bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Liver enhancer signature and regulatory network of non-alcoholic steatohepatitis resistance in humans with obesity

Yu-Han Hung[1], Ramja Sritharan[1], Marie-Claude Vohl[2], Olga Ilkayeva[3], Laurent Biertho[4], André Tchernof[4], Phillip J. White[3*], Praveen Sethupathy[1*]

[1]Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA
[2] Institute of Nutrition and Functional Foods (INAF), Université Laval, Québec City, QC, Canada
[3] Duke Molecular Physiology Institute, Duke University Medical Center, Division of Endocrinology and, Department of Medicine, and Department of Pharmacology and Cancer Biology, Durham NC USA
[4]Centre de recherche de l'Institut Universitaire de Cardiologie et de Pneumologie de Quebec-Universite Laval, Quebec, QC, Canada

*co-corresponding authors

Abstract

Non-alcoholic fatty liver disease (NAFLD), which often co-occurs with obesity and type 2 diabetes, is the most common cause of chronic liver disease worldwide. A subset of patients with NAFLD progress to the severe form known as non-alcoholic steatohepatitis (NASH), which increases the risk of developing hepatic fibrosis, cirrhosis, and hepatocellular carcinoma. The molecular underpinnings of the progression from NAFLD to NASH in patients is poorly understood. Active enhancer landscapes are known to determine cell states and can point to key transcription factors (TF), genes, and pathways in disease pathogenesis or resistance. Also, while super-enhancers have helped reveal key disease drivers in several cancer types, they remain undefined in NASH. To define the enhancer signature of NASH-prone (NP) and NASH-resistant (NR) phenotypes in humans, we performed chromatin run-on sequencing (ChRO-seq) analysis on liver biopsies of individuals with severe obesity who were stratified into either the NP or NR group. We defined the active enhancer signature, super-enhancer linked genes, and the candidate TF networks in human NP and NR livers. Notably, we showed that NR-activated genes that are linked to NR-specific enhancers/super-enhancers are involved in serine and glycine biosynthesis (*SHMT*, *BHMT*) as well as glycine utilization (*GLYAT*, *GATM*). Overall, this study has defined for the first time the active enhancer/super-enhancer landscape and the underlying TF programs of NASH resistance in humans with obesity.

Key words: non-alcoholic steatohepatitis (NASH), enhancer, super-enhancer, glycine metabolism

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is the most common cause of chronic liver disease, affecting ~25% of the population world-wide[1,2]. NAFLD is highly associated with many metabolic disorders including obesity, type 2 diabetes (T2D), and hyperlipidemia. Numerous risk factors have been identified, including body mass index (BMI), glycemic status, age, smoking, and specific genetic variants (such as rs738409 at *PNPLA3* locus)[3-5]. Clinically, NAFLD is characterized by abnormal accumulation of lipids in hepatocytes. In 10-20% of cases, NAFLD progresses to a more severe form known as non-alcoholic steatohepatitis (NASH)[2], which is marked by increased steatosis, hepatocyte ballooning, and lobular inflammation[6,7], which sets the stage for the development of fibrosis, and in some cases eventually cirrhosis, hepatocellular carcinoma, or liver failure requiring transplantation[8]. While NASH is the second leading cause and will soon be the primary cause for liver transplantation in the United States[9,10], there are currently no FDA-approved drugs for managing NASH[11]. Many drugs that have shown promise for ameliorating NASH in animal models have failed in human clinical trials[12], which speaks to the urgent need for research with human tissue that aims to discover new potential therapeutic targets for NASH.

While gene expression profiles related to NAFLD progression have been reported previously in several studies[13-18], the underlying transcriptional programming and the enhancer landscape that shapes gene expression in this disease remains completely unexplored. Regulatory elements (e.g., promoters and enhancers), which interact with transcription factors (TFs) to modulate activity of nearby genes, are critical for many fundamental biological processes including disease pathogenesis. The active enhancer landscape, in particular, most accurately distinguishes different cell states[19,20]. Furthermore, dense clusters of enhancers with remarkably high activity (known as super-enhancers) have been shown to mark genes that are critical for cell identity during development or cell behavior in disease pathogenesis[21,22]. Most of the current knowledge about TFs in NAFLD stems from animal models that may not fully recapitulate human NAFLD/NASH pathogenesis[23-25]. It is unknown how the active enhancer landscape and TF networks are re-wired during NASH development in humans. Also, while super-enhancers have helped identify key disease drivers in a number of different cancer types, they remain undefined and merit further investigation in NASH.

The newest generation of nascent RNA sequencing technology developed to study chromatin regulatory dynamics[26], chromatin run-on sequencing (ChRO-seq), provides an exceptional tool for characterizing the enhancer landscape and TF networks in human NASH. ChRO-seq not only provides direct measurements of gene transcription activity but also enables sensitive detection of active transcriptional regulatory elements (TREs, such as promoters and enhancers). Moreover, ChRO-seq detects TREs that are active, further distinguishing it from another widely-used method, ATAC-seq, which identifies all open chromatin loci, active or not. Also, in contrast to ChRO-seq, open chromatin profiling methods do not provide quantitative information about gene transcription. A major advantage of ChRO-seq relative to previous versions, known as GRO-seq and PRO-seq, is that it can be applied to archived, frozen tissues making it available for profiling human liver biopsy specimens. As an example of the utility of this method, recent studies have leveraged ChRO-seq to define transcriptional programs that underpin human cancer pathogenesis[26,27] as well as gut development in a human organoid model[28].

In this study, we applied ChRO-seq on human liver biopsies from two well-matched populations of individuals with severe obesity that are discordant for the NASH phenotype to define the active enhancer signature, super-enhancer linked genes, and the candidate TF networks that are associated with NASH-prone (NP) and NASH-resistant (NR) livers.

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Results

NASH resistant (NR) and NASH prone (NP) livers are associated with distinct profiles of gene transcription by RNA polymerases

Many NAFLD studies in humans employ non-obese individuals who have lower BMI, more favorable metabolic profiles and/or younger age as controls[18,29-32]. While such studies provide valuable information with translational implications, the presence of multiple confounders presents notable challenges for identifying factors specifically associated with NAFLD progression. Here we set out to study the enhancer landscape of NASH in humans using a NAFLD comparison population that is matched for the major risk factors. Specifically, we first identified six individuals who have severe obesity (BMI >45) without the presence of NASH. This cohort includes carriers of the PNPLA3 risk allele as well as persons with impaired glucose tolerance (IGT) and type 2 diabetes (T2D) -- we term this group NASH-resistant (NR) (**Supplementary Table 1; Figure 1A-B**). We then identified eight individuals with NASH who are well-matched for BMI, sex, PNPLA3 genotype, and glycemic status – we refer to this group as NASH-prone (NP) (**Supplementary Table 1; Figure 1A-B**).

