Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Cell. Author manuscript; available in PMC 2024 April 27.

Published in final edited form as:

Cell. 2023 April 27; 186(9): 1968–1984.e20. doi:10.1016/j.cell.2023.03.014.

Positive selection of somatically mutated clones identifies adaptive pathways in metabolic liver disease

Zixi Wang[1], **Shijia Zhu**[2], **Yuemeng Jia**[1], **Yunguan Wang**[3], **Naoto Kubota**[2], **Naoto Fujiwara**[2], **Ruth Gordillo**[4], **Cheryl Lewis**[5], **Min Zhu**[1], **Tripti Sharma**[1], **Lin Li**[1], **Qiyu Zeng**[1], **Yu-Hsuan Lin**[1], **Meng-Hsiung Hsieh**[1], **Purva Gopal**[6], **Tao Wang**[3], **Matt Hoare**[7], **Peter Campbell**[8], **Yujin Hoshida**[2], **Hao Zhu**[1]

[1]Children's Research Institute, Departments of Pediatrics and Internal Medicine, Center for Regenerative Science and Medicine, Simmons Comprehensive Cancer Center, Children's Research Institute Mouse Genome Engineering Core, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

[2]Liver Tumor Translational Research Program, Simmons Comprehensive Cancer Center, Division of Digestive and Liver Diseases, Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

[3]Quantitative Biomedical Research Center, Department of Population and Data Sciences, Center for the Genetics of Host Defense, University of Texas Southwestern Medical Center, Dallas, TX, 75390, USA.

[4]Touchstone Diabetes Center, Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, 75390, USA.

[5]Tissue Management Shared Resource, Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, TX, 75390, USA.

Correspondence and lead contact: Hao Zhu, Hao.Zhu@utsouthwestern.edu, Phone: (214)648-2850.

Author contributions

Z.W. and H.Z. conceived the project, performed the experiments, and wrote the manuscript.
S.Z., N.F., and Y.H. performed bioinformatic analysis and helped to conceptualize the paper.
N.K. reviewed the human NASH and cirrhosis liver histology.
R.G. performed the plasma assays and liver lipid analysis.
C.L. oversaw tissue sectioning and histologic assays.
Y.W., Y.J., and T.W. performed bioinformatic analysis for mouse experiments.
M.Z. and T.S. generated *Gpam* floxed mice.
L.L., Q.Z., Y.L., and M-H.H. performed animal and cell culture experiments.
P.G. performed the histologic analysis and NASH scoring.
M.H. and P.C. provided conceptual contributions to the paper.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of interests

Y.H. consults for Helio Genomics, Espervita Therapeutics, and Roche Diagnostics. Y.H. is a shareholder of Alentis Therapeutics and Espervita Therapeutics. M.H. consults for Spliceor, is on the speakers panel for Boston Scientific, and has research support from Pfizer and AstraZeneca. H.Z. consults for Alnylam Pharmaceuticals, Jumble Therapeutics, and Chroma Medicines, and serves on the SAB of Ubiquitix. H.Z. has research support from Chroma Medicines. H.Z. owns stock in Ionis and Madrigal Pharmaceuticals. M.H., H.Z., and P.C. consult for FL86 and Flagship Pioneering. M.H. and P.C. are co-inventors on patents on somatic mutants in liver disease, including *ACVR2A* and *GPAM*. Z.W., H.Z., and L.L. are co-inventors on patents on *GPAM, TBX3,* and *SMYD2* siRNAs.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[6]Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

[7]University of Cambridge Department of Medicine, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, UK and University of Cambridge Early Cancer Institute, Hutchison Research Centre, Cambridge Biomedical Campus, Cambridge, CB2 0XZ, UK.

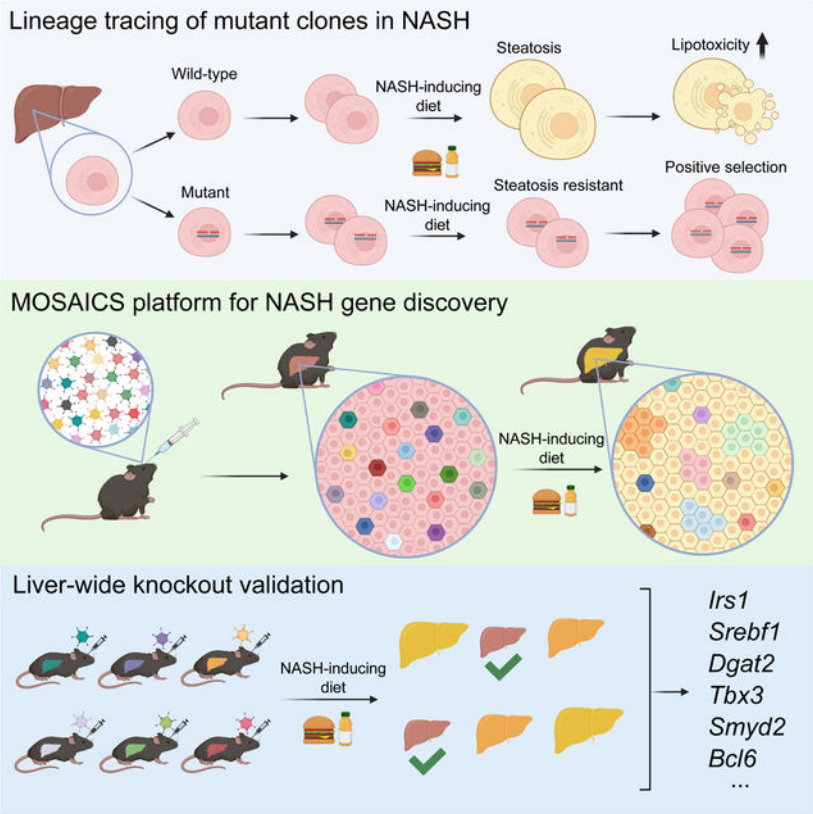
[8]Cancer Genome Project, Wellcome Sanger Institute, Hinxton, Cambridgeshire CB10 1SA, UK.

SUMMARY

Somatic mutations in non-malignant tissues accumulate with age and insult, but whether these mutations are adaptive on the cellular or organismal levels is unclear. To interrogate genes in human metabolic disease, we performed lineage tracing in mice harboring somatic mosaicism subjected to non-alcoholic steatohepatitis (NASH). Proof-of-concept studies with mosaic loss of *Mboat7,* a membrane lipid acyltransferase, showed that increased steatosis accelerated clonal disappearance. Next, we induced pooled mosaicism in 63 known NASH genes, allowing us to trace mutant clones side-by-side. This *in vivo* tracing platform, which we coined MOSAICS, selected for mutations that ameliorate lipotoxicity, including mutant genes identified in human NASH. To prioritize new genes, additional screening of 472 candidates identified 23 somatic perturbations that promoted clonal expansion. In validation studies, liver-wide deletion of *Bcl6, Tbx3,* or *Smyd2* resulted in protection against hepatic steatosis. Selection for clonal fitness in mouse and human livers identifies pathways that regulate metabolic disease.

Graphical Abstract

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



In Brief

An *in vivo* lineage tracing platform was developed to identify somatically mutated clones that are positively selected in chronic liver disease, revealing genes that modify hepatocyte fitness in the context of fatty liver disease in a high throughput manner.

INTRODUCTION

Somatic mutations are common in most individuals, and there is accumulating evidence that mutation burden increases with age and chronic tissue damage[1-4]. While the identity and abundance of these mutations are becoming increasingly understood through deep sequencing, fundamental questions about the relevance of these mutations remain unanswered. The detection of mutant clone expansion, recurrent mutations, or convergent mutations using sequencing provides correlative evidence for increased clonal fitness. However, such fitness increases can be caused by adaptive or pathogenic mechanisms, and it is uncertain if these ever contribute to organ health or function. Even though most somatically mutated clones are not fated to become cancerous, it is possible that increased proliferation/survival could be selfish and have no beneficial effects on tissue function. Therefore, it is unclear how somatic mutations contribute to organismal aging or disease pathogenesis, and whether or not somatic mutations can cause a reversal or adaptation to disease.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Recent evidence from human liver sequencing suggests that mutations could be adaptive. Our previous work indicates that some mutations in cirrhotic livers can result in the regenerative expansion of clones during injury[3], however, it is unclear if these expansion events protect against clinically relevant causes of liver disease. Non-alcoholic steatohepatitis (NASH) is becoming the leading cause of liver disease in the world[5]. NASH is usually conceptualized at the organismal and tissue levels, and less thought has been given to genetic heterogeneity between clones in the liver. In NASH livers, Ng et. al. identified convergent mutations in genes central to insulin signaling and lipogenesis[6]. These loss of function mutations in metabolic enzymes that generate hepatic lipids suggest that some somatic mutations can confer increased fitness through a reversal of the driving etiology of disease.

To understand the impact of somatic mutations at the cellular, tissue, and organismal levels, we developed mouse models that replicate a high density of mutations in the context of common liver diseases. The somatic mutations from human liver tissues were also the most positively selected in mouse models of fatty liver, but were not selected for in the absence of disease. Mechanistically, these mutations mitigated lipotoxic phenotypes, thereby increasing the survival of hepatocyte clones. These findings uncover the biological basis for positive selection of somatic mutations in NASH patient livers. We reasoned that identifying mutant cells with greater fitness than wild-type (WT) cells within diseased environments might nominate therapeutic targets. This encouraged us to explore genes beyond those that are known to be somatically mutated by performing additional *in vivo* CRISPR screens for genes that are dysregulated in chronic liver disease. These screens identified genes that when inhibited, promote liver fitness through the suppression of lipotoxicity. We propose that evolutionary selection in somatically mosaic tissues is a high throughput approach for the identification of adaptive metabolic disease pathways.

RESULTS

The fate of mosaic *Mboat7* knockout clones in the liver is diet-dependent

Fatty liver disease was modeled using a Western Diet (21.1% fat, 41% sucrose, 1.25% cholesterol by weight) supplemented with a high sugar solution (23.1g/L d-fructose and 18.9 g/L d-glucose), a combination hereafter designated as WD. We asked if mosaic mutations in a gene well known to drive NASH could also lead to fitness differences between hepatocytes. *Membrane-bound O-acyltransferase 7* (*Mboat7*) encodes a phospholipid synthetic enzyme identified through GWAS[7,8]. To measure the degree to which liver-wide *Mboat7* deletion promotes fatty liver disease, we used a high dose of AAV8-TBG-Cre to generate liver-wide knockout (KO) mice. We observed efficient deletion of the floxed exon and depletion of the intact *Mboat7* mRNA (Figure S1A,B) after AAV8-TBG-Cre. After 1.5 months of WD feeding, whole-liver *Mboat7* KO vs. WT mice had a dramatic increase in liver/body weight ratios (Figure 1A, Figure S1C,D, and Table S1). H&E showed macroscopic lipid droplets in pericentral hepatocytes of WD fed *Mboat7* KO livers, while only microscopic lipid droplets were observed in WD fed WT livers (Figure 1B). In the context of WD, *Mboat7* deficiency also led to increased transaminitis (Figure 1C,D), liver

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

lipids (Figure 1E,F), and lipidemia (Figure S1E,F). This confirmed that *Mboat7* normally protects against steatosis on the organ level.

To determine if small *Mboat7* KO and WT clones have fitness differences, we lineage traced hepatocytes in mosaic livers. To generate mosaic WT hepatocytes with Tomato reporter activation (control livers) and mosaic *Mboat7* deleted hepatocytes with Tomato reporter activation, we used low dose AAV8-TBG-Cre to achieve recombination in 10-15% of hepatocytes in *Mboat7*$^{+/+}$; *Rosa-Lox-stop-Lox-tdTomato* (hereafter *LSL-tdTomato*) and *Mboat7*$^{f/f}$; *LSL-tdTomato* het mice (Figure 1G). Tomato+ clones were traced under normal chow or WD for 4 months (Figure 1H). Body and liver weights increased with WD as expected (Figure S1G). The initial Tomato labeling percentages in both groups were similar (Figure 1I,J **left column)**. After 4 months of tracing with chow diets, Tomato+ clones increased in both *Mboat7*$^{+/+}$ and *Mboat7*$^{f/f}$ mice similarly (Figure 1I,J **middle column**), indicating a neutral effect of *Mboat7* loss in the liver under non-NASH conditions. After 4 months of tracing with WD, we observed a modest decrease in the percentage of Tomato+ cells in *Mboat7*$^{+/+}$ mice, but a more substantial decrease in *Mboat7*$^{f/f}$ mice (Figure 1I,J **right column**). Comparing WT and *Mboat7* KO clones in the WD setting showed that WT clones survived longer (Figure S1H). The same lineage tracing performed over 6 months showed almost complete disappearance of *Mboat7* KO clones (Figure S1I-K). Therefore, somatic mutations in an important lipid enzyme can have a large impact on clonal fitness in a diet-dependent fashion.

Pooled analysis of clonal evolution in somatically mutated fatty livers

The *Mboat7* observations supported the concept that mutations conferring fitness differences are subject to strong selection pressures within the fatty liver environment. This also showed that clone size provided a surrogate measure for a gene's influence on cellular health, not just proliferation, in metabolic liver disease. To expand our understanding of how mutations influence clonal dynamics in an unbiased fashion, we developed MOSAICS (Method Of Somatic AAV-transposon In vivo Clonal Screening), which is a CRISPR based method to generate pools of heterogeneous mutant cells within tissues such as the liver. MOSAICS is distinct from previous *Fah* KO based regeneration screening systems[9,10] because it is designed to assess a much higher density of mutant clones during homeostasis, and does not require rapid proliferation based selection of FAH expressing clones.

While AAVs are optimal for use in the liver, traditional AAVs cannot be used for screening because they do not genomically integrate, so their sgRNAs cannot be later quantified. The MOSAICS AAV vector carries a U6 promoter driven sgRNA element and a CAG promoter driven Sleeping Beauty 100 transposase (SB100)-P2A-Cre fusion protein (Figure 2A). The entire AAV payload is flanked by transposon inverted repeat (IR) sequences that enable genome integration of the payload mediated by SB100 (Figure 2A), and thus enabled long-term tracing of integrated sgRNAs. Prior studies have combined AAVs with transposons[11,12], but components of MOSAICS were engineered to meet the needs of studying mutations in the liver. We assessed the dose dependent effects of a MOSAICS vector containing an sgRNA against *Pten.* We injected MOSAICS AAV8-*sgPten* into doxycycline (dox)-inducible *TetO-Cas9* mice to generate somatic mutations in the liver,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

or into *LSL-tdTomato* mice to monitor the expression of SB100-P2A-Cre protein, which activates Tomato (Figure 2B). Increasing amounts of AAV could delete *Pten* or activate Tomato in an increasing number of hepatocytes (Figure 2C-E), showing that this system could be titrated to generate either mosaicism or liver-wide gene deletion.

After validation of the MOSAICS system, we aimed to determine if genes known to be important in NASH would have effects on clonal fitness. We generated an sgRNA library against 63 NASH genes including those identified through somatic mutation sequencing (*GPAM, ACVR2A, FOXO1*), GWAS (*MBOAT7, TM6SF2, GCKR*), germline sequencing (*PNPLA3, HSD17B13*), biochemical studies[13], or clinical stage drug targets (*ACC1/2, FXR, ASK1*) (Figure S2A). We generated mosaic livers by injecting the AAV library into Cas9 expressing mice, and then exposed the mice to either chow or WD. Analysis of enriched or depleted guides after a fixed time period enabled us to monitor clonal dynamics in both chow and WD conditions (Figure 3A). By including sgRNAs specifically enriched in WD conditions and excluding sgRNAs enriched in both chow and WD conditions, we were able to identify fatty liver specific fitness promoting mutations and exclude genes whose mutations induce a constitutive, fatty liver independent proliferation. Seven to eight independent mice were used for each group, and the results from individual mice were consistent (Table S2). While liver weight, body weight, and steatosis increased as expected on WD (Figure S2B,C), we did not detect gross liver tumors, suggesting that the somatic mutations did not cause rapid cancer development. We observed early fibrosis in the WD fed livers, indicating that clonal selection occurred in a microenvironment that included pathologic features of NASH (Figure S2D). We compared sgRNA distributions from different screening steps and found the distribution between plasmids/AAVs (the starting points) and chow/WD livers (the endpoints) were very similar (Figure S2E and Table S3), suggesting that most of the sgRNAs in the screen did not change in frequency, which is also true for other CRISPR screens[14], and indicated a lack of genetic drift.

