M., Faretta, M., and Cattoretti, G. Multiplex Staining by Sequential Immunostaining and Antibody Removal on Routine Tissue Sections. J Histochem Cytochem *65*, 431–444. doi.org/10.1369/0022155417719419.

189. Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. Nat. Methods *18*, 100–106. 10.1038/s41592-020-01018-x.

190. McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B.A., Karhohs, K.W., Doan, M., Ding, L., Rafelski, S.M., Thirstrup, D., et al. (2018). CellProfiler 3.0: Next-generation image processing for biology. PLOS Biol. *16*, e2005970. 10.1371/journal.pbio.2005970.

191. Hughes, T.K., Wadsworth, M.H., Gierahn, T.M., Do, T., Weiss, D., Andrade, P.R., Ma, F., de Andrade Silva, B.J., Shao, S., Tsoi, L.C., et al. (2020). Second-Strand Synthesis-Based Massively Parallel scRNA-Seq Reveals Cellular States and Molecular Features of Human Inflammatory Skin Pathologies. Immunity *53*, 878-894.e7. 10.1016/j.immuni.2020.09.015.

192. Soneson, C., and Robinson, M.D. (2018). Bias, robustness and scalability in single-cell differential expression analysis. Nat. Methods *15*, 255–261. 10.1038/nmeth.4612.

Acknowledgements

We thank Maria Alimova in the Center for the Development of Therapeutics, for her expertise with high-throughput imaging. We thank the Swanson Biotechnology Center at the Koch Institute, which encompasses the Flow Cytometry, Histology, and Genomics & Bioinformatics Core facilities (NCI P30-CA14051). We thank the Department of Comparative Medicine for mouse husbandry support. We thank Sven Holder and members of the Hope Babette Tang (1983) Histology Facility for histology support. C.N.T. is supported by a Fannie and John Hertz Foundation fellowship and a National Science Foundation Graduate Research Fellowship (1122374). M.S.S is supported by the National Institutes of Health Grant T32DK007191. J.E.S. is supported by a National Institutes of Health F32 Fellowship (F32DK128872). Ö.H.Y. is supported by National Institutes of Health Grant R01CA245314. W.G. is supported by National Institutes of Health Grants R01DK090311, R01DK105198, R24OD017870. C.N.T., J.E.S, Ö.H.Y., and A.K.S receive support from the MIT Stem Cell Initiative through Fondation MIT.

Competing Interests

A.K.S. reports compensation for consulting and/or SAB membership from Honeycomb Biotechnologies, Cellarity, Ochre Bio, FL86, Relation Therapeutics, Senda Biosciences, IntrECate biotherapeutics, Bio-Rad Laboratories, and Dahlia Biosciences unrelated to this work. C.N.T., M.S.S., J.E.S., Ö.H.Y., W.G., and A.K.S have filed a patent related to this work.

Data and materials availability

All code used for scRNA-seq analysis are accessible at Zenodo under XX. MATCHA is available on Github for download and use at XX. scRNA-seq digital gene expression matrices, metadata, and interactive visualization tools can be found as study XX through the Alexandria Project, a Bill & Melinda Gates Foundation–funded portal (part of the Single Cell Portal hosted by the Broad Institute of MIT and Harvard). FASTQs have been uploaded to GEO at accession number XX.