The livers from NP subjects displayed the classic NASH phenotype, with higher steatosis levels and the presence of hepatocyte ballooning, lobular inflammation and tissue fibrosis (**Figure 1A**). In contrast, the livers from NR subjects displayed lower levels of steatosis, lacked hepatocyte ballooning, and exhibited little-to-no lobular inflammation and fibrosis (**Figure 1A**). Individuals across both NP and NR groups exhibited very similar values of BMI, circulating lipids, glycemia and frequency of the risk allele of rs738409 at the *PNPLA3* locus (**Figure 1A-B**). Moreover, despite having healthier livers, individuals with the NR phenotype were significantly older than those with NP phenotype. Therefore, the resistance to NASH among the NR subjects is not simply attributable to a difference in time for disease progression. We hypothesized that the molecular phenotype of NR livers holds important clues to limiting NASH progression. Numerous studies have demonstrated that disease-specific enhancers/super enhancers mark nearby genes that are critical for defining disease behavior[21,22]. However, the enhancer landscape of NASH progression in humans has not been characterized. To bridge this important knowledge gap, we sought to define the enhancer signatures of both the NR and NP phenotypes.

We first performed chromatin run-on sequencing (ChRO-seq) on the liver biopsies from individuals with severe obesity who were stratified into either the NP or NR group (**Figure 1C**), thereby generating information on the transcription level of genes, active promoters, and active enhancers in a single assay. Principal component analysis (PCA) of transcriptional signal at annotated gene bodies showed a modest separation between NR and NP livers (**Figure 1D**). Differential transcription analysis identified 879 and 651 genes that are more transcriptionally active in NP and NR, respectively (**Figure 1E**; $p < 0.05$, padj < 0.2, normalized counts > 100, log2FC < or > 0 by DESeq2). The genes that are more transcriptionally active in NP than NR (hereafter, referred to as NP-activated genes) are enriched in pathways involved in neutrophil activation, immunity, and cytokine responses (**Figure 1F**). Those genes include *IL1B*[33], *IL6*[34], *CCL2* (encoding MCP-1)[35,36], *CCL20*[16], *HIF1A*[37], *TREM1*[38], *PKM*[39], *CXCR2* [40], *CRP*[41,42], *IL32*[43], and *CD44*[44] (**Figure 1H**). The genes that are more transcriptionally active in NR than NP (hereafter, referred to as NR-activated genes) are enriched in pathways involved in fatty acid metabolism (*CPT2, ACADL, AIG1, ACAA2, ACSM5, ACOX2, BDH2, HAO2*) and xenobiotic processes (*CYP2J2*)[45](**Figure 1G-H**). Overall, the transcriptionally activated genes and pathways in NP or NR are consistent with the literature[7], mostly based on studies of mouse models of NASH, describing candidate molecular players in NAFLD and/or NASH development.

NR and NP livers have unique active enhancer signatures

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ChRO-seq enables sensitive detection of active transcriptional regulatory elements (TREs), including promoters and enhancers. In this study, we defined a total of 136,863 active TREs across all samples, with an average size 397.1 bp (**Supplementary Figure 1A-B**). The identified active TREs were then classified as either promoter (n = 49,853; TREs overlapping -1000/+200 windows around transcriptional start sites) or enhancer (n = 87,010; TREs not overlapping -1000/+200 windows around transcriptional start sites). Hierarchical clustering analysis of promoter and enhancer profiles showed that sample stratification was not driven by *PNPLA3* genotype (**Figure 2A-B**). More importantly, it demonstrated that enhancer profiles separate the NR and NP phenotypes more faithfully than the promoter profiles (**Figure 2A-B**).

Next, we performed DESeq2 analysis to define TREs that are strongly associated with NP or NR livers (**Figure 2C**; $p < 0.05$, padj < 0.2, log2FC < -1 or > 1 by DESeq2). We identified 1,447 and 1,833 TREs that are significantly activated in NP and NR, respectively (**Figure 2C**). Given that enhancers more strongly discriminate between NP and NR (**Figure 2A**), we focused the analyses on the subset of differentially active TREs that are enhancers, defined as NP-specific (n=753) and NR-specific enhancers (n=1,407) (**Figure 2C**). We next performed enhancer density analysis for each transcribed gene (average normalized counts > 100 across all the samples) by counting the number of NP- or NR-specific enhancers within a window of +/-100 kb around the annotated TSS. Several key observations were made in this analysis. First, we identified very few cases of genes nearby to both NP- and NR-specific enhancers, indicating as expected that the sets of genes associated with NP- or NR-specific enhancers are largely mutually exclusive (**Supplementary Figure 1C**). Second, we found that genes that are associated with NP- or NR-specific enhancers are heavily skewed to be more transcriptionally active in NP or NR, respectively (**Figure 2D-G**). Third, genes with a greater density of NP- or NR-specific enhancers nearby exhibited greater transcriptional activation in NP and NR, respectively (**Figure 2D-E**). Together, these data indicate that the enhancer profiles are, indeed, shaping the transcriptional landscapes of each phenotype (NP/NR), which establishes the reliability of the enhancer signatures for further detailed analyses.

We identified 250 genes that are associated with 1 or more NP-specific enhancers and also exhibit significantly higher transcriptional activity in NP (**Figure 2F**), including *CD44* (**Supplementary Figure 1D**) and *THBS1* (**Figure 2H**). Hepatic CD44 was observed previously to be significantly upregulated in patients with NASH and demonstrated as a key factor in developing NASH in mice[44]. Hepatic *THBS1* was reported to be elevated in patients with obesity and advanced NASH and was downregulated after bariatric surgery[46]. Through ChRO-seq, we identified nearby enhancers that are NP-specific and thus likely contribute to transcriptional activation of these genes in NP livers. On the other hand, we identified 244 genes that are associated with 1 or more NR-specific enhancers and also exhibit significantly higher transcriptional activity in NR (**Figure G**), including *HAO2* (**Supplementary Figure 1E**), *BHMT1* and *BHMT2* (**Figure 2I**). A role for HAO2 in promoting lipid catabolism was recently demonstrated as a mechanism to suppress tumor progression in patients with renal carcinoma[47]. Our data links *HAO2* to NAFLD for the first time and uncovers its potential role in contributing to NASH resistance. BHMT1 participates in one-carbon metabolism and contributes to synthesis of glycine/methionine, the dysregulation of which has been implicated in NAFLD/NASH development[48,49]. Here, our data not only provide evidence in humans for the potential role of these genes in NASH resistance but also reveal NR-specific enhancers that likely modulate their transcriptional activity (**Figure 2I**).