The five genes associated with the highest levels of clonal expansion in the WD but not the chow group were *Acvr2a, Gpam, Dgat2, Srebf1,* and *Irs1* (Figure 3B,C and Table S2); each is important in NASH (Figure 3D). *Dgat2* and *Gpam* are critical enzymes in triglyceride synthesis, and *Dgat2* is currently being inhibited in NASH clinical trials[15,16]. *Srebf1* is a major transcriptional regulator of lipogenesis[17-19]. *Irs1* is an insulin receptor substrate and when lost, leads to suppressed lipogenesis due to reduced insulin signaling[20-22]. Interestingly, *Acvr2a* and *Gpam* are two of the most recurrently mutated genes in cirrhotic tissues from alcoholic liver disease and NASH patients[6]. *Foxo1* gain of function hotspot mutations are also positively selected in human tissues; consistently, *Foxo1* loss of function mutations were negatively selected in the MOSAICS screen (Table S2). To further validate the screening results with a smaller set of genes that allowed for deeper sgRNA coverage, we fate mapped a mini-pool of sgRNAs against *Acvr2a, Irs1, Srebf1,* and *Dgat2* as well as 4 non-targeting sgRNAs (Figure 3E). After 6 months, the control sgRNAs in the WD fed livers decreased in abundance compared to those in chow fed livers, whereas each of the guides against *Acvr2a, Irs1, Srebf1,* and *Dgat2* were positively selected (Figure 3F and Table S2). The depletion of control guides was likely due to the increased turnover of hepatocytes in WD fed conditions, as was also seen for Tomato in Figure 1I (right vs. middle column).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

These data show that multiple mutations that impair lipid accumulation result in a fitness advantage for clones in the NASH environment.

Gpam KO clones are more fit and reduce liver steatosis

We further investigated mutations in *Gpam,* because it appeared as a top hit from the functional NASH screen described above and from human somatic mutation sequencing efforts[6]. GPAM catalyzes the acylation of glycerol-3-phosphate with acyl-coenzyme A (CoA) to generate CoA and lysophosphatidic acid (LPA), the rate-limiting step in triacylglycerol synthesis (Figure 3D). However, no one has studied mosaic *Gpam* WT vs. KO clones in the fatty liver microenvironment. Because conditional KO mice required to study clonal dynamics were not available, we generated *Gpam* floxed mice targeting exon 3, whose deletion is predicted to lead to a frameshift and a premature stop codon. Saturating doses of AAV8-TBG-Cre led to the deletion of exon 3 and almost complete depletion of *Gpam* mRNA one week after injection (Figure S3A,B). Using these floxed mice, we asked how WT and KO clones would fare under normal and WD conditions. In similar fashion as the *Mboat7* experiments, we used low dose AAV8-TBG-Cre to induce Tomato in 10-15% of hepatocytes in *LSL-tdTomato* het mice (control group), or to induce Tomato with *Gpam* deletion in 10-15% of hepatocytes in $Gpam^{f/f}$; *LSL-tdTomato* het mice, then gave these mosaic mice either chow or WD (Figure 4A,B). After 6 months of WD, the expected increases in body and liver weights were observed (Figure S3C). Although there were similar initial Tomato labeling frequencies in both groups (Figure 4C,D left column), and a similar level of *Gpam* WT and KO clonal expansion in the setting of chow (Figure 4C,D, middle column), we observed an increase in *Gpam* KO clones in the setting of WD (Figure 4C,D, right column). These data demonstrated the diet and steatosis dependent effects of *Gpam* mutations, which is in accord with the *Mboat7* tracing experiments (Figure 1H-J).

To determine why *Gpam* mutant clones were more fit, we investigated whole-liver *Gpam* KO mice induced with AAV8-TBG-Cre. Modest body and liver weight differences were observed after 3 months of WD (Figure S3D,E and Table S1), but KO mice showed a trend toward decreased liver/body weight ratios (Figure 4E), decreased steatosis (Figure 4F), and a trend toward improved transaminitis (Figure S3F,G). We used a NAFLD activity scoring system tailored to rodent histology to quantify steatosis and liver injury[23]. This showed that KO livers had reduced macrovesicular steatosis, microvesicular steatosis, and hepatocyte hypertrophy (Figure 4G). Liver triglyceride but not cholesterol levels were decreased in the WD fed KO group (Figure 4H,I), consistent with *Gpam*'s role as a rate limiting enzyme in triglyceride synthesis. Interestingly, plasma cholesterol but not triglyceride was decreased in the KO group (Figure S3H,I). We asked if increased survival or proliferation were responsible for the clonal expansion of *Gpam* deficient hepatocytes in the WD setting. The frequency of TUNEL positive nuclei in WT and *Gpam* KO livers were similar and low (Figure S3J,K). In WT livers, some apoptotic hepatocytes were filled with macroscopic lipid droplets, whereas in KO livers, the apoptotic hepatocytes did not have these characteristics (Figure S3L). Similarly, WT and KO liver sections showed comparable frequencies of proliferating hepatocytes (Figure 4J,K), suggesting that liver-wide *Gpam* KO does not lead to hyperproliferation. However, proliferating hepatocytes in WT livers were more likely to be found near portal triads, where steatosis was less pronounced, whereas in KO livers,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

proliferating hepatocytes were more evenly distributed between portal and central regions (Figure 4K,L), suggesting that steatotic hepatocytes were less likely to divide. To further investigate mechanisms of clonal disappearance, we compared the proliferation rates of Tomato+/*Gpam* KO and Tomato-/WT clones in mosaic livers (Figure 4M,N). We found that Ki67+ hepatocyte nuclei were more frequent in the Tomato+/*Gpam* KO compartment than in the Tomato-/WT compartment. Altogether, our data and previous studies show that *Gpam* loss of function mutations lead to clonal expansions in NASH, and that *Gpam* is a promising therapeutic target in NASH.

Somatic mosaic screening of candidate genes identifies putative targets for NASH

Positive selection of specific clones in the NASH environment suggests that some somatic mutations can promote fitness through altered lipid accumulation or resistance to lipotoxicity. Therefore, clone size/sgRNA abundance can be a surrogate measure for a gene's influence on cellular fitness in NASH livers. Based on fitness competition, MOSAICS represents a different way of identifying NASH genes. This concept was strengthened by the fact that the same mutated genes such as *Gpam* and *Acvr2a* were selected for in both mouse and human tissues[6]. Equipped with MOSAICS, we sought to identify unexpected fitness-promoting mutations in NASH. We reasoned that transcription or epigenetic factors might have the strongest impact on a broad array of pathways, so we analyzed human expression data to nominate candidate factors whose activities are altered in chronic liver disease (Figure S4). We performed modeling of putative transcription factor activities using gene expression data from NAFLD/NASH (72 patients; GSE130970)[24] and HCV cirrhosis cohorts (216 patients; GSE15654)[25]. Since transcription factor gene expression levels do not always reflect their functional importance, we chose those with the highest or lowest activities based on the induction or suppression of their downstream targets (See methods for details)[26,27]. The epigenetic genes were ranked only based on differential expression. To further enrich for the most influential factors, candidate genes were also associated with histologic features such as fibrosis, inflammation, ballooning, and steatosis, as well as temporal events such as Child-Pugh score, liver decompensation, cancer development, and death. Using this pipeline, we identified 217 transcriptional and 255 epigenetic regulators that have the highest likelihood of having an impact in human NAFLD/NASH (Table S4).

We next created sgRNA libraries corresponding to these two gene sets, and used them to generate somatically mutated mouse livers (Figure 3A). After 6 months, we observed liver steatosis and fibrosis in WD but not in chow groups, as expected (Figure S5A-F). SgRNA distributions were largely maintained before and after screening (Figure S5G,H and Table S3). The sgRNAs associated with the most clonal expansion in WD livers targeted 13 and 10 genes in each screen, respectively (Figure 5A-C and Table S2). Several genes were previously connected with fatty liver disease such as the glucocorticoid receptor *Nr3c1*[28,29], the mineralocorticoid receptor *Nr3c2*[30,31], the transcription factor *Zbtb16*[32], the DNA methyltransferase *Dnmt1*[33], and the E3 ubiquitin ligase adaptor *Keap1*[34,35], but most of the top genes had no connections with liver disease or NASH.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Conditional deletion methods for rapid liver disease phenotyping

To further ascertain which of the top genes, when genetically ablated in all hepatocytes, have the most positive influences on metabolic health, we developed rapid conditional KO approaches. We cloned the most enriched sgRNA for each of the top genes into the MOSAICS vector, then delivered saturating doses of these AAVs with individual guides into Cas9 expressing mice to achieve single gene, whole-liver deletion (Figure 2C,D). We first tested this system using the four known NASH genes that were the most positively selected hits from the 63 NASH gene screen (Figure 3). High titers of MOSAICS AAV8-sgRNAs against *GFP* (control 1), *LacZ* (control 2), *Acvr2a, Irs1, Srebf1* or *Dgat2* (Figure S6A) were injected into Cas9 expressing mice, which were then fed WD for 3 months (Figure 5D). The remaining intact mRNAs of the targeted genes in KO livers fell to 25-50% of the levels in WT livers, indicating efficient somatic gene editing (Figure S6B). After 3 months of WD, each conditional KO mouse was metabolically phenotyped (Table S1). *Irs1, Srebf1,* and *Dgat2* liver KO mice had reduced liver to body weight ratios (Figure 5E,F and Figure S6C). Furthermore, the protection against fatty liver disease was evident from improved plasma lipids, decreased liver lipid content (Figure 5G-I and Figure S6D-F), and reduced steatosis (Figure 5J and Figure S7A,B). These findings correlated with the degree of liver/body weight ratio reduction, a surrogate for fat accumulation. This showed that CRISPR liver-wide conditional deletion was an effective and rapid method of organ level validation.

Tbx3, Bcl6, and *Smyd2* loss showed protection against fatty liver disease

An advantage of this KO system was that it allowed us to rank the impact of the top hits ($p < 0.05$), and to compare unknown vs. canonical NASH genes. Among all 20 KO mice associated with the transcription and epigenetic factor screens (Table S1), more than half of the KO models showed significant reductions in liver mass and none showed significant increases (Figure S6C). This suggested that mutations that drive clonal expansions more often led to reduced hepatic lipid accumulation. Among all KO models, deletion of the transcription factors *Tbx3* and *Bcl6,* as well as the epigenetic factor *Smyd2,* showed significant reductions in body weights (Figure 5E) and the most substantial reductions in liver/body weight ratios compared to controls after 3 months of WD (dark dots in Figure 5F). *Bcl6* is a transcriptional repressor that is enriched in the fed state and cooperates with PPARα to suppress the induction of fasting transcription, and *Bcl6* loss has been shown to increase insulin sensitivity and repress NASH[36-38]. Neither *TBX3*, a T-box transcription factor[39], nor *SMYD2*, a protein lysine methyltransferase with histone and non-histone methylation targets, has been studied in fatty liver disease. These three KO models showed the most significant reductions in liver injury as measured by ALT (Figure 5G), a trend toward reduced AST (Figure S6D), and the most significant reductions in liver triglyceride levels (Figure 5H). While *Tbx3* and *Smyd2* KO mice had reduced liver cholesterol, *Bcl6* KO mice had increased liver cholesterol (Figure 5I). We also observed reduced plasma cholesterol in *Smyd2* KO mice and increased plasma triglyceride in *Tbx3* KO mice, while the other KO models showed modest or no changes in these parameters (Figure S6E,F). Compared to controls and other KO models, there were clear improvements in hepatic steatosis in *Tbx3, Bcl6,* and *Smyd2* KO livers (Figure 5J and Figure S7C,D). The phenotypes of *Tbx3, Bcl6,* and *Smyd2* KO mice were comparable to *Srebf1* and *Dgat2* KO mice, suggesting potent regulation of lipid metabolism by these three genes.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

To examine late stage NASH features, we fed some CRISPR KO models for 9 months with WD (Figure S8 and Table S1). In male *Tbx3* KO mice, liver weight and liver/body weight ratios decreased, while body weights were unchanged (Figure S8A). Significant decreases of ALT and AST were also observed (Figure S8B), suggesting reduced liver damage. In male *Smyd2* KO mice, liver/body weight ratios decreased, but liver and body weights were unchanged (Figure S8A), and ALT/AST levels were unchanged (Figure S8B). In the plasma, both *Tbx3* and *Smyd2* KOs showed moderate increases in triglyceride but cholesterol levels remained similar to controls (Figure S8C). These data showed that both KO models protect the liver from steatosis, but male *Tbx3* KO had more pronounced effects under long-term WD (Figure S8D). Although the study was underpowered to detect statistical differences in fibrosis in KOs vs. controls, we observed trends toward decreased fibrosis in both KO models (Figure S8D,E).

Compared to males, females in the 9-month study were more resistant to NASH. We observed a trend towards decreased liver/body weight ratios in *Tbx3* KO livers and a significant decrease in *Smyd2* KO livers (Figure S8F). We did not see improvements for ALT and AST in *Tbx3* KO mice but did see modest improvements in *Smyd2* KO mice (Figure S8G). Plasma triglyceride levels were increased in *Tbx3* KO but not in *Smyd2* KO mice, while cholesterol in both KO cohorts were unaffected (Figure S8H). Both KO livers showed improvements in steatosis (Figure S8I). There were no improvements for fibrosis in *Tbx3* KO livers but a trend of decreased fibrosis in *Smyd2* KO livers (Figure S8I,J). These data suggested that in females, *Smyd2* KO mice were more protected from steatosis than *Tbx3* KO mice after long-term WD. Gender differences in the progression of NASH have also been observed in *Bcl6* deficient mice[40].

Transcriptional analysis revealed diverse mechanisms of lipid regulation

To investigate the expression changes in WD fed livers carrying fitness promoting mutations, we performed RNA-seq on control (sg*GFP,* sg*LacZ*), *Irs1, Srebf1, Tbx3, Bcl6*, and *Smyd2* KO livers generated with MOSAICS AAVs (Table S5). We observed massive expression changes in *Irs1* KO livers and fewer changes in *Tbx3* and *Bcl6* KO livers (Figure 6A), which is congruent with *Irs1* KO phenotypes showing the most substantial decreases in body weight, liver weight, inflammation, and steatosis (Figure 5E-J), and supports the known regulation of lipogenesis by insulin signaling. *Srebf1* KO livers showed fewer gene expression changes compared with *Irs1, Tbx3,* and *Bcl6* KO livers; *Smyd2* KO livers showed the smallest number of expression changes (Figure 6A). Comparing differentially regulated genes in control vs. KO livers showed a relatively large intersection between *Tbx3* and *Irs1* KO livers (Figure 6B), suggesting shared regulatory circuits. The intersections among *Bcl6, Tbx3,* and *Srebf1* KO livers were smaller (Figure 6C), suggesting unique transcriptional roles. Using Gene Set Enrichment Analysis (GSEA), we observed reduced inflammatory pathways in *Irs1, Srebf1,* and *Tbx3* KO livers and reduced proliferation in all KO compared with control livers (Figure 6D), suggesting decreased cell turnover in KO livers. For lipid metabolism pathways in GSEA analysis, we observed some shared but mostly unique expression patterns for each KO group (Figure 6D). We further examined the expression of genes involved in lipogenesis and fatty acid oxidation. Deletion of *Irs1, Srebf1, Tbx3, Bcl6,* and *Smyd2* each led to decreased expression of fatty acid and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

triglyceride synthesis genes, but to different extents (Figure 6E). We also observed decreases in cholesterol synthesis genes in *Irs1* and *Smyd2* KO livers, but a dramatic upregulation of these genes in *Srebf1* KO livers (Figure 6E). Since *Srebpl* is a master transcription factor for cholesterol synthesis, this increase was potentially due to a compensatory effect on cholesterol production. For lipid degradation pathways, strong decreases in β-oxidation genes were observed in *Irs1* and *Bcl6* KO livers (Figure 6F, **upper panel**), likely because these livers had the least triglyceride content among all groups (Figure 5H). We observed upregulation of many TCA cycle genes in *Tbx3* and *Smyd2* KO livers (Figure 6F, **lower panel**), suggesting that these livers have higher rates of acetyl-CoA consumption. We also analyzed the mRNA expression of collagen, a major component of fibrosis, and observed the downregulation of multiple collagen mRNAs in KO livers (Figure 6G). These data indicate shared and unique mechanisms by which mutant hepatocytes converge on decreased steatosis and fibrosis.