Super enhancers identify candidate driver genes and pathways of NR or NP phenotypes

Super enhancers, or enhancer hotspots, are hyper-active regulatory regions defined by dense clusters of individual enhancers. Accumulating evidence shows that enhancer hotspots serve to control the transcription of genes especially critical for cell fate identity during development and cell behavior in disease progression[21]. To identify enhancer hotspots in NP or NR, we used an algorithm

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Methods) that we have recently developed[27,28], which itself is adapted from a previous pipeline for super enhancer detection[22].

The enhancer hotspot analysis revealed 56 enhancer hotspots formed by NP-specific enhancers (**Figure 3A; Supplementary Figure 2A**) and 95 enhancer hotspots formed by NR-specific enhancers (**Figure 3F; Supplementary Figure 2B**). We assigned these phenotype-specific enhancer hotspots to the closest annotated TSSs (only those genes with an average gene transcription activity > 100 normalized counts were included) (**Supplementary Figure 2A-B**). We observed that the genes linked to NP-enhancer hotspots are on average more transcriptionally active in NP livers compared to those linked to non-enhancer hotspots (**Figure 3B**). Indeed, only 45% of genes associated with non-enhancer hotspots are significantly activated in NP compared to NR, whereas 70% of genes associated with NP-enhancer hotspots are significantly activated in NP (**Figure 3B**). Very similar trends were observed for genes associated with NR-enhancer hotspots (**Figure 3G**). For both NP and NR, the enhancer hotspots are generally in close proximity to their nearest gene, except in a few cases where the distance is >25kb (**Figure 3C & H**). Together, these observations support the strong regulatory impact of enhancer hotspots on nearby genes in both NP and NR livers.

The NP-activated genes that are associated with NP-enhancer hotspots (**Figure 3C**) may represent those with particularly critical etiological roles in NASH development. We observed that these genes are enriched in the p53 pathway (Panther 2016; P = 0.0003), which is consistent with a reported role for p53-mediated apoptosis in NASH[43,50,51]. Notably, while the transcriptional activity of p53 itself is similar between NR and NP livers (data not shown), several direct targets of p53, including *CDKN1A* [52], *GADD45B*, and *GRAMD4*, as well as apoptosis-related factors such as *THBS1* and *SPI1*, are significantly increased (**Figure 3C**). Specifically, the *GADD45B* locus is associated with two NP-enhancer hotspots, which we found occurs only rarely (**Figure 3D**). Overall, our analysis suggests that the p53 network is hyper-active in NP livers. We also showed that NP-enhancer hotspots are associated with members of AP-1 (*FOS*, *FOSB*, *JUNB*)[53] and the Notch signaling activator *HES1* (**Figure 3E**)[54,55], both of which have been reported as key molecular players in hepatic fibrosis in both humans and *in vivo* mouse models. Most importantly, this enhancer hotspot analysis also revealed novel mechanistic insights into NASH development. For example, we identified *MAFK*, which can partner with members of AP-1[56] as well as *BACH* members[57], as a candidate driver of NASH. Also, although activation of Wnt signaling is known to drive liver fibrosis [41], the downstream key mediators are not completely understood. In this analysis, we uncovered Wnt signaling mediators, *EXT1*[58] and *CSRNP1*[59,60], neither of which has been previously linked to NAFLD or NASH and thus warrant detailed investigation. Other NP-enhancer hotspot associated genes such as *NFIL3*/*E4BP4* (ER stress[61]), *LITAF* (pro-inflammation and pro-fibrosis[62]), *KDM6B*/*JMJD3* (immune activation[63,64]) may also warrant functional characterization in the context of NASH development.

The NR-activated genes that are associated with NR-enhancer hotspots (**Figure 3H**) may represent those with hepatoprotective roles, thereby contributing to NASH resistance. Interestingly, many of these genes are involved in amino acid production and metabolism, particularly glycine synthesis and utilization. Specifically, we uncover that *GLYAT*, which carries out glycine conjugation during detoxification (**Figure 3I**)[65], and *GATM*, which mediates creatine biosynthesis from glycine (**Figure 3J**), could be critical in contributing to NASH resistance. Other genes identified through this analysis include those involved in bile acid synthesis and signaling (*AKR1D1*, *NR1I2/PXR*)[66,67], anti-oxidative and anti-ER stress (*CAT*, *PON1*, *SESN2*), selenium biology (*SELENOP*) and hormone signaling (*GHR* and *AR*). Among these genes, the reduction of *PON1*[68], *SELENOP*[69] and *ARK1D1*[70] have been shown to correlate to severity of NAFLD in human patients. Mouse models with deficiency in *Cat*[71], *Sesn2*[72], *Ghr*[73] and androgen receptor (*Ar*)[74,75] spontaneously develop NAFLD development or facilitate NAFLD progression. Consistent with those observations, we have identified them for the first time to be linked to NR-specific enhancer hotspots and serve as candidate contributors to NASH resistance in human patients.

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Define active TF gene networks in NP livers

TFs initiate transcription of downstream genes by binding to sequence-specific motifs present within nearby active enhancers. To identify candidate master TF regulators relevant to NP livers, we performed motif enrichment analysis within NP-specific enhancers (**Figure 4A**; $p < 0.05$, q-value < 0.05, enrichment fold change > 1.5 by HOMER). The TFs that are highly expressed in NP livers (normalized counts > 500 in NP) and exhibit significantly enriched binding motifs in NP-specific enhancers are shown in Figure 4B (top 10 motifs based on P value) and Supplementary Figure 3. The NP associated TFs include ATF members (e.g. ATF3, ATF4)[24,76,77], AP-1 members (e.g. JUN, JUNB, JUND)[53,78], STAT members (e.g. STAT3, STAT1)[79,80], and MAFK (**Figure 4B**), several of which have been implicated in mouse models of NAFLD formation and/or disease progression to NASH. Notably, we observed that many of the NP-associated TFs are themselves significantly activated at the transcriptional level in NP compared to NR (**Figure 4C**), and even associated with NP-specific enhancer hotspots (**Figure 3C**), which suggests that activation of TF networks in NP livers is achieved by transcriptional activation of NP associated TFs as well as the enhancers containing the corresponding binding motifs.

We next sought to define major TF-gene networks (cistromes) in NP livers. Specifically, we selected ATF3, JUNB, STAT3 and MAFK as TFs of interest (due in part to their dissimilarity of binding motifs) and then identified NP-activated genes that are associated with (within +/-100 kb window from TSS) NP-specific enhancers containing one or more of their binding motifs (**Figure 4D**). Through this analysis, we identified genes that are shared across different TF cistromes or uniquely associated with a particular TF (**Figure 4D**). For example, *CDKN1A*, an NP-specific enhancer hotspot linked gene, is associated with all four selected TFs (**Figure 4D-E**). On the other hand, the NOTCH signaling mediator *HES1*, another NP-specific enhancer hotspot linked gene, is uniquely associated with *STAT3* (**Figure 4D**), indicating potential crosstalk between STAT and NOTCH signaling in NP livers. This analysis also elucidated ATF3 and JUNB as novel candidate TFs that act on *CSRNP1* (**Figure 4D, F**) and MAFK as a novel candidate TF acting on *FOS* (**Figure 4D, G**). In general, we observed that the genes likely regulated by each of the selected TFs are enriched in different pathways, thereby potentially contributing to distinct aspects of the NP phenotype (Figure 4H). For example, the top ranked pathway for ATF3-associated NP-activated genes is reactive oxygen species generation, whereas the top ranked pathways for MAFK-associated NP-activated genes is cytokine-mediated signaling (**Figure 4H**). These findings provide a glimpse into how different gene regulatory mechanisms relate to each other to form a complex regulatory program that may promote NASH development and progression (an example of an NP-specific regulatory module is shown in **Figure 4I**).

Define active TF gene networks in NR livers

To identify candidate master TF regulators relevant to NR livers, we performed motif enrichment analysis within NR-specific enhancers (**Figure 5A**; $p < 0.05$, q-value < 0.05, enrichment fold change > 1.5 by HOMER). The TFs that are highly expressed in NR livers (normalized counts > 500 in NR) and exhibit significantly enriched binding motifs in NR-specific enhancers are defined as NR associated TFs (**Figure 5B**). Among these NR associated TFs, HNF4A[43], HNF1B[81], FOXA2[82] and AR[74,75] have reported roles in maintaining liver metabolism and/or preventing NAFLD/NASH development. Interestingly, the genes encoding all of the NR associated TFs, with the exception of AR, are transcribed at similar levels between NR and NP (**Figure 5C**). This suggests that activation of NR-associated TF networks relies primarily on enhancing the activity of the TFs at enhancers without affecting their own transcription levels (unlike what was observed in NP).

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To define major TF-gene networks (cistrome) in NR livers, we selected HNF4A, HNF1B, FOXA2 and AR as TFs of interest and then identified NR-activated genes that are associated with NR-specific enhancers containing one or more of their binding motifs (**Figure 5D**). Notably, *GATM* and *GLYAT*, which are linked to NR-specific enhancer hotspots (**Figure 3G-H**), are associated with all four selected TFs (**Figure 5D**), highlighting their potential importance to NASH resistance. By performing pathway enrichment analysis for each of the NR-associated TF networks, we observed enrichment in monocarboxylic acid and fatty acid oxidation (**Figure 5E**). We also performed protein-protein interaction (PPI) enrichment analysis with all genes associated with at least one of the four selected motifs, which revealed key modules regulating peroxisomal protein import, xenobiotic metabolism and glycine, serine and threonine metabolism (**Figure 5F**). Two genes in the latter module, *SARDH* (**Figure 5G**) and *SHMT1* (**Figure 5H**), are uniquely associated with AR and HNF4A (**Figure 5D**), respectively, whereas the third gene, *GATM* (**Figure 3J**), is associated with all four selected NR associated TFs (**Figure 5D**).

NR-specific enhancers and the associated TF network activates the hepatic serine and glycine synthesis pathways

The enrichment of glycine, serine and threonine (GST) metabolism in NR-associated TF networks (**Figure 5F**) supports very recent studies linking glycine[49] and serine[83] deficiency to NAFLD/NASH development in human patients. Given that circulating amino acid profiles can be influenced by the status of hepatic amino acid metabolism[65,84,85], we measured amino acids and related metabolites in the blood of NR and NP subjects (**Figure 6A**). We found that NR subjects exhibit a significant increase in serine (1.38 fold increase; $p = 0.028$) and a modest trend for higher glycine levels (1.27 fold increase; $p = 0.14$) in circulation compared to NP subjects (**Figure 6A**), which is consistent with our findings generated from the analysis of NR-associated gene networks (**Figure 5F**).

The increase in circulating serine in NR subjects (**Figure 6A**) motivated us to examine further the GST metabolism pathway (KEGG pathway hsa00260). We found that NR-activated genes are very strongly associated with serine and glycine biosynthesis (**Figure 6B**). Genes involved in serine biosynthesis, particularly *PHGDH* ($p < 0.05$, padj < 0.2 by DESeq2), and to a lesser extent *PSPH* ($p < 0.05$, padj $= 0.23$ by DESeq2), are significantly activated in NR compared to NP livers (**Figure 6B**). Serine is known to support glycine synthesis through enzymatic activity of SHMT1 (**Figure 6B**), which is also activated in NR livers. Interestingly, a number of genes associated with glycine production (*BHMT*, *DMGHD* and *SARDH*) are also highly significantly activated in NR compared to NP livers (**Figure 6B**). BHMT is a core component of the one-carbon cycle that provides methyl groups required for phosphatidylcholine synthesis and VLDL secretion[86]. The remarkable transcriptional activation of NR-specific enhancer hotspot associated genes, *GLYAT* and *GATM*, also points to glycine conjugation-mediated acyl-CoA clearance and creatine synthesis as potentially critical processes for resisting the development of NASH (**Figure 6B**; $p < 0.05$, padj < 0.2 by DESeq2). *GAMT*, another gene involved in creatine synthesis ($p < 0.05$, padj $= 0.21$ by DESeq2), is also highly activated in NR compared to NP livers (**Figure 6B**). Almost all of the NP-activated genes in the GST metabolism pathway exhibit levels of transcriptional activity that are significantly correlated with circulating serine levels (**Figure 6B**). Overall, these findings suggest that serine and glycine synthesis, as well as glycine utilization and metabolism, likely serve as important mechanisms underlying NASH resistance. Our analysis pinpoints particularly important genes in the glycine/serine metabolism pathway and also identifies NR-specific enhancers/enhancer hotspots and candidate TFs relevant for their activation (**Figure 6B**).

Discussion

In this study, ChRO-seq was leveraged to define for the first time the active enhancer signature of NASH-prone (NP) and NASH-resistant (NR) phenotypes in a patient cohort with severe obesity.

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Although whole-transcriptome profiling has been used previously to characterize the gene expression landscape of NAFLD/NASH in humans[13-15], the chromatin activity patterns, enhancer signature, and TF networks that drive gene expression in NASH remain largely undefined. The results of this study fill this important knowledge gap.

The enhancer hotspot analysis of NP livers highlighted the p53 pathway and the associated processes (e.g., apoptosis and cellular senescence) as key candidate drivers of NASH development[15,50,87]. We found that *CDKN1A* (p21), which is a direct downstream target of p53 and has been implicated in NAFLD progression[88], is linked to an enhancer hotspot that is uniquely active in NP livers and also harbors binding motifs for several NP-associated TFs (e.g., AP-1, ATF3, STAT3, and MAFK). We observed that apoptosis mediators *THBS1* and *GADD45B* are also associated with NP-enhancer hotspots, the former of which has been linked to NAFLD progression in both humans and mice[46,89] but the latter of which has not previously been reported in the context of NAFLD or NASH and therefore warrants more detailed characterization.