A selective small molecule inhibitor of SMYD2 can prevent fatty liver disease

We sought to determine if chemical SMYD2 inhibition could serve as orthogonal validation of the genetic findings. We used AZ505, a selective SMYD2 inhibitor with an IC_{50} of 0.12 μM[41] (>600 fold selectivity). WD was fed to WT mice at 8 weeks of age, followed by treatment with vehicle control or AZ505 with a dose of 10mg/kg (intraperitoneal injection) three times per week (Figure 7A, **left panel** and Table S1). Two months after continued WD and inhibitor treatment, we observed significant reductions in body weight, liver weight, and liver/body weight ratios in AZ505 vs. vehicle treated mice (Figure 7B-D). Smaller, less pale livers were observed in AZ505 treated mice, and liver histology showed decreased steatosis in contrast to a wide range of macro and microscopic lipid droplet deposition in vehicle treated livers (Figure 7E). AZ505 treated mice showed improved transaminitis (Figure 7G,H) and a significant decrease in liver triglycerides (Figure 7K), whereas the liver cholesterol (Figure 7L) and plasma lipids (Figure 7I,J) showed modest increases. We also gave AZ505 or vehicle to mice fed normal chow diets (Figure 7A, **right panel** and Table S1). After two months, we observed no significant impact on body or liver weights (Figure 7B,C), although there was a decrease in the liver/body ratio in AZ505 treated mice (Figure 7D). No changes were observed in liver histology (Figure 7F), AST/ALT (Figure 7G,H), or plasma lipids (Figure 7I,J). Thus, AZ505's protective effects while on WD were not due to NAFLD-independent influences on the liver. Because AZ505 binds to the active center of SMYD2[41], these experiments also showed that SMYD2 enzymatic activity, and not just protein levels, contributes to fatty liver phenotypes. Altogether, several lines of orthogonal evidence suggested that SMYD2 is a promising NASH target.

Discussion

Mutant clone expansion and recurrent mutations within and between individuals provide compelling evidence for increased mutant clone fitness in chronic liver disease[6]. Fatty liver disease has not been conceptualized as a disease of genetic mosaicism, so the biological impact of clonal heterogeneity has not yet been functionally explored. We established the MOSAICS platform to model widespread somatic mosaicism. The advance associated with this technology is that it allows for *in vivo* screening with a much greater clone density,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and in a low-proliferation setting, in contrast to the *Fah* KO screening model. Tracing of *Mboat7* and *Gpam* mutant hepatocytes and the results from the MOSAICS screens support the hypothesis that lipid laden cells have a fitness disadvantage compared to cells with less fat. This environmental pressure to reduce lipid accumulation fosters the selection of adaptive mutations within NASH livers.

Just because expanding mutant clones are more fit does not mean that the causative mutations, if applied to the entire tissue, can also increase organismal health. Genomic data and *in vivo* screening in isolation are not sufficient to clarify how fitness is promoted, and whether or not this type of fitness is selfish or adaptive. The rapid whole-liver KO method allowed head to head comparisons between more than 20 different types of KO mice. A significant subset of positively selected somatic mutations, when induced in the entire liver, led to improvements in fatty liver phenotypes. It is likely that the clonal expansions associated with these mutations are caused by increases in cell survival, in addition to modest proliferation changes. In humans, these fitness improvements are enough to drive measurable levels of positive selection over decades.

Remarkably, multiple lines of evidence suggest that these mutations do not promote cancer since many of these genes were not identified in large cancer sequencing efforts, and our *in vivo* experiments did not reveal increased tumorigenesis. While liver-wide fitness increases were often observed, in some cases such as *Irs1* deletion, this could come at the expense of organismal fitness in the form of insulin resistance or diabetes. Thus, mutations that lead to clonal expansion may not always be suitable therapeutic targets.

Convergent and recurrent somatic mutations suggest that in disease conditions, humans are in effect performing pooled genetic screens that select for mutations that promote fitness in fatty liver disease. Because adaptive mutations are rare and clonal expansion takes decades, it is likely that beneficial clones cannot expand enough to prevent or reverse NASH development within the lifetime of the patient. Nevertheless, somatic mosaicism will likely reveal adaptive mechanisms and potential therapeutic targets. This led us to hypothesize that testing a larger genetic space could uncover additional unexpected pathways in metabolic liver disease. Using evolutionary selection between genetically heterogeneous clones in cell culture systems has been a mainstay of cancer gene discovery. However, metabolic disease has not identified a facile method of gene discovery using phenotypic screens, *in vitro* or *in vivo.* Ultimately, it is difficult to phenotype many genetic strains of mice or humans, but it is much easier to phenotype thousands of genetically distinct and competing cells. The output of MOSAICS validates somatic mutation selection as a gene discovery engine in metabolic disease.

Some of the most important human metabolic disease genes have been identified using germline genetic methods such as GWAS and whole exome sequencing[42-44]. However, germline mutations are limited in the amount of genetic space that can be explored because lethal developmental phenotypes select against mutations that could have a benefit in specific tissues later in life. This study, in tandem with somatic sequencing studies from multiple labs, shows that somatic mutations represent an alternate source for important

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

disease genes. Positive selection of mutant clones in MOSAICS mice, and in human tissues, is an effective way of identifying metabolic disease regulators and therapeutic targets.

Limitations of this study

While somatic mutations and the MOSAICS platform can potentially identify therapeutic targets, the ultimate clinical significance of the genes identified is unknown. While not all of these hits will have high therapeutic value, it is possible that they have important biological influences in liver disease. Some of the genes identified in this study, when deleted or inhibited, may not cause improvements in all metabolic parameters associated with fatty liver disease. Furthermore, some of the hits may not have an impact on late stage NASH features.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact—Further information and requests for resources and reagents will be fulfilled by the lead contact, Hao Zhu (hao.zhu@utsouthwestern.edu).

Materials Availability—Plasmids, mouse strain, and other unique reagents generated in this study will be distributed upon request from the lead contact, Hao Zhu (hao.zhu@utsouthwestern.edu), after completion of relevant Materials Transfer Agreements.

Data and Code Availability

- All RNA-seq data have been deposited to GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- There is no original code in this paper.
- Any additional information required to reanalyze the data in this paper will be made available by the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse strains and breeding—All mice were handled in accordance with the guidelines of the Institutional Animal Care and Use Committee at UT Southwestern. All experiments were done in an age and sex controlled fashion unless otherwise noted. All mice used in this study were male except for those in Figure S8F-J. C57BL/6 strain background mice were used for all experiments. *Mboat7*$^{tm1a(KOMP)Wtsi}$/*H* mice, obtained from KOMP, were originally generated by the Wellcome Trust Sanger Institute. *Gpam*$^{f/f}$ mice were generated in the CRI Mouse Genome Engineering Core at UT Southwestern. In brief, CAS9 protein, synthetic sgRNA, and single-stranded DNA containing *Gpam* exon3 flanked by LoxP sites and homology arms, were co-injected into C57BL/6 mouse zygotes, which were then implanted into CD-1 mice. Genotyping and Sanger sequencing was used to confirm homologous recombination in the genome-edited pups. *LSL-tdTomato* (strain #007914) and *Rosa-rtTA; TetO-Cas9* mice (#029415) were obtained from The Jackson Laboratory. Mice homozygous for both *Rosa-rtTA* and *TetO-Cas9* were used to ensure a high Cas9 expression level in the liver. Western Diet (WD) used for NAFLD/NASH modeling is

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

described in[45]. It is composed of high fat solid food (ENVIGO #TD. 120528) and high sugar water containing 23.1g/L d-fructose (Sigma-Aldrich #F0127) and 18.9 g/L d-glucose (Sigma-Aldrich #G8270).

Lineage tracing in floxed mice—For whole-liver deletion of *Mboat7* or *Gpam*, $5*10^{10}$ genomic copies of commercially produced AAV8-TBG-Cre (Addgene #107787) or control AAV8-TBG-GFP (Addgene #105535) in 100μl saline was injected retro-orbitally into *Mboat7*$^{f/f}$ or *Gpam*$^{f/f}$ mice at 8 weeks of age. For mosaic deletion of *Mboat7* or *Gpam* and Tomato labeling of hepatocytes, $0.125*10^{10}$ genomic copies of AAV8-TBG-Cre in 100μl saline was injected retro-orbitally into *Mboat7*$^{f/f}$*; LSL-tdTomato* het, *Gpam*$^{f/f}$*; LSL-tdTomato* het, or control *LSL-tdTomato* het mice at 8 weeks of age. One week after injection of low dose AAV8-TBG-Cre, 7 mice from each group were collected to determine the initial Tomato labeling percentage in the liver. The remaining mice were divided into chow or WD groups and traced for another 4 or 6 months.

METHOD DETAILS

Fluorescent imaging and image processing—For fluorescent imaging, liver pieces were fixed in buffered formalin (Fisherbrand #245-685) for 24h with gentle shaking at 4°C and then transferred into 30% sucrose (w/v) solution for another 24h with shaking at 4°C. The livers were then embedded and frozen in Cryo-Gel (Leica #39475237), and sectioned at a thickness of 16μm. Images were taken using a Zeiss Axionscan Z1 system in the UTSW Whole Brain Microscopy Facility to visualize Tomato clones. To statistically analyze the percentage of Tomato+ cells, black and white fluorescent images were taken from the same slide using an Olympus IX83 microscope at 4x magnification. Two different fields were taken for each liver. The percentage of Tomato+ cells (bright areas) was analyzed using ImageJ.

H&E, immunohistochemistry (IHC), immunofluorescence (IF), TUNEL, and Sirius Red staining—Liver pieces were fixed in buffered formalin (Fisherbrand #245-685) for 24h with gentle shaking at 4°C and then transferred to 70% EtOH for another 24h with shaking at 4°C. Paraffin embedding, liver sectioning (4μm thickness), and H&E staining were performed at the UT Southwestern Tissue Management Shared Resource Core. For IHC staining, the following primary antibodies were used: PTEN (CST #9559, IHC 1:200); Ki67 (Abcam #AB15580, IHC 1:200). IHC was performed as previously described[46]. Briefly, paraffin-embedded sections were dewaxed in xylene and hydrated using ethanol gradients. The slides were then boiled in antigen retrieval buffer (10mM sodium citrate, 0.05% Tween 20, pH 6.0) for 20min and soaked in 3% hydrogen peroxide (in methanol) for 10min. After blocking with 5% goat serum for 1h at room temperature, the slides were incubated with primary antibody overnight at 4°C. After washing, the slides were incubated with secondary antibody at room temperature for 0.5h. The secondary antibody was coupled with HRP using ABC-HRP Kit (Vector laboratories, #PK-6101). The slides were developed using the DAB Kit (Vector laboratories, #SK-4100). For IF staining, the following primary antibodies were used: RFP (Rockland #600-401-379, IF 1:500); Ki67 (Invitrogen #14-5698-82, IF 1:500); HNF4a (Abcam #ab41898, IF 1:500), and the following secondary antibodies were used: Goat anti-rat IgG (H&L) Alexa Fluor

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Plus 488 (Invitrogen #A-48262, IF 1:500); Donkey anti-rabbit IgG (H&L) Alexa Fluor 594 (Invitrogen #A-21207, IF 1:500); Goat anti-mouse IgG2a Alexa Fluor 647 (Invitrogen #A-21241, IF 1:500). IF was performed on paraffin embedded mouse liver sections using the same protocol as IHC except that secondary antibodies were substituted by Alexa Fluor conjugated antibodies. TUNEL staining was performed on paraffin embedded liver sections using *In Situ* Cell Death Detection Kit, Fluorescein (Roche #C755B40) according to the manufacturer's protocol. Sirius Red staining was performed on paraffin embedded liver sections using the Picro Sirius Red Staining Kit (Abcam #ab150681) according to the manufacturer's protocol. QuPath software was used to quantify TUNEL staining and IHC staining of Ki67. ImageJ was used to quantify Sirius Red staining.

Plasma and liver metabolic assays—Blood was taken using heparinized tubes from the inferior vena cava immediately after sacrificing the mouse, and then transferred into 1.5ml tubes and centrifuged at 2000g for 15min at 4°C. The supernatant after centrifugation (plasma) was analyzed for AST, ALT, cholesterol, and triglyceride (Manufacturer's Reference Numbers 8433815, 1655281, 1669829, and 1336544, respectively) using a fully automated OCD Vitros 350 dry chemistry analyzer following the protocols provided by the reagent kit manufacturer (Ortho Clinical Diagnostics, Raritan, NJ) at the UT Southwestern Metabolic Phenotyping Core. 100-150mg of liver per mouse was weighed and used for lipid extraction and quantification at the UT Southwestern Metabolic Phenotyping Core. Briefly, flash frozen tissue samples were homogenized with 2:1 chloroform:methanol mixture (v/v) using a multiplexed automatic tissue disruptor (TissueLYser II, Qiagen, Germantown, MD). The organic extract was transferred to a 5 ml graduated flask and the total volume was brought up to 5 ml. Total cholesterol and triacylglycerol analyses were performed in triplicate using 100μl and 25μl of lipid extracts, respectively. Total cholesterol and triacylglycerol concentrations were determined by commercial enzymatic colorimetric assays following the protocols described by the manufacturer (Infinity™ Cholesterol Liquid Stable Reagent #TR13421; Infinity™ Triglycerides Liquid Stable Reagent #TR22421; Matrix plus™ Chemistry Reference Kit #NC9592194).

MOSAICS reagent construction—The MOSAICS plasmid uses the pX602 plasmid as a backbone. The sequence between the two AAV ITRs were removed using the NsiI and NotI restriction enzymes. The following fragments were cloned between the two AAV ITRs: the first SB100 binding IR, a U6 driven sgRNA scaffold, a CAG promoter driven SB100-P2A-Cre fusion cDNA with a beta-globin poly(A) signal, and the second SB100 binding IR. For library construction, mouse candidate genes for all of the *in vivo* screens were generated by using the mouse homologs of the human genes shown in Table S4. A few genes were not included in the mouse gene lists due to the lack of a homolog or because they were known tumor suppressor genes. The individual sgRNA sequences corresponding to mouse candidate genes were extracted from the Brie library[47] or obtained from the GUIDES server (http://guides.sanjanalab.org/). and synthesized by CustomArray. Most genes had 5 distinct sgRNAs, 4 from Brie and 1 from GUIDES. A few genes had 4 targeting sgRNAs due to the overlap of sgRNA sequences from Brie and GUIDES. See Table S2 for the sgRNA sequences of targeted genes. The library construction protocol was previously described[48]. Briefly, synthesized oligonucleotide libraries were amplified by PCR, purified using a PCR

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Purification Kit (Qiagen, #28104), and assembled in BsaI digested MOSAICS vector using DNA Assembly Kit (NEB, #E5520A). 1μl of the assembled vector was then electroporated into 25μl competent cells (Lucigen, #60052-2). After recovery in SOC medium for 1h, bacteria were spread on a 245*245mm LB agar plate and incubated at 37°C overnight. The bacteria were then harvested for plasmid preparation using the HiSpeed Plasmid Maxi Kit (Qiagen, #12663). Each sgRNA maintained a >1000-fold representation during construction. For individual sgRNA cloning, forward and reverse primers were annealed and fused to BsaI digested MOSAICS plasmid using T4 ligase. See Table S6 for the primers associated with individual sgRNAs used in this paper.