Several of the NP-activated genes identified in our study have been observed previously to be elevated in humans with NAFLD/NASH. These include *HES1*[54,55], AP-1[53], *CRP*[41,42], *TNFRSF1B* (also known as p75)[90,91] and *MYO1C* [92]. Also, several genes implicated in NAFLD/NASH pathogenesis in mouse models, such as *Nfil3* (encoding E4bp2)[61], *Pr2x4*[93] and *Dusp1* (encoding Mkp-1)[94], were identified in our study for the first time as candidate drivers of NASH in humans. Most importantly, our study uncovered a core set of genes (e.g. *CSRNP1*, *HBEGF*, *SERPINB1*) that have not been previously linked directly to NAFLD/NASH. *CSRNP1* is known to augment Wnt signaling and promote apoptosis[59,60]. HBEGF can serve as a ligand for EGFR, which is known to mediate the activation of hepatic stellate cells in a diet-induced mouse model of liver injury[95]. Hbegf overexpression induces fibrosis in mouse pancreas[96]; whether HBEGF contributes to liver fibrosis through activation of EGFR signaling remained to be tested. SERPINB1, which is a liver derived factor and secreted into the circulation, regulates islet beta cell proliferation[97]. The role of SRPINB1 in the context NAFLD/NASH context remains unclear and merits future studies.

Based on the current literature, it is possible that some of the NP-associated TFs uncovered in our study (e.g. AP-1, HES1 and ATF3) exert cell-type specific roles in the liver during NASH progression. For example, in mouse NAFLD/NASH models, Jun/AP-1 was observed to promote cell survival in hepatocytes but drive pro-inflammatory/pro-fibrotic programs in non-parenchymal cells[23,53]. Another notable example is HES1, which has been shown to exhibit cell type specific expression patterns in human livers (hepatocytes vs. non-parenchymal cells) during the course of disease progression from healthy to steatosis to NASH[54]. More recently, HES1 in a human study[55] and Atf3 in a mouse study[24] were suggested to play a critical role in scar-associated macrophages. Active enhancer profiles, which are known to be cell-type specific, may be a key underlying mechanism for directing cell-type specific roles of TFs. In the future, it will be critical to extend the work of this study to achieve single-cell resolution on TF networks.

Our study also identified genes that may be pivotal for NASH resistance in humans. We identified several NR associated TFs, including HNF4A, HNF1B, FOXA2 and AR. While these TFs are already known to be important for several liver functions, we pinpoint in this study specific downstream genes that are likely especially critical for driving the NR phenotype in humans. Among these genes are *PON1*[68], *ETNK2*[15] and *IGFBP2*[98], which have been reported as promising markers that distinguish NASH from simple steatosis or healthy livers. Here we have identified the enhancers and transcription factors associated with the regulation of these genes.

While numerous studies in animal models have reported a wide array of genes that may confer beneficial effects in the liver toward mitigating NASH, which of them if any are relevant to humans remains unclear. An important contribution of our study is to help determine which of these many

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

candidates are most pertinent in humans. Several genes that have been implicated in mouse models and are highlighted here for the first time as part of the NR-associated TF network in humans include: *Lias* (alleviating oxidative stress)[99], *Sesn2* (ameliorating ER-stress)[72], *Cygb* (reducing stellate cell activation)[100], and *Smo* (preventing injury-induced apoptosis)[101].

The most intriguing and novel findings are from the analysis defining NR-specific enhancer hotspot associated genes, which are enriched in glycine-related functions: glycine synthesis (*BHMT*), glycine conjugation-mediated acyl-CoA clearance (*GLYAT*), and glycine-based creatine synthesis (*GATM*). Our results also highlight the importance of serine and glycine synthesis, as almost all the genes involved in these pathways (e.g., *PHGDH*, *SHMT*, *BHMT*, *SARDH*) are significantly activated in NR livers, associated with NR-specific enhancer hotspots, and/or positively correlated with circulating serine levels. Our findings are consistent with two human studies, one of which indicated serine deficiency in human NAFLD/NASH patients[83] and the other showing that glycine deficiency is a risk factor for liver fibrosis[102]. Our observations also support recent studies that suggest serine-, glycine- or creatine-based treatments for NASH prevention[49,103,104]. The novelty of the present study is that we pinpointed particularly important genes in the glycine/serine metabolism pathway, in part by association with NR-specific super-enhancers, and the TF networks most relevant for their activation.

Overall, by leveraging the cutting-edge genome-scale technique ChRO-seq on human liver biopsies from a cohort of obese individuals, we defined for the first time active enhancer signatures that distinguish the NR from the NP phenotype. We were able to identify genes that are linked to NP- or NR-specific enhancers/enhancer hotspots and offer a basis for understanding how enhancer status is modulated to regulate transcriptional activity of genes during NAFLD progression. We also describe the re-wiring of TF networks in NASH relative to NAFLD. Some of the findings in this study help uncover the specific genes and networks reported in animal models that may be most relevant in humans. Other findings represent completely novel candidate molecular factors that have not previously been linked to NAFLD/NASH and may serve as potential therapeutic targets for managing NASH.

Methods

Human sample description

Tissue specimens, blood, and plasma samples were obtained from the Biobank of the Institut Universitaire de Cardiologie et de Pneumologie de Québec (Québec Heart and Lung Institute, QHLI, Québec City, QC, Canada) according to institutionally approved management modalities. All participants provided written, informed consent. Participants were recruited during a preoperative visit for bariatric surgery and met the inclusion/exclusion criteria for the procedure[105]. The study population included 18 female patients of European ancestry with severe obesity (BMI > 45 kg/m^2) aged 29-56 years from the eastern provinces of Canada who underwent bariatric surgery at the Institut Universitaire de Cardiologie et de Pneumologie de Québec (Québec Heart and Lung Institute, QHLI, Québec City, QC, Canada) and were not on diabetes medications. Blood samples were obtained at the time of admission for surgery. Liver tissues were obtained during bariatric surgery by incisional biopsy of the left lobe and were not cauterized. The sampling procedure and position was standardized among surgeons. A portion of the liver tissues were rapidly frozen in liquid nitrogen and another portion was fixed for histology grading by trained pathologists according to[106]. The participants used for the study were selected based on the presence or absence of NASH defined as steatosis alongside both lobular inflammation and ballooning. 6 out of the 18 samples showed no indication of hepatocyte ballooning and hepatic fibrosis were categorized as NASH resistant (NR). 8 out of the 18 samples with indications of hepatocyte ballooning, hepatic fibrosis, lobular inflammation, and portal inflammation were categorized as NASH prone (NP). The remaining 4 samples in this

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cohort satisfied neither NR nor NP criteria and thus were excluded from NR vs. NP comparison throughout the entire study. The NR and NP groups were matched for BMI, sex, glucose tolerance, PNPLA3 genotype and smoking status. Additional sample and subject details are provided in Supplementary Table 1.

Genotyping

Patatin-like phospholipase domain-containing protein 3 (PNPLA3) genotyping for the Ile148Met variant associated with hepatic steatosis (rs738409) was performed on genomic DNA extracted from the blood buffy coat using the GenElute Blood Genomic DNA kit (Sigma, St. Louis, MO, USA). rs738409 was genotyped using validated primers and TaqMan probes (Applied Biosystems, Waltham, MA). PNPLA3 genotypes were determined using 7500 Fast Real-Time PCR System (Applied Biosystems, Waltham, MA) and analyzed using a high-throughput array technology (QuantStudio 12K Flex system), coupled with Taqman OpenArray technology (Life Technologies).

Plasma metabolite measurements

Plasma amino acids were measured by targeted metabolomics methods, as previously described[107,108]. Briefly, plasma amino acid profiling was performed by tandem mass spectrometry (MS/MS). All MS analyses employed stable-isotope-dilution with internal standards from Isotec, Cambridge Isotopes Laboratories, and CDN Isotopes.

Chromatin isolation

Chromatin isolation for length extension chromatin run-on sequencing (ChRO-seq) was performed as previously described[26,27]. To isolate chromatin from pulverized frozen liver biopsies, samples were incubated in 1X NUN buffer [20 mM HEPES, 7.5 mM MgCl2, 0.2 mM EDTA, 0.3 M NaCl, 1M urea, 1% NP-40, 1 mM DTT, 50 units/mL RNase Cocktail Enzyme Mix (Thermo Fisher Scientific, Waltham, MA; AM2286), 1X Protease Inhibitor Cocktail (Roche, Basel, Switzerland; 11873580001)] in an Eppendorf Thermomixer (Eppendorf, Hamburg, Germany) set at 12°C and 2000 rpm for 30 minutes. The chromatin was pelleted by centrifugation at 12,500 × g for 30 minutes at 4°C, washed by 50 mM Tris-HCl (pH 7.5) containing 40 units/mL SUPERase inhibitor and stored in chromatin storage buffer (50 mM Tris-HCl pH 8.0, 25% glycerol, 5 mM magnesium acetate, 0.1 mM EDTA, 5 mM DTT, and 40 units/mL SUPERase In RNase Inhibitor). To solubilize the chromatin into the storage buffer, samples were loaded into a Bioruptor (Diagenode, Denville, NJ) and sonicated with repeated cycles (10 minutes per cycle, consisted of 10 rounds of 30s on and 30s off). Chromatin solubilized in storage buffer was stored at −80°C until usage for library construction.

ChRO-seq library preparation

Library preparation for length extension ChRO-seq was performed as previously described[26,27]. To perform run-on reaction with the solubilized chromatin, samples were mixed with an equal volume of 2X run-on reaction mix [10 mM Tris-HCl pH 8.0, 5 mM MgCl2, 1 mM DTT, 300 mM KCl, 400 μM ATP, 0.8 μM CTP, 400 μM GTP, 400 μM UTP, 40 μM Biotin-11-CTP (Perkin Elmer, Waltham, MA; NEL542001EA), 100 ng yeast tRNA (VWR, 80054–306), 0.8 units/μL SUPERase In RNase Inhibitor, 1% (w/v) Sarkosyl]. The run-on reaction was incubated in an Eppendorf Thermomixer at 37°C for 5 min (700 rpm) and stopped by adding Trizol LS (Life Technologies, Carlsbad, CA; 10296–010). RNA samples were precipitated by GlycoBlue (Ambion, Austin, TX; AM9515) and resuspended in diethylpyrocarbonate (DEPC)-treated water. To perform base hydrolysis reaction, RNA samples in DEPC water were heat denatured at 65°C for 40 s and incubated in 0.2N NaOH on ice for 4 min. Base hydrolysis reaction was stopped by neutralizing with Tris-HCl pH 6.8. Nascent RNA was enriched using streptavidin beads (NEB, Ipswich, MA; S1421) followed by RNA extraction using

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Trizol. To perform adapter ligations, nascent RNA samples were processed through the following steps: (i) 3′ adapter ligation with T4 RNA Ligase 1 (NEB, Ipswich, MA; M0204), (ii) RNA binding with streptavidin beads (NEB, Ipswich, MA; S1421) followed by RNA extraction with Trizol, (iii) 5′ de-capping with RNA 5′ pyrophosphohydrolase (NEB, Ipswich, MA; M0356), (iv) 5′ end phosphorylation using T4 polynucleotide kinase (NEB, Ipswich, MA; M0201), (iv) 5′ adapter ligation with T4 RNA Ligase 1 (NEB, Ipswich, MA; M0204). The 5' adaptor contained a 6-nucleotide unique molecular identifier (UMI) to allow for bioinformatic detection and elimination of PCR duplicates. Streptavidin bead binding followed by Trizol RNA extraction was performed again before final library construction. To generate ChRO-seq library, cDNA was generated through a reverse-transcription reaction using Superscript III Reverse Transcriptase (Life Technologies, Carlsbad, CA; 18080–044) and amplified using Q5 High-Fidelity DNA Polymerase (NEB, Ipswich, MA; M0491). Finally, ChRO-seq libraries were sequenced (5′ single end; single-end 75×) using the NextSeq500 high-throughput sequencing system (Illumina, San Diego, CA) at the Cornell University Biotechnology Resource Center.

ChRO-seq mapping and visualization

ChRO-seq mapping was performed with an established pipeline[109]. Briefly, read quality was assessed using FastQC. PCR deduplication was performed by collapsing UMIs using PRINSEQ lite 0.20.2[110]. Adapters were trimmed from the 3' end of remaining reads using cutadapt 1.16 with a maximum 10% error rate, minimum 2 bp overlap, and minimum 20 quality score. Processed reads with a minimum length of 15 bp were mapped to the hg38 genome modified with the addition of a single copy of the human Pol I ribosomal RNA complete repeating unit (GenBank: U13369.1) using Burrows-Wheeler Aligner (BWA) 0.7.13[111]. The location of the RNA polymerase active site was represented by a single base that denotes the 5' end of the nascent RNA, which corresponds to the position on the 3' end of each sequenced read. Supplementary Table 2 provides the mapping statistics of the ChRO-seq experiments. As denoted in the table, 4 samples in this patient cohort did not satisfy the criteria for NR or NP category based on their clinical profiles. The ChRO-seq data of these 4 samples were used only in the dREG analysis for calling active promoters and enhancers (see method section: Transcription activity transcriptional regulatory elements), but not elsewhere. To visualize ChRO-seq signal, data were converted to bigwig format using bedtools and UCSC bedGraphToBigWig. Bigwig files within a sample category were then merged and normalized to a total signal of 1×10^6 reads. Genomic loci snapshots were generated using R package Gviz[112].