AAV production and purification—AAV8 was produced using AAV-Pro 293T cells (Takara #632273) cultured in one or more 15cm dishes. Cells were plated one day before transfection at 50% confluence, which would allow the cells to reach 80-90% confluence the next day. For transfection of one 15cm dish, 10μg MOSAICS vector, 10μg pAAV2/8 (Addgene #112864) and 20μg pAdDeltaF6 (Addgene #112867) plasmids were mixed with 1ml Opti-MEM medium in one tube. In another tube, 160μl PEI solution (1mg/ml in water, pH7.0, powder from ChemCruz #sc-360988) was mixed with 1ml Opti-MEM medium. The solutions from both tubes were then mixed and incubated for 10min before adding to cell culture. 48h after transfection, the cells were scraped off the dish and collected by centrifugation at 500g for 10min. The supernatant was disinfected and discarded, and the cell pellets were lysed in 1.5ml/15cm dish lysis buffer (PBS supplemented with NaCl powder to final concentration of 200mM, and with CHAPS powder to final concentration of 0.5% (w/v)). The cell suspension was put on ice for 10min with intermittent vortexing, and then centrifuged at 20,000g for 10min at 4°C. The supernatant containing the AAV was collected. To set up the gravity column for AAV purification, 0.5ml of AAV8-binding slurry beads (ThermoFisher #A30789), enough to purify AAV from three 15cm dishes, was loaded into an empty column (Bio-Rad #731-1550). After the beads were tightly packed at the bottom, they were washed with 5ml of wash buffer (PBS supplemented with NaCl powder to a final concentration of 500mM). The supernatant containing AAV was then loaded onto the column. After all of the supernatant flowed through, the beads were washed with 10ml wash buffer twice. The AAV was then eluted with 3ml elution buffer (100mM glycine, 500mM NaCl in water, pH 2.5) and the eluate was immediately neutralized with 0.12ml 1M Tris-HCl (pH 7.5-8.0). The AAV was concentrated by centrifugation at 2000g for 3-5min at 4°C using an 100k Amicon Ultra Centrifugal Filter Unit (Millipore #UFC810024). After centrifugation, the volume of AAV should be equal to or less than 0.5ml. The concentrated AAV was diluted with 4-5ml AAV dialysis buffer (PBS supplemented with powders to final concentrations of 212mM NaCl and 5% sorbitol (w/v)) and centrifuged at 2000g for 3-5min at 4°C. The dilution and centrifugation processes were repeated 3 times. The final concentrated AAV was transferred into a 1.5ml tube and centrifuged at 20,000g for 5min to remove debris. The supernatant was aliquoted, flash frozen using liquid nitrogen, and stored at −80°C.

MOSAICS screening and single gene deletion in inducible Cas9 mice—To functionally validate the MOSAICS vector, a *Pten* sgRNA (see Table S6 for the *Pten* sgRNA primers) was cloned into the vector and the corresponding AAV8 was produced.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Different volumes of concentrated AAV8-*sgPten* were diluted to a final volume of 100μl using saline, then retro-orbitally injected into *Rosa-rtTA; TetO-Cas9* double homozygous mice 3 days after dox water (1mg/ml dox) was initiated at 6.5 weeks of age. After determination of *Pten* KO efficiency using IHC and confirmation of SB100-Cre fusion protein expression and recombination using fluorescence imaging with *LSL-tdTomato* mice (Figure 2C), the concentration of this AAV8-*sgPten* virus was used as the standard. The relative concentrations of the other AAVs, including library AAVs and single gene targeting AAVs were all compared to AAV8-*sgPten* by qPCR using a pair of primers within the Cre sequence of the vector (See Table S6 for primer sequences). The absolute genomic copy numbers were not determined. For the AAV libraries used for the NASH, transcription factor, and epigenetic factor screens, a volume corresponding to 7μl of AAV8-*sgPten* (Figure 2D,E) was injected. For the AAV mini-pool containing 8 sgRNAs (Figure 3E), a volume corresponding to 1μl of AAV8-*sgPten* (Figure 2D,E) was injected. For AAVs targeting individual genes (Figure S6A), a volume corresponding to 20μl AAV8-*sgPten* (Figure 2D,E) was injected. All AAVs were diluted to a final volume of 100μl using saline prior to injection. Dox water was withdrawn 10 days after injection and then chow or WD was given to mice for the specified times.

Genomic DNA extraction, sgRNA amplification, and amplicon library construction—To extract genomic DNA containing the integrated sgRNA, the entire liver (except a small piece used for sectioning and H&E staining) was minced into about $1mm^3$ pieces using a blade and weighed. Small nodules observed in some epigenetic factor screening livers given WD (<=3 nodules per liver in 5 out of 8 livers) were excluded from samples being processed for genomic DNA extraction. Minced liver in two volumes (w/v) of homogenizing buffer (100mM NaCl, 25mM EDTA, 0.5% SDS, 10mM Tris-HCl, pH 8) was transferred into a glass Wheaton Dounce Tissue Grinder and stroked 50 times or until no bulk tissues were seen. After homogenizing, 200μl chow fed liver lysate or 300μl WD fed liver lysate was transferred to a 15ml tube for genomic DNA extraction using the Blood & Cell Culture DNA Midi Kit (Qiagen #13343) according to the manufacturer's protocol. The remaining lysates were frozen in −80°C as backup samples. Briefly, 10ml Buffer G2 from the kit, 100μl Proteinase K (Roche #03115828001, or Proteinase K from the Qiagen kit) and 100μl RNase A (Invitrogen #12091-021) were added to the 15ml tube containing the lysate and digested in a 50°C water bath overnight. The next day, the tubes were centrifuged at 4000g for 10min and the lipid layer on the top was discarded. The remaining supernatant was loaded on the column, washed, and genomic DNA elution/precipitation were performed according to the manufacturer's protocol. The precipitated DNA was resuspended in 100μl 10mM Tris (pH 8.0) and shaken on a 55°C shaker for 2h to help it dissolve. The amplicon library preparation protocol for high-throughput sequencing was adapted from [48]. Briefly, 5μg genomic DNA, 5μl general forward primer mix (5μM), 5μl barcode specific reverse primer (5μM), 1μl Q5 DNA polymerase, 10μl Q5 buffer, 10μl HighGC buffer, 1μl dNTP, and water was mixed for a 50μl PCR reaction, and two reactions were made for each genomic DNA sample. The PCR cycle was 95°C 3min-(95°C 30s-60°C 30s-72°C 20s)*n-72°C 2min. The PCR cycle number was pre-optimized using the same PCR reactions with a smaller volume. The cycle numbers that gave a weak but sharp band on the DNA gel were used. In the final PCR reaction, 23 cycles were used for preparing the NASH gene, transcription

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

factor, and epigenetic factor screens, and 30 cycles were used for preparing the 8 guide mini-pool validation screen. After PCR, the two tubes of reactions with the same genomic DNA template were combined (total 100μl) and 70μl was resolved on a DNA gel. The 250bp band corresponding to the amplicon was cut and purified using the QIAquick Gel Extraction Kit (Qiagen #28704). The DNA concentration was determined using Qubit kit (Invitrogen #Q32853) and high-throughput sequencing was performed using an Illumina NextSeq500 system at the CRI at UT Southwestern Sequencing Facility.

Bioinformatic analysis of MOSAICS screening results—The reads from the sequencing of amplicon libraries described above were trimmed with cutadapt (version 1.9.1) to remove the excessive adaptor sequences so that only the sgRNA sequences were retained. The 5' sequences were trimmed with the options -O 32 --discard-untrimmed -g CTTTATATATCTTGTGGAAAGGACGAAACACCG. The 3' sequences were trimmed with the options -O 12 -a GTTTTAGAGCTAGAAATAGCA. The abundance of each sgRNA was calculated with the count function in MAGeCK (version 0.5.6) with the default options[14]. The trimmed fastq files were assigned to chow-fed and WD-fed groups and uploaded together with library files containing sgRNA sequences and targeted gene names to a server preloaded with MAGeCK. The enrichment of each sgRNA was calculated with the test function in MAGeCK.

RNA-seq library preparation and transcriptome analysis of mouse fatty livers —Total liver RNA was extracted from 4 sg*GFP,* 2 sg*LacZ,* 5 *Irs1,* 5 *Srebf1,* 7 *Tbx3,* 5 *Bcl6* and 6 *Smyd2* KO livers using TRIzol reagent (Invitrogen #15596026) followed by purification using the RNeasy Mini kit (Qiagen #74014). Briefly, a liver fragment with a volume of about 3*3*3 mm^3 from each sample was homogenized in 1ml Trizol, followed by adding 200μl chloroform and vortexed. After centrifugation at 20,000g for 10min at 4°C, 350μl supernatant from each sample was transferred to a new tube and mixed with equal volume of 75% EtOH, and then loaded on an RNeasy column. The following wash steps using RW1 and RPE buffers and RNA precipitation step were performed according to the manufacturer's protocol. RNA-seq libraries were prepared with the SMARTer Stranded Total RNA Sample Prep Kit - HI Mammalian (Takara #634875). 75 bp single-end sequencing was performed using an Illumina NextSeq500 system at the CRI at UT Southwestern Sequencing Facility. Alignment, quantification, and differential expression analysis were performed using the QBRC_BulkRnaSeqDE pipeline (https://github.com/QBRC/QBRC_BulkRnaSeqDE). Briefly, the alignment of reads to the mouse reference genome (mm10) was done using (v2.7.2b)[49]. FeatureCounts (v1.6.4)[50] was then used for gene count quantification. Differential expression analysis was performed using the R package DEseq2 (v1.26)[51]. Cutoff values of absolute fold change greater than 2 and FDR<0.05 were used to select for differentially expressed genes between sample group comparisons. Finally, GSEA was carried out with the R package fgsea (v1.14.0) using the 'KEGG' and 'Hallmark' libraries from MsigDB.

Drug treatment—AZ505 powder was purchased from MCE (#HY-15226). 40mg/ml AZ505 in 100% DMSO stock solution was made, aliquoted, and stored at −80°C. Before use, the stock solution was diluted 40x using 0.9% sodium chloride (Baxter #2F7123).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

WD was given to B6 mice at 8 weeks of age for a total of 2 months. AZ505 or vehicle (0.9% sodium chloride containing 2.5% DMSO) treatment started one day after WD and was administered intraperitoneally at a dose of 10mg/kg three times per week until one day before euthanasia and collection of livers.

Transcriptomic analysis of human NASH livers—Dataset 1 (NAFLD/NASH cohort): We downloaded the RNA-Seq transcriptome profiles of biopsied liver tissues (GEO: GSE130970)[24] and analyzed 72 NAFLD/NASH patients with a range of disease severities (NASH activity scores of 1 to 6) and relevant histological features, i.e., steatosis, inflammation, fibrosis, and hepatocyte ballooning. The raw sequence reads were aligned to the GENCODE human reference genome (GRCh37, p13) using the STAR aligner (ver 2.6.1b)[49], and gene-level count data were generated by the featureCounts function in the Subread package (ver 1.6.1)[50] and the GENCODE genome annotation (GRCh37, v19)[52]. The count data were normalized using "Relative Log Expression" normalization (RLE) implemented in the DESeq2 package[51]. Dataset 2 (HCV cirrhosis cohort): We analyzed the microarray gene expression profile of formalin-fixed needle biopsy specimens from the livers of 216 patients with hepatitis C-related early-stage (Child-Pugh class A) cirrhosis (GEO: GSE156540)[25]. This cohort was prospectively followed for a median of 10 years at an Italian center with relevant time event outcomes collected, including child, death, HCC and decomposition. Male and female patients were included in these studies.

Curating transcription and epigenetic factors with putative pathogenic activity in NAFLD/NASH—Gene expression of transcription factors do not directly reflect their functional activities, due to the low correlation between gene expression and protein abundance as well as the co-regulation with co-factors. Transcription factor activities can be estimated using enrichment of their downstream targets[26,27]. However, as we observed in our previous study[53], traditional methods such as Gene Set Enrichment Analysis (GSEA) overestimated the number of significant transcription factors, largely confounded by the large number of overlaps among putative target genes of different TFs[54]. Therefore, we performed global modeling of putative transcription factors[55] with gene expression data from NAFLD/NASH (GEO: GSE130970)[24] and HCV cirrhosis patients (GEO: GSE15654)[25], to directly infer the transcription factor activities from their downstream targets with adjustment for overlapping targets. A similar linear regression-based model was previously proposed to predict transcription factor regulatory activities and motifs from yeast gene expression data[56]. The method regresses the fold-change of a gene on its putative regulatory transcription factor(s). The coefficient (Z score) of a transcription factor, estimated using genome-wide fold changes and predicted targets of all transcription factors, represents the regulatory activity change of the transcription factor across all the liver patients.

The regression model is defined as following:

$$f_g = \sum_i \alpha_i T_{ig} + \sum_j \beta_j M_{jg} + c$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

where f_g is the fold change of *g-th* gene between two conditions; T_{ig} is the number of binding sites of *i-th* TF on the promoter of the *g-th* gene; M_{ig} is the number of binding sites of the *j-th* microRNA on the 3' UTR of the *g-th* gene; and α_i, β_j and c (a constant) can be inferred based on the values of $f_g T_{ig}$ and M_{jg} for all the genes in the RNA-seq data. The Z scores of coefficients α_i and β_j represent the activity changes of the *i-th* TF and j-th microRNA. Global transcription factor binding sites represented by 190 position-weighted matrices (PWMs) covering 500 mammalian TFs were based on the union of JASPAR[57], TRANSFAC[58], and additional motifs from chromatin immunoprecipitation with DNA microarray and ChIP-seq data collected by Balwierz et al.[55]. The initial regression analysis was done using ISMARA before further integrative analysis with patient clinical histological features and time event outcome; sample-specific transcription factor activity was estimated by the same regression model, where the fold changes were calculated between a single sample and all the samples combined together.

In the NAFLD/NASH cohort, the sample-level activities for each TF were associated with the four histological features, including fibrosis, inflammation, ballooning, and steatosis based on Pearson correlation. In the HCV cirrhosis cohort, the activities for each TF were used to perform outcome analysis on four time events, including child, death, HCC and decomposition using cox proportional regression model. The p-values were calculated for both analyses respectively, followed by the calculation of False Discovery Rate (FDR) for multiple testing correction.

Similarly, to screen the putative pathogenic epigenetic regulators, we also calculated the Pearson correlation between the gene expression of epigenetic regulators and the four histological features (fibrosis, inflammation, ballooning, and steatosis) in the NAFLD/NASH cohort, and used gene expression of known epigenetic regulators as independent variables to perform the cox proportional regression on four time events (Child-Pugh, death, HCC, decompensation) in the HCV cirrhosis cohort. The p-values were calculated for both analyses respectively, followed by the calculation of FDR for multiple testing correction.