Transcription activity of genes

Quantification of gene transcription activity was based on hg38 GENCODE v33 annotations. Read counts of gene loci were quantified using the R package bigwig (https://github.com/andrelmartins/bigWig). To quantify transcription activity of gene loci, stranded ChRO-seq reads within gene coordinates were counted, with exclusion of reads within 500 b downstream of transcription start site (TSS) to avoid bias generated by the RNA polymerase pausing at the promoters. Genes with gen body smaller than 1 kb were excluded from all the gene body related analysis. Principal components analysis (PCA) of genes was performed using counts with rlog transformation. In differential transcription analysis, ChRO-seq raw counts of gene loci were analyzed through DESeq2 1.30.1[113], which models based on negative binomial distribution, to define genes that are significantly activated in NASH-resistant or NASH-prone samples. Genes significantly activated in one specific group were defined by criteria of log2 foldchange > 0 (or < 0), normalized counts > 100, $P < 0.05$, padj < 0.2 (Wald test; DESeq2).

Transcription activity transcriptional regulatory elements (TREs)

To identify active TREs across all the samples, bigwig files of the same strand from all samples were merged. This merged dataset was analyzed through dREG[109,114], the peak-calling algorithm that detects short bidirectional expression patterns across the entire genome, to define active TREs. To

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

quantify activity of TREs, total reads from both strands of each TRE were counted using the R package bigwig (https://github.com/andrelmartins/bigWig). Classification of TREs into promoters and enhancers was based on hg38 GENCODE v33 annotations. TREs that had at least 1 b overlapping with the window of 1000 b upstream and 200 b downstream of TSS were defined as promoter regions; the remaining TREs were defined as enhancers. Raw counts after rlog transformation were used for PCA and hierarchical clustering analysis that computes pairwise correlations between samples based on promoter or enhancer profiles. To identify TREs that are uniquely active in NASH-resistant or NASH-prone samples, ChRO-seq raw counts of TREs were analyzed through DESeq2 1.30.1 [113]. TREs associated with NASH-resistant phenotype were defined by criteria of log2 foldchange > 1 (or < -1), P < 0.05, padj < 0.2 (Wald test; DESeq2). TREs associated with NASH-prone phenotype were defined by criteria of log2 foldchange < -1), P < 0.05, padj < 0.2 (Wald test; DESeq2). In enhancer density analysis, the enhancers present within +/-100 kb window from annotated TSSs of genes (longest isoform if there are multiple; hg38 GENCODE v33) were counted.

Transcription factor motif enrichment analysis

HOMER [115] was used to determine enrichment of sites corresponding to known motifs of a given set of enhancer peaks. Specifically, findMotifsGenome.pl function, with “given” as the size parameter, was used to query NASH-resistant enhancer peaks (using all enhancers that were not identified as NASH-resistant as background) and NASH-prone enhancer peaks (using all enhancers that were not identified as NASH-resistant as background). “Known motif” results generated from HOMER were reported in this study. To identify enhancers that contain binding motif(s) for a specific transcription factor, HOMER function annotatePeaks.pl (“given” as the size parameter; genome build hg38) was used.

Defining enhancer hotspots

NASH-resistant and NASH-prone enhancer hotspots were defined using an analysis pipeline described previously[28]. The enhancer hotspots in this study were identified by the criteria similar to the studies describing ‘super-enhancers’. Briefly, the enhancers uniquely activated in one specific group (NASH-resistant or -prone samples) were stitched based on proximity of distance (<12.5 kb). To quantify the activity of stitched enhancers, counts of each enhancer were normalized by DESeq2 and summed for each stitched enhancer. To determine which stitched enhancers were qualified as ‘enhancer hotspots’, the stitched enhancers were first ranked by their activity, creating a curve of activity to rank plot. A tangent line was then applied to the curve. The ones above the cutoff point that is determined by the tangent line were defined as enhancer hotpots, whereas the ones below the cutoff point were non-enhancer hotspots.

Gene set analysis

Pathway analyses were performed using Enrichr[116-118] and KEGG pathway[119]. Protein-protein interaction enrichment analysis, which is based on Molecular Complex Detection (MCODE) algorithm, was performed using Metascape[120].

Statistics

Statistical analyses were performed using R (4.0.4). Statistical significance of blood measurements between groups was determined using unpaired two sample T-test. In sequencing studies, statistical significance was determined using DESeq2, where p values were calculated by Wald test and the p values were adjusted using Benjamin and Hochberg (BH) method. Other statistical methods including unpaired two-sample Wilcoxon test and two-sided Pearson correlation test were used as indicated in the figure legends. The p values in pathway enrichment analysis by Enrichr were

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

calculated using Fisher's exact test. Values of $p < 0.05$ (and adjusted $p < 0.2$ if in sequencing studies) were considered statistically significant unless otherwise noted. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Data availability

Raw and processed data generated in the sequencing studies will be made available upon publication.

Acknowledgement

The Authors acknowledge the invaluable collaboration of the surgery team, bariatric surgeons and Biobank staff of the IUCPQ. The authors also thank the Cornell University Biotechnology Resource Center for generating ChRO-seq genomic dataset.

Competing interests

A.T. and L.B. receive funding from Johnson & Johnson Medical Companies, Medtronic, Bodynov and GI Windows for studies on bariatric surgery. A.T. and L.B. acted as consultants for Bausch Health and Novo Nordisk. No conflicts of interest were declared by the other authors.

Author Contributions

Conceptualization, P.S., P.J.W., Y.-H.H.; Tissue collection and biobank resources, L.B., A.T; Genotyping, M.C.V.; Metabolomics, O.I.; ChRO-seq library preparation, R.S; Bioinformatic analyses and data curation, Y.-H.H.; Writing (original draft), Y.-H.H., P.S.; Review and editing, Y.-H.H., P.S., P.J.W., R.S., A.T.; Supervision, P.S., P.J.W.

Funding

This work was supported by American Diabetes Association Pathways to Stop Diabetes Awards (1-16-ACE-47 to P.S. and 1-16-INI-17 to P.J.W.) and a Borden Scholars Award through Duke University (P.J.W.). This study was performed with support from the Research Chair in Bariatric and Metabolic Surgery at Laval University (L.B. and A.T.). The Biobank is supported by the IUCPQ foundation.