QUANTIFICATION AND STATISTICAL ANALYSIS

The data in most panels reflect multiple experiments performed on different days using mice derived from different litters. Variation in all panels is indicated using standard deviation presented as mean ± SD. Two-tailed unpaired Student's t-tests were used to test the significance of differences between two groups. Statistical significance is displayed as ns (not significant, or p>=0.05), * (p<0.05), ** (p<0.01), *** (p<0.001), **** (p<0.0001) unless specified otherwise. Image analysis for quantification was blinded.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Sam Wang and Andrew Hsieh for constructive comments on the manuscript; J. Shelton (UTSW Histopathology Core), E. Nwoka, and C. Moxon (UT Southwestern Tissue Management Shared Resource) for histopathology; D. Ramirez (UTSW Whole Brain Microscopy Facility, RRID:SCR_017949) for whole section

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

imaging; J. Xu, Y. J. Kim and J. Lyu (CRI Sequencing Core) for sequencing; S. Hacker, A. Walker, and J.I. Gamayot for metabolic phenotyping assays. pAdDeltaF6 and pAAV2/8 plasmids were gifts from James Wilson. T.W. is supported by NIH (R01CA258584). C.L. and the UT Southwestern Tissue Management Shared Resource are supported by the Simmons Comprehensive Cancer Center Support Grant (P30CA142543). Y.H. is supported by NIH (R01CA233794) and Cancer Prevention and Research Institute of Texas (CPRIT grants RR180016). H.Z. is supported by the Pollack Foundation, the NIH (R01AA028791, R01DK125396), CPRIT (RP220614) a Simmons Comprehensive Cancer Center Cancer & Obesity Translational Pilot Award, and the Emerging Leader Award from the Mark Foundation For Cancer Research (#21-003-ELA).

Bibliography

1. Martincorena I, and Campbell PJ (2015). Somatic mutation in cancer and normal cells. Science 349, 1483–1489. 10.1126/science.aab4082. [PubMed: 26404825]
2. Brunner SF, Roberts ND, Wylie LA, Moore L, Aitken SJ, Davies SE, Sanders MA, Ellis P, Alder C, Hooks Y, et al. (2019). Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. Nature 574, 538–542. 10.1038/s41586-019-1670-9. [PubMed: 31645727]
3. Zhu M, Lu T, Jia Y, Luo X, Gopal P, Li L, Odewole M, Renteria V, Singal AG, 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. [PubMed: 30955891]
4. Lawson ARJ, Abascal F, Coorens THH, Hooks Y, O'Neill L, Latimer C, Raine K, Sanders MA, Warren AY, Mahbubani KTA, et al. (2020). Extensive heterogeneity in somatic mutation and selection in the human bladder. Science 370, 75–82. 10.1126/science.aba8347. [PubMed: 33004514]
5. Diehl AM, and Day C (2017). Cause, pathogenesis, and treatment of nonalcoholic steatohepatitis. N. Engl. J. Med 377, 2063–2072. 10.1056/NEJMra1503519. [PubMed: 29166236]
6. Ng SWK, Rouhani FJ, Brunner SF, Brzozowska N, Aitken SJ, Yang M, Abascal F, Moore L, Nikitopoulou E, Chappell L, et al. (2021). Convergent somatic mutations in metabolism genes in chronic liver disease. Nature 598, 473–478. 10.1038/s41586-021-03974-6. [PubMed: 34646017]
7. Buch S, Stickel F, Trépo E, Way M, Herrmann A, Nischalke HD, Brosch M, Rosendahl J, Berg T, Ridinger M, et al. (2015). A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. Nat. Genet 47, 1443–1448. 10.1038/ng.3417. [PubMed: 26482880]
8. Xia M, Chandrasekaran P, Rong S, Fu X, and Mitsche MA (2021). Hepatic deletion of Mboat7 (LPIAT1) causes activation of SREBP-1c and fatty liver. J. Lipid Res 62, 100031. 10.1194/jlr.RA120000856. [PubMed: 32859645]
9. Wuestefeld T, Pesic M, Rudalska R, Dauch D, Longerich T, Kang T-W, Yevsa T, Heinzmann F, Hoenicke L, Hohmeyer A, et al. (2013). A Direct in vivo RNAi screen identifies MKK4 as a key regulator of liver regeneration. Cell 153, 389–401. 10.1016/j.cell.2013.03.026. [PubMed: 23582328]
10. Jia Y, Li L, Lin Y-H, Gopal P, Shen S, Zhou K, Yu X, Sharma T, Zhang Y, Siegwart DJ, et al. (2022). In vivo CRISPR screening identifies BAZ2 chromatin remodelers as druggable regulators of mammalian liver regeneration. Cell Stem Cell 29, 372–385.e8. 10.1016/j.stem.2022.01.001. [PubMed: 35090595]
11. Ye L, Park JJ, Dong MB, Yang Q, Chow RD, Peng L, Du Y, Guo J, Dai X, Wang G, et al. (2019). In vivo CRISPR screening in CD8 T cells with AAV-Sleeping Beauty hybrid vectors identifies membrane targets for improving immunotherapy for glioblastoma. Nat. Biotechnol 37, 1302–1313. 10.1038/s41587-019-0246-4. [PubMed: 31548728]
12. Cooney AL, Singh BK, and Sinn PL (2015). Hybrid nonviral/viral vector systems for improved piggyBac DNA transposon in vivo delivery. Mol. Ther 23, 667–674. 10.1038/mt.2014.254. [PubMed: 25557623]
13. Samuel VT, and Shulman GI (2018). Nonalcoholic fatty liver disease as a nexus of metabolic and hepatic diseases. Cell Metab. 27, 22–41. 10.1016/j.cmet.2017.08.002. [PubMed: 28867301]
14. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M, and Liu XS (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. 15, 554. 10.1186/s13059-014-0554-4. [PubMed: 25476604]
15. Calle RA, Amin NB, Carvajal-Gonzalez S, Ross TT, Bergman A, Aggarwal S, Crowley C, Rinaldi A, Mancuso J, Aggarwal N, et al. (2021). ACC inhibitor alone or co-administered with a DGAT2

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

inhibitor in patients with non-alcoholic fatty liver disease: two parallel, placebo-controlled, randomized phase 2a trials. Nat. Med 27, 1836–1848. 10.1038/S41591-021-01489-1. [PubMed: 34635855]

16. Yenilmez B, Wetoska N, Kelly M, Echeverria D, Min K, Lifshitz L, Alterman JF, Hassler MR, Hildebrand S, DiMarzio C, et al. (2022). An RNAi therapeutic targeting hepatic DGAT2 in a genetically obese mouse model of nonalcoholic steatohepatitis. Mol. Ther 30, 1329–1342. 10.1016/j.ymthe.2021.11.007. [PubMed: 34774753]
17. Wang X, Briggs MR, Hua X, Yokoyama C, Goldstein JL, and Brown MS (1993). Nuclear protein that binds sterol regulatory element of low density lipoprotein receptor promoter. II. Purification and characterization. J. Biol. Chem 268, 14497–14504. [PubMed: 8314806]
18. Briggs MR, Yokoyama C, Wang X, Brown MS, and Goldstein JL (1993). Nuclear protein that binds sterol regulatory element of low density lipoprotein receptor promoter. I. Identification of the protein and delineation of its target nucleotide sequence. J. Biol. Chem 268, 14490–14496. [PubMed: 8390995]
19. Brown MS, and Goldstein JL (1997). The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell 89, 331–340. 10.1016/s0092-8674(00)80213-5. [PubMed: 9150132]
20. Sun XJ, Rothenberg P, Kahn CR, Backer JM, Araki E, Wilden PA, Cahill DA, Goldstein BJ, and White MF (1991). Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein. Nature 352, 73–77. 10.1038/352073a0. [PubMed: 1648180]
21. Kubota N, Kubota T, Itoh S, Kumagai H, Kozono H, Takamoto I, Mineyama T, Ogata H, Tokuyama K, Ohsugi M, et al. (2008). Dynamic functional relay between insulin receptor substrate 1 and 2 in hepatic insulin signaling during fasting and feeding. Cell Metab. 8, 49–64. 10.1016/j.cmet.2008.05.007. [PubMed: 18590692]
22. Kubota N, Kubota T, Kajiwara E, Iwamura T, Kumagai H, Watanabe T, Inoue M, Takamoto I, Sasako T, Kumagai K, et al. (2016). Differential hepatic distribution of insulin receptor substrates causes selective insulin resistance in diabetes and obesity. Nat. Commun 7, 12977. 10.1038/ncomms12977. [PubMed: 27708333]
23. Liang W, Menke AL, Driessen A, Koek GH, Lindeman JH, Stoop R, Havekes LM, Kleemann R, and van den Hoek AM (2014). Establishment of a general NAFLD scoring system for rodent models and comparison to human liver pathology. PLoS ONE 9, e115922. 10.1371/journal.pone.0115922. [PubMed: 25535951]
24. Hoang SA, Oseini A, Feaver RE, Cole BK, Asgharpour A, Vincent R, Siddiqui M, Lawson MJ, Day NC, Taylor JM, et al. (2019). Gene expression predicts histological severity and reveals distinct molecular profiles of nonalcoholic fatty liver disease. Sci. Rep 9, 12541. 10.1038/S41598-019-48746-5. [PubMed: 31467298]
25. Hoshida Y, Villanueva A, Sangiovanni A, Sole M, Hur C, Andersson KL, Chung RT, Gould J, Kojima K, Gupta S, et al. (2013). Prognostic gene expression signature for patients with hepatitis C-related early-stage cirrhosis. Gastroenterology 144, 1024–1030. 10.1053/j.gastro.2013.01.021. [PubMed: 23333348]
26. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102, 15545–15550. 10.1073/pnas.0506580102. [PubMed: 16199517]
27. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, and Mesirov JP (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics 27, 1739–1740. 10.1093/bioinformatics/btr260. [PubMed: 21546393]
28. Wang Z, Miu K-K, Zhang X, Wan AT-Y, Lu G, Cheung H-H, Lee H-M, Kong AP-S, Chan JC-N, and Chan W-Y (2020). Hepatic miR-192-3p reactivation alleviates steatosis by targeting glucocorticoid receptor. JHEP Rep. 2, 100179. 10.1016/j.jhepr.2020.100179. [PubMed: 33134908]
29. Moreira RPP, Gomes LG, Madureira G, Mendonca BB, and Bachega TASS (2014). Influence of the A3669G Glucocorticoid Receptor Gene Polymorphism on the Metabolic Profile of Pediatric Patients with Congenital Adrenal Hyperplasia. Int. J. Endocrinol 2014, 594710. 10.1155/2014/594710. [PubMed: 25050120]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

30. Pizarro M, Solís N, Quintero P, Barrera F, Cabrera D, Rojas-de Santiago P, Arab JP, Padilla O, Roa JC, Moshage H, et al. (2015). Beneficial effects of mineralocorticoid receptor blockade in experimental non-alcoholic steatohepatitis. Liver Int. 35, 2129–2138. 10.1111/liv.12794. [PubMed: 25646700]

31. Wada T, Miyashita Y, Sasaki M, Aruga Y, Nakamura Y, Ishii Y, Sasahara M, Kanasaki K, Kitada M, Koya D, et al. (2013). Eplerenone ameliorates the phenotypes of metabolic syndrome with NASH in liver-specific SREBP-1c Tg mice fed high-fat and high-fructose diet. Am. J. Physiol. Endocrinol. Metab 305, E1415–25. 10.1152/ajpendo.00419.2013. [PubMed: 24129399]

32. Liška F, Landa V, Zídek V, Mlejnek P, Šilhavý J, Šimáková M, Strnad H, Trnovská J, Škop V, Kazdová L, et al. (2017). Downregulation of plzf gene ameliorates metabolic and cardiac traits in the spontaneously hypertensive rat. Hypertension 69, 1084–1091. 10.1161/HYPERTENSIONAHA.116.08798. [PubMed: 28396530]

33. Kutay H, Klepper C, Wang B, Hsu S, Datta J, Yu L, Zhang X, Majumder S, Motiwala T, Khan N, et al. (2012). Reduced susceptibility of DNA methyltransferase 1 hypomorphic (Dnmt1N/+) mice to hepatic steatosis upon feeding liquid alcohol diet. PLoS ONE 7, e41949. 10.1371/journal.pone.0041949. [PubMed: 22905112]

34. Mohs A, Otto T, Schneider KM, Peltzer M, Boekschoten M, Holland CH, Hudert CA, Kalveram L, Wiegand S, Saez-Rodriguez J, et al. (2021). Hepatocyte-specific NRF2 activation controls fibrogenesis and carcinogenesis in steatohepatitis. J. Hepatol 74, 638–648. 10.1016/j.jhep.2020.09.037. [PubMed: 33342543]

35. Ramadori P, Drescher H, Erschfeld S, Schumacher F, Berger C, Fragoulis A, Schenkel J, Kensler TW, Wruck CJ, Trautwein C, et al. (2016). Hepatocyte-specific Keap1 deletion reduces liver steatosis but not inflammation during non-alcoholic steatohepatitis development. Free Radic. Biol. Med 91, 114–126. 10.1016/j.freeradbiomed.2015.12.014. [PubMed: 26698665]

36. Senagolage MD, Sommars MA, Ramachandran K, Futtner CR, Omura Y, Allred AL, Wang J, Yang C, Procissi D, Evans RM, et al. (2018). Loss of transcriptional repression by BCL6 confers insulin sensitivity in the setting of obesity. Cell Rep. 25, 3283–3298.e6. 10.1016/j.celrep.2018.11.074. [PubMed: 30566857]

37. Chikada H, Ida K, Nishikawa Y, Inagaki Y, and Kamiya A (2020). Liver-specific knockout of B cell lymphoma 6 suppresses progression of non-alcoholic steatohepatitis in mice. Sci. Rep 10, 9704. 10.1038/s41598-020-66539-z. [PubMed: 32546802]

38. Sommars MA, Ramachandran K, Senagolage MD, Futtner CR, Germain DM, Allred AL, Omura Y, Bederman IR, and Barish GD (2019). Dynamic repression by BCL6 controls the genome-wide liver response to fasting and steatosis. eLife 8. 10.7554/eLife.43922.

39. Willmer T, Cooper A, Peres J, Omar R, and Prince S (2017). The T-Box transcription factor 3 in development and cancer. Biosci. Trends 11, 254–266. 10.5582/bst.2017.01043. [PubMed: 28579578]

40. Nikkanen J, Leong YA, Krause WC, Dermadi D, Maschek JA, Van Ry T, Cox JE, Weiss EJ, Gokcumen O, Chawla A, et al. (2022). An evolutionary trade-off between host immunity and metabolism drives fatty liver in male mice. Science 378, 290–295. 10.1126/science.abn9886. [PubMed: 36264814]

41. Ferguson AD, Larsen NA, Howard T, Pollard H, Green I, Grande C, Cheung T, Garcia-Arenas R, Cowen S, Wu J, et al. (2011). Structural basis of substrate methylation and inhibition of SMYD2. Structure 19, 1262–1273. 10.1016/j.str.2011.06.011. [PubMed: 21782458]

42. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, Boerwinkle E, Cohen JC, and Hobbs HH (2008). Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat. Genet 40, 1461–1465. 10.1038/ng.257. [PubMed: 18820647]

43. Abul-Husn NS, Cheng X, Li AH, Xin Y, Schurmann C, Stevis P, Liu Y, Kozlitina J, Stender S, Wood GC, et al. (2018). A Protein-Truncating HSD17B13 Variant and Protection from Chronic Liver Disease. N. Engl. J. Med 378, 1096–1106. 10.1056/NEJMoa1712191. [PubMed: 29562163]

44. Cohen JC, Boerwinkle E, Mosley TH, and Hobbs HH (2006). Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N. Engl. J. Med 354, 1264–1272. 10.1056/NEJMoa054013. [PubMed: 16554528]

45. Tsuchida T, Lee YA, Fujiwara N, Ybanez M, Allen B, Martins S, Fiel MI, Goossens N, Chou HI, Hoshida Y, et al. (2018). A simple diet- and chemical-induced murine NASH model with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

rapid progression of steatohepatitis, fibrosis and liver cancer. J. Hepatol 69, 385–395. 10.1016/j.jhep.2018.03.011. [PubMed: 29572095]

46. Zhu H, Shah S, Shyh-Chang N, Shinoda G, Einhorn WS, Viswanathan SR, Takeuchi A, Grasemann C, Rinn JL, Lopez MF, et al. (2010). Lin28a transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. Nat. Genet 42, 626–630. 10.1038/ng.593. [PubMed: 20512147]
47. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol 34, 184–191. 10.1038/nbt.3437. [PubMed: 26780180]
48. Canver MC, Haeussler M, Bauer DE, Orkin SH, Sanjana NE, Shalem O, Yuan G-C, Zhang F, Concordet J-P, and Pinello L (2018). Integrated design, execution, and analysis of arrayed and pooled CRISPR genome-editing experiments. Nat. Protoc 13, 946–986. 10.1038/nprot.2018.005. [PubMed: 29651054]
49. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. 10.1093/bioinformatics/bts635. [PubMed: 23104886]
50. Liao Y, Smyth GK, and Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. 10.1093/bioinformatics/btt656. [PubMed: 24227677]
51. Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]
52. Frankish A, Diekhans M, Ferreira A-M, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J, Armstrong J, et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. 47, D766–D773. 10.1093/nar/gky955. [PubMed: 30357393]
53. Topol A, Zhu S, Hartley BJ, English J, Hauberg ME, Tran N, Rittenhouse CA, Simone A, Ruderfer DM, Johnson J, et al. (2016). Dysregulation of miRNA-9 in a Subset of Schizophrenia Patient-Derived Neural Progenitor Cells. Cell Rep. 15, 1024–1036. 10.1016/j.celrep.2016.03.090. [PubMed: 27117414]
54. Friedman RC, Farh KK-H, Burge CB, and Bartel DP (2009). Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 19, 92–105. 10.1101/gr.082701.108. [PubMed: 18955434]
55. Balwierz PJ, Pachkov M, Arnold P, Gruber AJ, Zavolan M, and van Nimwegen E (2014). ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. Genome Res. 24, 869–884. 10.1101/gr.169508.113. [PubMed: 24515121]
56. Conlon EM, Liu XS, Lieb JD, and Liu JS (2003). Integrating regulatory motif discovery and genome-wide expression analysis. Proc Natl Acad Sci USA 100, 3339–3344. 10.1073/pnas.0630591100. [PubMed: 12626739]
57. Wasserman WW, and Sandelin A (2004). Applied bioinformatics for the identification of regulatory elements. Nat. Rev. Genet 5, 276–287. 10.1038/nrg1315. [PubMed: 15131651]
58. Matys V, Fricke E, Geffers R, Gössling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, et al. (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. Nucleic Acids Res. 31, 374–378. 10.1093/nar/gkg108. [PubMed: 12520026]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights:

1. Clones with *Mboat7* loss of function mutations are selected against in NASH livers.
2. Clones with *Gpam* loss of function mutations are positively selected for in NASH.
3. *In vivo* screening was built to identify genes that alter cell fitness and steatosis.
4. Screening of transcription and epi-factors identified unexpected NASH target genes.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

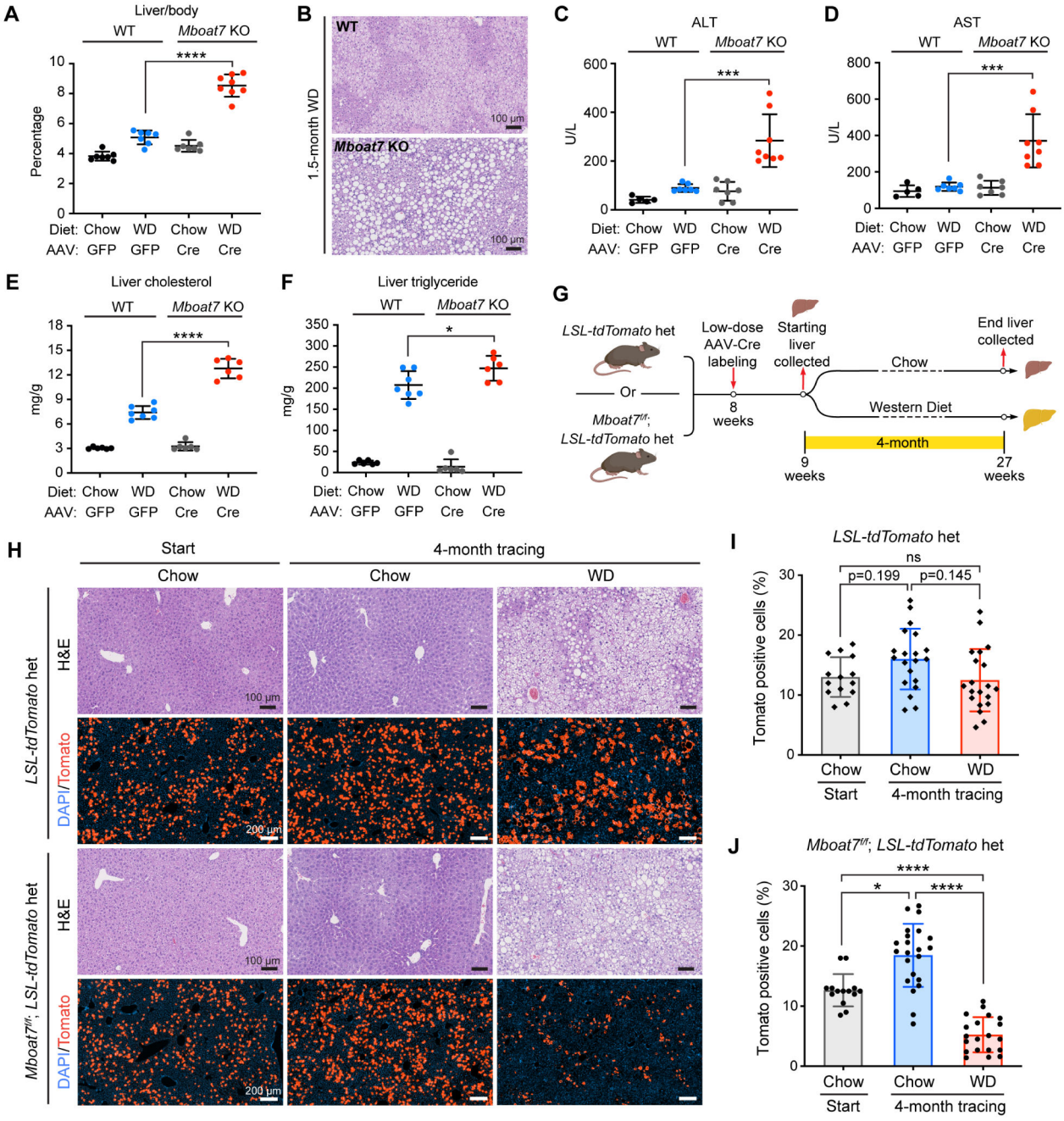


Figure 1. *Mboat7* loss of function mutations, which caused lipid accumulation, led to decreased clonal fitness.

A. Liver/body weight ratios of liver-specific *Mboat7* WT and KO mice fed with 1.5 months of chow or WD (n = 7, 7, 7, 8 mice for each group). These mice were given high doses of AAV8-TBG-Cre to generate liver-wide *Mboat7* deletion in almost all hepatocytes.

B. Representative H&E staining of *Mboat7* WT and KO liver sections after 1.5 months of WD.

C-D. Liver function testing with plasma AST and ALT (n = 5, 7, 7, 8 mice for each group).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

E-F. Cholesterol and triglyceride measurements from liver tissues (n = 6, 7, 6, 6 mice for each group).

G. Schema for the mosaic *Mboat7* lineage tracing experiment. *LSL-tdTomato* het or *Mboat7*$^{f/f}$*; LSL-tdTomato* het mice were injected with a low dose of AAV8-TBG-Cre to generate mosaic Tomato+ WT hepatocytes in control mice, or Tomato+ *Mboat7* mutant hepatocytes in experimental mice. These mice were then fed with either chow or WD for 4 months. Livers were collected one week after AAV8-TBG-Cre and 4 months after chow or WD was initiated.

H. Representative H&E and fluorescent images of liver sections at the beginning and end of lineage tracing.

I. Quantification of Tomato+ cells from *LSL-tdTomato* het liver sections in **H** (n = 7, 10, 10 mice for each group).

J. Quantification of Tomato+ cells from *Mboat7*$^{f/f}$*; LSL-tdTomato* het liver sections in **H** (n = 7, 11, 10 mice for each group). Each dot in **I** and **J** represents one image field; two fields from each mouse liver are shown. Statistical analysis in **I** and **J** were performed on averaged image data from individual mice.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

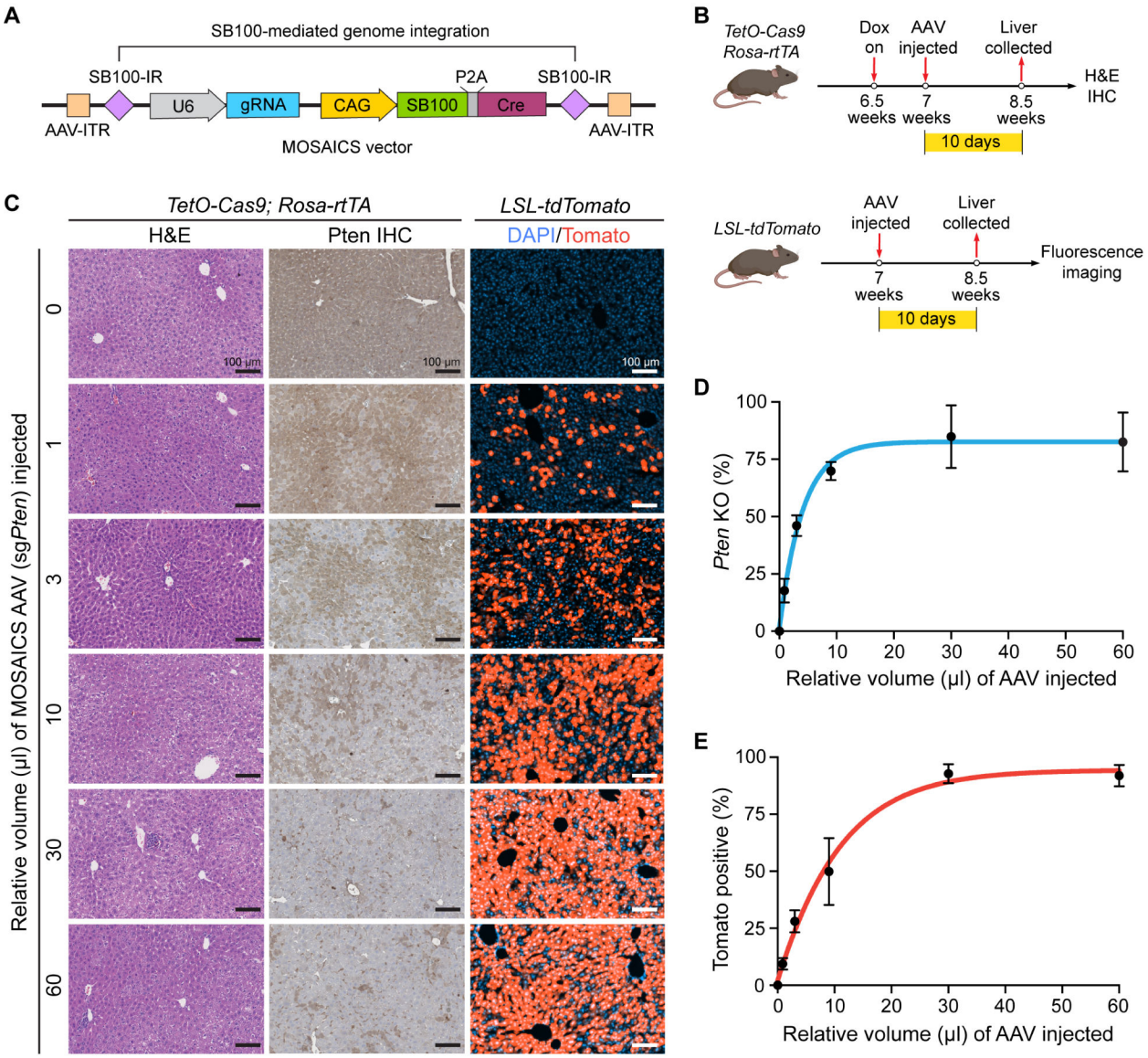


Figure 2. MOSAICS platform for generating and tracing somatic mutations *in vivo*.

A. MOSAICS AAV delivery, gene perturbation, and sgRNA tracing. The U6-sgRNA element and CAG promoter-driven SB100-P2A-Cre fusion protein are flanked by inverted repeat (IR) sequences, enabling SB100 transposase mediated genomic integration.

B. Schema for functional validation of the MOSAICS platform. The MOSAICS AAV carrying a *Pten* sgRNA was IV injected into dox-inducible Cas9 expressing mice (*TetO-Cas9* homo; *Rosa-rtTA* homo) to query the generation of *Pten* deficient hepatocytes. The same AAV was also injected into *LSL-tdTomato* homo mice to test SB100-P2A-Cre fusion protein expression.

C. Liver sections from the validation mice described in B. H&E and PTEN IHC staining showed that the frequency of PTEN deficient hepatocytes correlated with the amount of AAV injected. Fluorescent images of *LSL-tdTomato* liver sections showed that the frequency of hepatocytes expressing Tomato correlated with the amount of AAV injected.

D. Quantification of *Pten* KO cells shown in **C** (n = 3 mice for each AAV concentration).

E. Quantification of Tomato+ cells shown in **C** (n = 3 mice for each AAV concentration).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

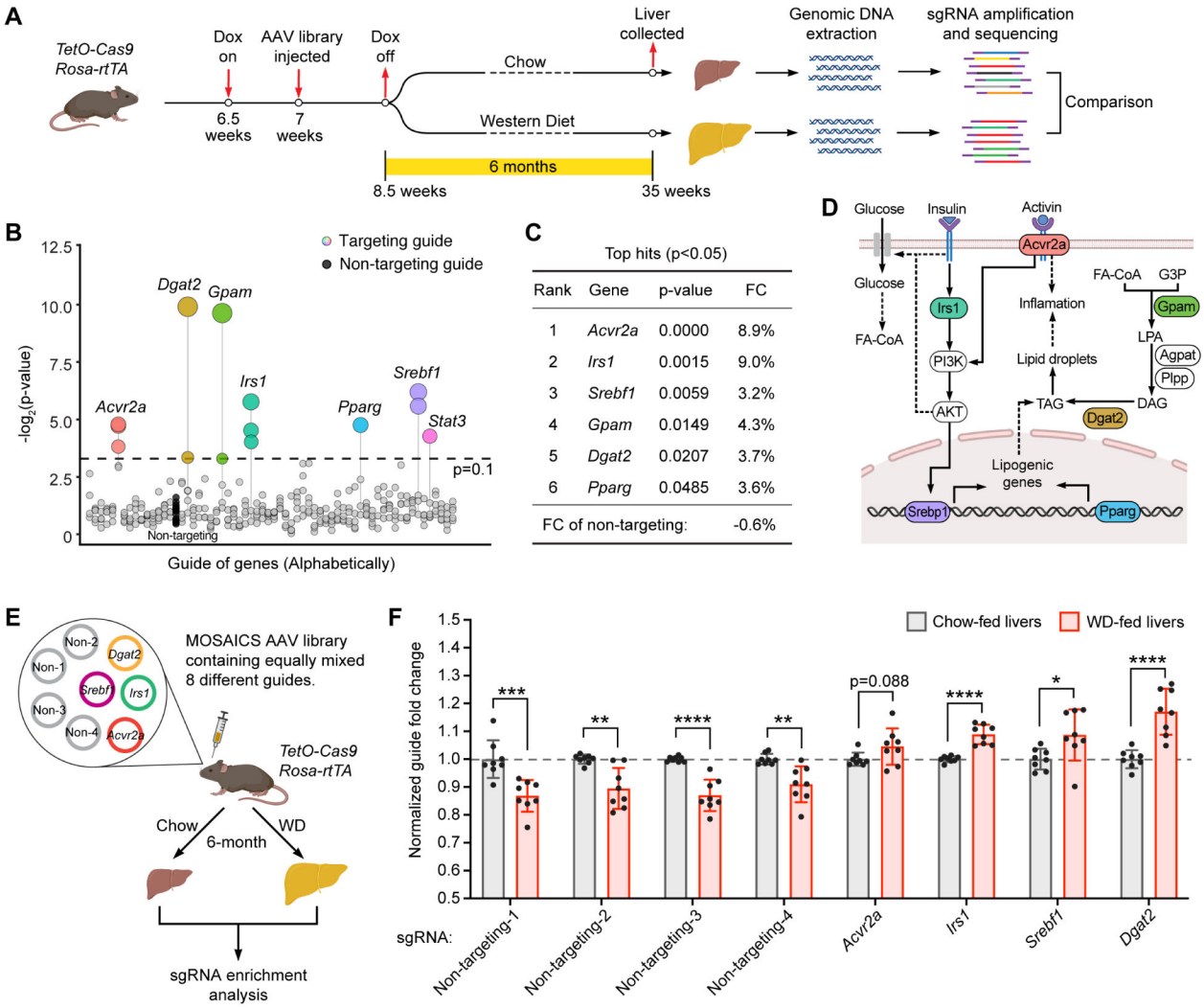


Figure 3. Lineage tracing of mosaic mutant hepatocytes demonstrated that mutations that suppress lipogenesis are positively selected in fatty livers.

A. Schema for pooled tracing of mutant clones under different dietary conditions. MOSAICS AAVs carrying sgRNA libraries were injected into Cas9 expressing mice. Ten days after gene perturbation, Cas9 was turned off by dox withdrawal, and chow or WD was given to mice for 6 months. Genomic DNA was extracted and sgRNA sequences were amplified and quantified. sgRNAs that were enriched in fatty but not in normal livers were investigated.

B. sgRNAs enriched in WD but not in chow fed livers. Each circle represents one sgRNA. Different sgRNAs targeting the same genes were aligned vertically. Circle sizes correlate to $-\log_2(p)$. Control sgRNAs were drawn as filled black circles.

C. Genes associated with sgRNAs enriched in WD fed mice ($p < 0.05$). See Table S2 for raw data.

D. Pathways in which the enriched genes (listed in **C**) are involved.

E. Validation of the MOSAICS platform for the most positively selected mutations. Four non-targeting sgRNAs and the sgRNAs targeting *Acvr2a, Irs1, Srebf1* and *Dgat2* (Table S6) were cloned into the MOSAICS vector. The eight vectors were mixed before being used for

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

AAV production. Cas9 mice were injected with the 8-sgRNA AAV library and treated with chow or WD for 6 months. Genomic DNA was extracted and sgRNAs were sequenced.
F. Normalized sgRNA reads in chow or WD fed livers (n = 8 and 8 mice for each group). See Table S2 for raw data.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

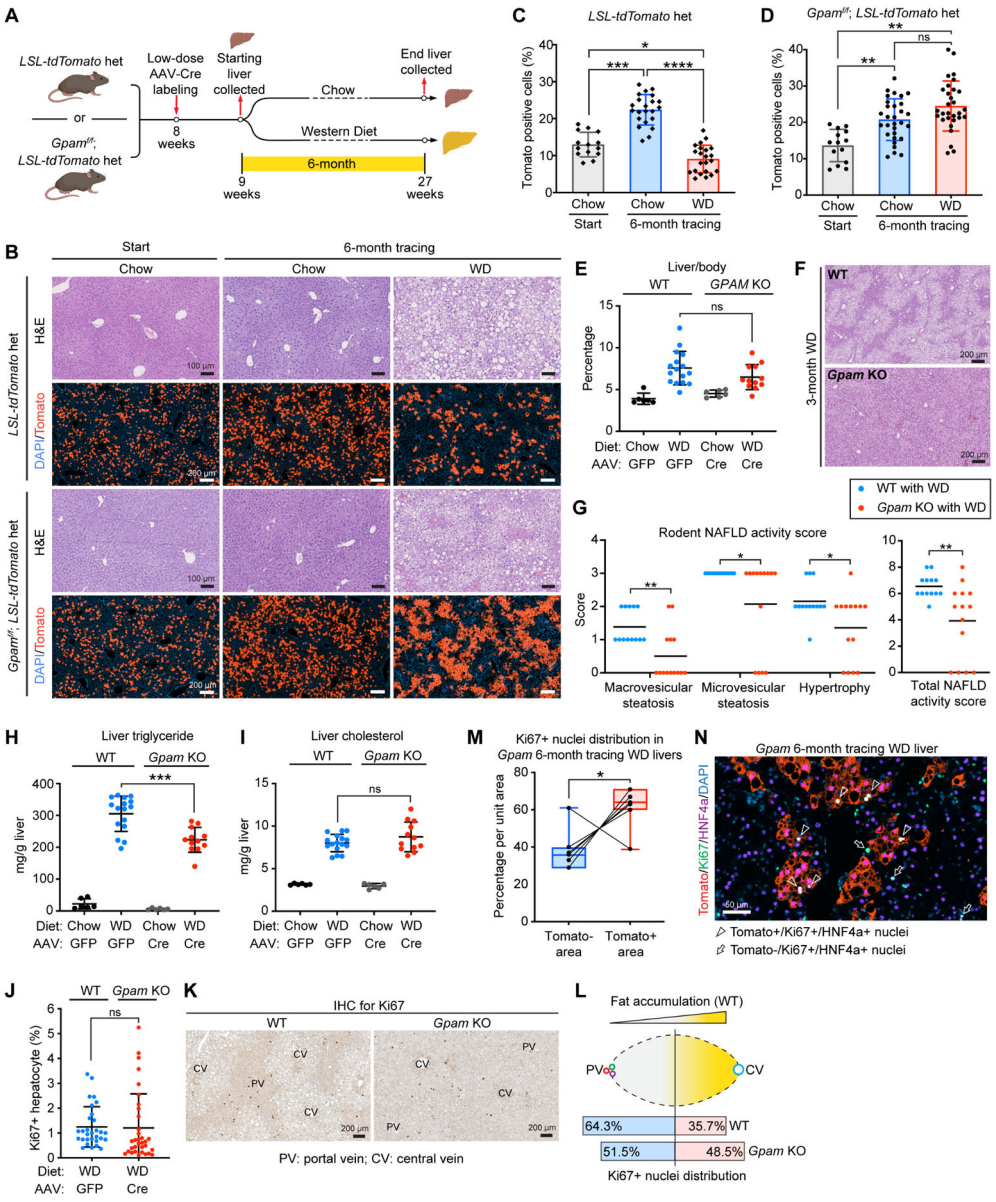


Figure 4. *Gpam* loss of function mutations, which suppressed lipogenesis, led to increased clonal fitness.

A. Schema for the mosaic *Gpam* lineage tracing experiment. *LSL-tdTomato* het mice or *Gpam*$^{f/f}$*; LSL-tdTomato* het mice were injected with a low dose of AAV8-TBG-Cre to generate mosaic Tomato+ hepatocytes in control mice, or Tomato+ and *Gpam* mutant hepatocytes in experimental mice. These mice were then fed with either chow or WD for 6 months. Livers were collected one week after AAV injection and 6 months after chow or WD was initiated.

B. Representative H&E and fluorescent images of liver sections at the beginning and end of lineage tracing.

C. Quantification of Tomato+ cells from *LSL-tdTomato* het liver sections in **B** (n = 7, 11, 11 mice for each group). Each dot represents one image field; two fields from each mouse liver were analyzed. The same time zero group of mice was used in Figure 1I.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

D. Quantification of Tomato+ cells from *Gpam*$^{f/f}$*; LSL-tdTomato* het liver sections in **B** (n = 7, 15, 15 mice for each group). Each dot represents one image field; two fields from each mouse liver were shown. Statistical analysis in **C** and **D** were performed on averaged image data from individual mice.

E. Liver/body weight ratios of liver-specific *Gpam* WT and KO mice fed with 3 months of WD (n = 6, 16, 6, 12 mice for each group). These mice were given high doses of AAV8-TBG-Cre to generate liver-wide *Gpam* deletion in almost all hepatocytes.

F. Representative H&E staining of *Gpam* WT and KO liver sections after 3 months of WD.

G. NAFLD activity score of the H&E sections in **F**. The right panel represents the total NAFLD activity score, which is the sum of the three scores on the left (n = 13 and 14 mice for each group).

H-I. Triglyceride and cholesterol measurements from liver tissues (n = 6, 16, 6, 12 for each group).

J. Quantification of Ki67+ hepatocytes from liver sections of *Gpam* WT and KO mice fed with 3 months of WD (n = 10 and 11). Each dot represents one image field; three fields from each mouse liver were analyzed. Statistical analysis was performed on averaged image data from individual mice.

K. Representative Ki67 IHC staining of *Gpam* WT and KO liver sections after 3 months of WD.

L. Distribution of Ki67+ hepatocytes in the portal vein (PV) half and central vein (CV) half of the lobule.

M. Quantification of Ki67+ hepatocytes from IF staining of *Gpam* mosaic liver sections from **B** (n = 7). 60-130 Ki67+/HNF4a+ nuclei were counted per liver depending on the abundance of Ki67+ nuclei in each liver section. In each randomly selected area that was analyzed, all Ki67+ nuclei were counted. The percentage of Tomato-/Ki67+/HNF4a+ nuclei and Tomato+/Ki67+/HNF4a+ nuclei per unit area were calculated. The sum of these two groups of nuclei was defined as 100%.

N. Representative IF co-staining of Tomato, Ki67, and HNF4a in *Gpam* mosaic liver sections from **B**.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

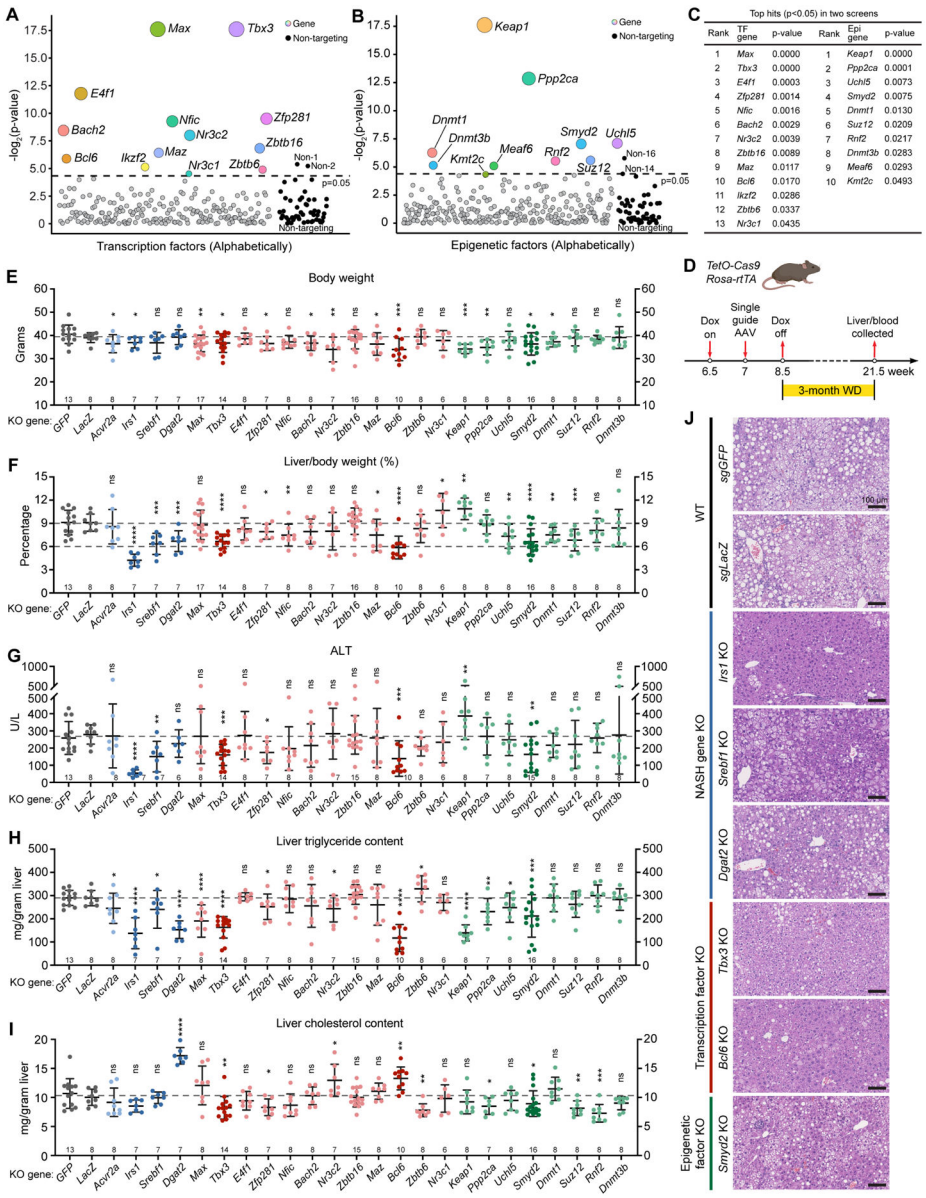


Figure 5. Somatic mosaic screening of transcription and epigenetic factors identified putative therapeutic targets for NASH.

A. Results for transcription factor screening. MOSAICS vectors carrying transcription factor targeting sgRNA libraries were injected into Cas9 expressing mice. Chow or WD was fed to mice for 6 months. The genes corresponding to enriched sgRNAs in WD fed but not chow fed livers were drawn as colored circles with sizes correlating to $-\log_2(p)$. Control sgRNAs were drawn as filled black circles.

B. Results for epigenetic factor screening. The methods and color scheme are the same as in **A**.

C. List of the genes corresponding to enriched sgRNAs ($p < 0.05$) in both MOSAICS screens.

D. How we assess KO phenotypes of the top genes under WD conditions. A MOSAICS AAV carrying an individual sgRNA was injected into Cas9 mice such that each mouse had

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

one gene deleted in the liver. Dox was withdrawn 10 days after AAV injection and WD was given for 3 months before sacrifice.

E-F. Body weight and liver/body ratios of control (sg*GFP* and sg*LacZ*) and liver-specific KO mice fed with 3 months of WD. Gray dots represent control mice, blue dots represent liver-specific KO mice for known NASH genes, red dots represent transcription factor KO mice, and green dots represent epigenetic factor KO mice. Darker dots represent mice that have the most significant differences in liver/body weight ratios. Each dot represents one mouse, and the n is denoted at the bottom of each plot.

G. Liver function testing using plasma ALT. The color scheme is the same as in **E**. The n is denoted at the bottom of the plot.

H-I. Liver triglyceride and cholesterol analysis.

J. Representative H&E images of liver sections are shown for the mice described in **E**.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

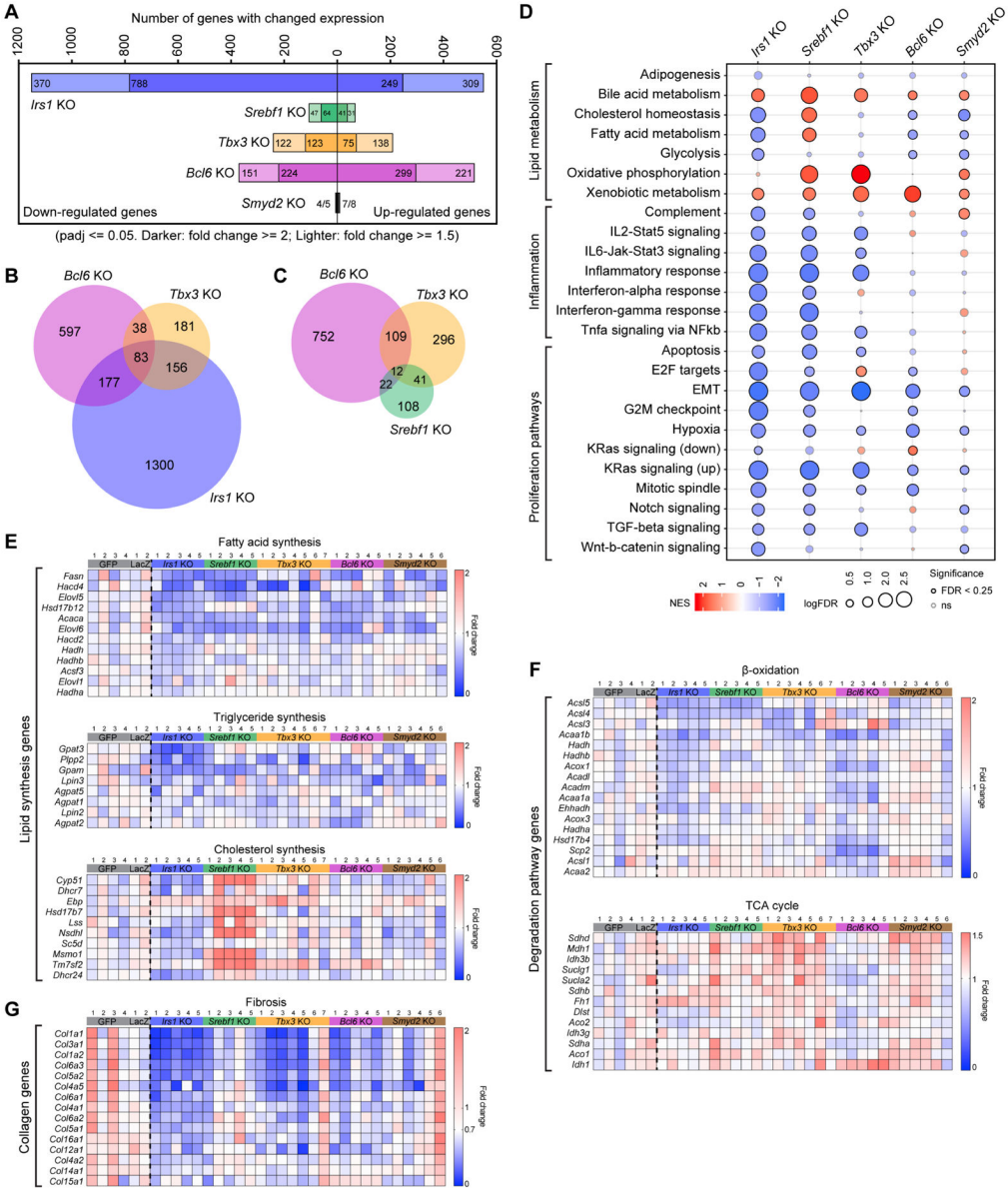


Figure 6. Transcriptional analysis of *Irs1*, *Srebf1*, *Bcl6*, *Tbx3*, and *Smyd2* KO livers after WD.

A. The number of genes with altered expression in the RNA-seq data when comparing control (sg*GFP* and sg*LacZ*) and KO livers after 3 months of WD. Darker and lighter colored bars represent the number of differentially expressed genes with a fold change of >=2 and >=1.5, respectively. Genes with statistically significant fold change differences of less than 1.5 were not included here.

B. Venn diagram showing the shared and unique gene numbers with changed expression (fold change >=1.5) in *Bcl6, Tbx3,* and *Irs1* KO livers.

C. Venn diagram showing the shared and unique gene numbers with changed expression (fold change >=1.5) in *Bcl6, Tbx3*, and *Srebf1* KO livers.

D. Hallmark pathway enrichment analysis of RNA-seq data from **A**.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

E. Heatmaps showing the fold changes of differentially expressed genes in fatty acid, triglyceride, and cholesterol synthesis pathways. The average expression levels of control samples (four sg*GFP* and two sg*LacZ*) were normalized to 1 for each gene.
F. Heatmaps showing the fold changes of differentially expressed genes in β-oxidation and TCA cycle pathways. The normalization method is the same as in **E**.
G. Heatmaps showing the fold changes of differentially expressed collagen genes. The normalization method is the same as in **E**.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

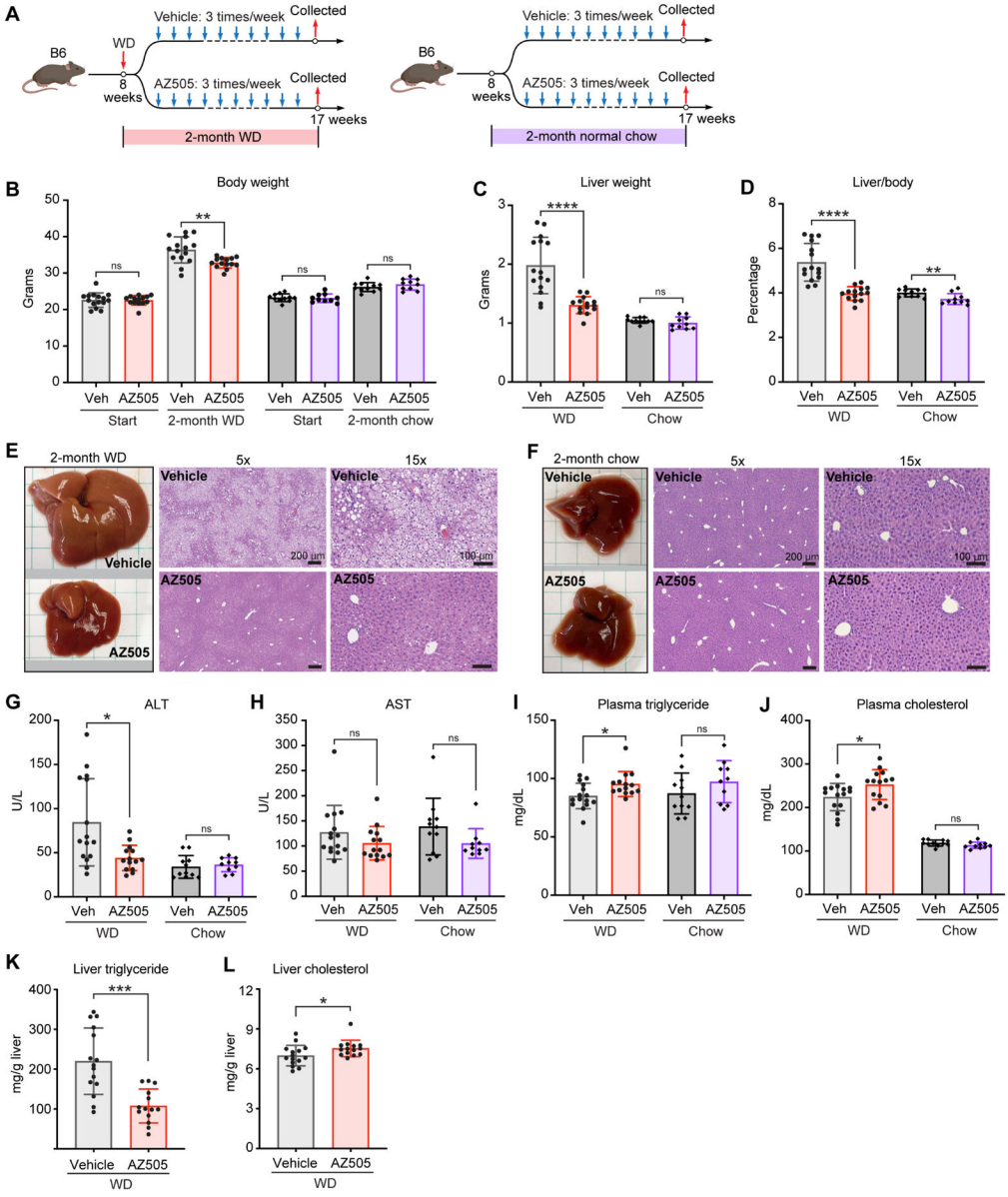


Figure 7. AZ505, a selective SMYD2 inhibitor, ameliorated fatty liver disease.

A. Pre-clinical testing of AZ505 in NASH models. WD or chow was given at 8 weeks of age for 2 months. AZ505 treatment started one day after WD or chow was initiated. Vehicle (2.5% DMSO in saline) or AZ505 (10 mg/kg) was given to mice intraperitoneally 3 times per week.

B-D. Body weight, liver weight and liver/body weight ratios of vehicle or AZ505 treated mice before and after 2 months of WD or chow (n = 15 and 14 mice on WD, light gray and red bars; n = 11 and 10 mice on chow, dark gray and purple bars).

E-F. Representative liver pictures and H&E of liver sections for vehicle or AZ505 treated mice after 2 months of WD or chow.

G-H. Liver function analysis using plasma ALT and AST (n = 15 and 14 mice on WD; n = 11 and 10 mice on chow).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

I-J. Plasma triglyceride and cholesterol analysis (n = 15 and 14 mice on WD; n = 11 and 10 mice on chow).
K-L. Liver triglyceride and cholesterol analysis of the mice on WD (n = 15 and 14).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Pten	Cell Signaling	#9559; RRID: AB_390810
Ki67 (for IHC)	Abcam	#AB15580; RRID: AB_443209
Ki67 (for IF)	Invitrogen	#14-5698-82; RRID: AB_10854564
RFP	Rockland	#600-401-379; RRID: AB_2209751
HNF4a	Abcam	#ab41898; RRID: AB_732976
Goat anti-rat IgG (H&L) Alexa Fluor Plus 488	Invitrogen	#A-48262; RRID: AB_2896330
Donkey anti-rabbit IgG (H&L) Alexa Fluor 594	Invitrogen	#A-21207; RRID: AB_141637
Goat anti-mouse IgG2a Alexa Fluor 647	Invitrogen	#A-21241; RRID: AB_2535810
Bacterial and virus strains		
Stellar Competent Cells	Takara	#636766
E. cloni 10G ELITE Electrocompetent Cells	Lucigen	#60052-2
AAV8-TBG-Cre	Addgene	#107787
AAV8-TBG-GFP	Addgene	#105535
Chemicals, peptides, and recombinant proteins		
DMEM	Thermo Fisher	#SH30022FS
Opti-MEM	Life Technologies	#31985070
FBS	Sigma Aldrich	#F0926
D-(+)-Glucose	Sigma Aldrich	#G8270
D-(−)-Fructose	Sigma Aldrich	#F0127
Sucrose	Sigma Aldrich	#S8501
D-Sorbitol	Sigma Aldrich	#85529
Formalin	Fisher scientific	#245-685
PEI	ChemCruz	#sc-360988
CHAPS hydrate	Sigma Aldrich	#C0570
AMPure XP	Beckman Coulter	#NC9933872
Cryo-Gel	Leica	#39475237
PORO CaptureSelect AAV8 Affinity Resin	Thermo Fisher	#A30789
Proteinase K	Roche	#03115828001
RNase A	Invitrogen	#12091-021
Doxycycline	Fisher scientific	#50550400
AZ505	MCE	#HY-15226
0.9% sodium chloride	Baxter	#2F7123
TRIzol	Thermo Fisher	#15596018
Critical commercial assays		
In Situ Cell Death Detection Kit, Fluorescein	Roche	#C755B40

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Picro Sirius Red Staining Kit	Abcam	#ab150681
ABC-HRP Kit	Vector laboratories	#PK-6101
DAB Kit	Vector laboratories	#SK-4100
Qubit™ dsDNA HS and BR Assay Kits	Invitrogen	#Q32853
Blood & Cell Culture DNA Midi Kit	Qiagen	#13343
QIAquick PCR Purification Kit	Qiagen	#28104
QIAquick Gel Extraction Kit	Qiagen	#28704
HiSpeed Plasmid Maxi Kit	Qiagen	#12663
RNeasy Mini kit	Qiagen	#74014
DNA Assembly Kit	NEB	#E5520A
SMARTer Stranded Total RNA Sample Prep Kit - HI Mammalian	Takara	#634875
General Chemistry Aspartate Aminotransferase (AST) For VITROS 250/950	Ortho Clinical Diagnostics	#8433815
General Chemistry Alanine Aminotransferase (ALT) For VITROS 250/950	Ortho Clinical Diagnostics	#1655281
General Chemistry Cholesterol For VITROS 250/950	Ortho Clinical Diagnostics	#1669829
General Chemistry Triglycerides For VITROS 250/950	Ortho Clinical Diagnostics	#1336544
VITROS Chemistry Products Calibrator Kit 2	VITROS Chemistry	#1662659
VITROS Chemistry Products Calibrator Kit 3	VITROS Chemistry	#1290709
VITROS Chemistry Products Performance Verifier I	VITROS Chemistry	#8067324
VITROS Chemistry Products Performance Verifier II	VITROS Chemistry	#8231474
Infinity™ Cholesterol Liquid Stable Reagent	Thermo Scientific	#TR13421
Infinity™ Triglycerides Liquid Stable Reagent	Thermo Scientific	#TR22421
Matrix plus™ Chemistry Reference Kit	Verichem Laboratories	#NC9592194
Matrix plus™ Chemistry Reference Kit	Verichem Laboratories	#NC9592194
Deposited data		
mRNA seq on Cas9 KO mice	This study	GSE212916
Gene expression data from NAFLD/NASH patients	Hoang et al., 2019	GSE130970
Gene expression data from HCV cirrhosis patients	Hoshida et al., 2013	GSE15654
Experimental models: Cell lines		
AAV-Pro 293T	Takara	#632273
Experimental models: Organisms/strains		
Mboat7 f/f: C57BL/6N-*Mboat7*$^{tm1a(KOMP)Wtsi/Mmucd}$	KOMP	Design ID: 45294
Gpam f/f	This study	N/A
LSL-tdTomato: B6.Cg-*Gt(ROSA)26Sor*$^{tm14(CAG-tdTomato)Hze}$/J	The Jackson Lab	#007914
Rosa-rtTA; TetO-Cas9 : B6;129S4-*Gt(ROSA)26Sor*$^{tm1(rtTA*M2)Jae}$*Col1a1*$^{tm1(tetO-cas9)Sho}$/J	The Jackson Lab	#029415
Recombinant DNA		
MOSAICS vector	This study	N/A

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pAAV2/8	Addgene	#112864
pAdDeltaF6	Addgene	#112867
Software and algorithms		
QuPath	Qupath (v0.3.2)	https://qupath.github.io/
ImageJ	ImageJ (v1.53a)	https://imagej.net/ij/
cutadapt	cutadapt (v1.9.1)	https://cutadapt.readthedocs.io/en/stable/
MAGeCK	Li et al., 2014	https://sourceforge.net/p/mageck/wiki/Home/
FeatureCounts	Liao et al., 2014	https://subread.sourceforge.net/
DEseq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
R package fgsea	fgsea (v1.14.0)	https://bioconductor.org/packages/release/bioc/html/fgsea.html
Other		
Poly-Prep Column	Bio-Rad	#731-1550
Amicon Ultra Centrifugal Filter Unit	Millipore	#UFC810024