References

1 Loomba, R. & Sanyal, A. J. The global NAFLD epidemic. *Nat Rev Gastroenterol Hepatol* **10**, 686-690, doi:10.1038/nrgastro.2013.171 (2013).

2 Younossi, Z. M. *et al.* Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* **64**, 73-84, doi:10.1002/hep.28431 (2016).

3 Trepo, E. & Valenti, L. Update on NAFLD genetics: From new variants to the clinic. *J Hepatol* **72**, 1196-1209, doi:10.1016/j.jhep.2020.02.020 (2020).

4 Rutledge, S. M. & Asgharpour, A. Smoking and Liver Disease. *Gastroenterol Hepatol (N Y)* **16**, 617-625 (2020).

5 Younossi, Z. *et al.* Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. *Nat Rev Gastroenterol Hepatol* **15**, 11-20, doi:10.1038/nrgastro.2017.109 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

6 Kleiner, D. E. *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313-1321, doi:10.1002/hep.20701 (2005).
7 Loomba, R., Friedman, S. L. & Shulman, G. I. Mechanisms and disease consequences of nonalcoholic fatty liver disease. *Cell* **184**, 2537-2564, doi:10.1016/j.cell.2021.04.015 (2021).
8 Dulai, P. S. *et al.* Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. *Hepatology* **65**, 1557-1565, doi:10.1002/hep.29085 (2017).
9 Wong, R. J. *et al.* Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. *Gastroenterology* **148**, 547-555, doi:10.1053/j.gastro.2014.11.039 (2015).
10 Diehl, A. M. & Day, C. Cause, Pathogenesis, and Treatment of Nonalcoholic Steatohepatitis. *N Engl J Med* **377**, 2063-2072, doi:10.1056/NEJMra1503519 (2017).
11 Vuppalanchi, R., Noureddin, M., Alkhouri, N. & Sanyal, A. J. Therapeutic pipeline in nonalcoholic steatohepatitis. *Nat Rev Gastroenterol Hepatol* **18**, 373-392, doi:10.1038/s41575-020-00408-y (2021).
12 Ratziu, V. & Friedman, S. L. Why do so many NASH trials fail? *Gastroenterology*, doi:10.1053/j.gastro.2020.05.046 (2020).
13 Baselli, G. A. *et al.* Liver transcriptomics highlights interleukin-32 as novel NAFLD-related cytokine and candidate biomarker. *Gut* **69**, 1855-1866, doi:10.1136/gutjnl-2019-319226 (2020).
14 Govaere, O. *et al.* Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis. *Sci Transl Med* **12**, doi:10.1126/scitranslmed.aba4448 (2020).
15 Suppli, M. P. *et al.* Hepatic transcriptome signatures in patients with varying degrees of nonalcoholic fatty liver disease compared with healthy normal-weight individuals. *Am J Physiol Gastrointest Liver Physiol* **316**, G462-G472, doi:10.1152/ajpgi.00358.2018 (2019).
16 Chu, X. *et al.* CCL20 is up-regulated in non-alcoholic fatty liver disease fibrosis and is produced by hepatic stellate cells in response to fatty acid loading. *J Transl Med* **16**, 108, doi:10.1186/s12967-018-1490-y (2018).
17 Gerhard, G. S. *et al.* Transcriptomic Profiling of Obesity-Related Nonalcoholic Steatohepatitis Reveals a Core Set of Fibrosis-Specific Genes. *J Endocr Soc* **2**, 710-726, doi:10.1210/js.2018-00122 (2018).
18 Hotta, K. *et al.* Identification of core gene networks and hub genes associated with progression of non-alcoholic fatty liver disease by RNA sequencing. *Hepatol Res* **47**, 1445-1458, doi:10.1111/hepr.12877 (2017).
19 Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108-112, doi:10.1038/nature07829 (2009).
20 Nord, A. S. *et al.* Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* **155**, 1521-1531, doi:10.1016/j.cell.2013.11.033 (2013).
21 Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-947, doi:10.1016/j.cell.2013.09.053 (2013).
22 Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319, doi:10.1016/j.cell.2013.03.035 (2013).
23 Liu, X. *et al.* Identification of Lineage-Specific Transcription Factors That Prevent Activation of Hepatic Stellate Cells and Promote Fibrosis Resolution. *Gastroenterology* **158**, 1728-1744 e1714, doi:10.1053/j.gastro.2020.01.027 (2020).
24 Seidman, J. S. *et al.* Niche-Specific Reprogramming of Epigenetic Landscapes Drives Myeloid Cell Diversity in Nonalcoholic Steatohepatitis. *Immunity* **52**, 1057-1074 e1057, doi:10.1016/j.immuni.2020.04.001 (2020).
25 Cave, M. C. *et al.* Nuclear receptors and nonalcoholic fatty liver disease. *Biochim Biophys Acta* **1859**, 1083-1099, doi:10.1016/j.bbagrm.2016.03.002 (2016).
26 Chu, T. *et al.* Chromatin run-on and sequencing maps the transcriptional regulatory landscape of glioblastoma multiforme. *Nat Genet* **50**, 1553-1564, doi:10.1038/s41588-018-0244-3 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

27 Dinh, T. A. *et al.* Hotspots of Aberrant Enhancer Activity in Fibrolamellar Carcinoma Reveal Candidate Oncogenic Pathways and Therapeutic Vulnerabilities. *Cell Rep* **31**, 107509, doi:10.1016/j.celrep.2020.03.073 (2020).

28 Hung, Y. H. *et al.* Chromatin regulatory dynamics of early human small intestinal development using a directed differentiation model. *Nucleic Acids Res* **49**, 726-744, doi:10.1093/nar/gkaa1204 (2021).

29 Arendt, B. M. *et al.* Altered hepatic gene expression in nonalcoholic fatty liver disease is associated with lower hepatic n-3 and n-6 polyunsaturated fatty acids. *Hepatology* **61**, 1565-1578, doi:10.1002/hep.27695 (2015).

30 Moylan, C. A. *et al.* Hepatic gene expression profiles differentiate presymptomatic patients with mild versus severe nonalcoholic fatty liver disease. *Hepatology* **59**, 471-482, doi:10.1002/hep.26661 (2014).

31 Younossi, Z. M. *et al.* A genomic and proteomic study of the spectrum of nonalcoholic fatty liver disease. *Hepatology* **42**, 665-674, doi:10.1002/hep.20838 (2005).

32 Younossi, Z. M. *et al.* Hepatic gene expression in patients with obesity-related non-alcoholic steatohepatitis. *Liver Int* **25**, 760-771, doi:10.1111/j.1478-3231.2005.01117.x (2005).

33 Kamari, Y. *et al.* Lack of interleukin-1alpha or interleukin-1beta inhibits transformation of steatosis to steatohepatitis and liver fibrosis in hypercholesterolemic mice. *J Hepatol* **55**, 1086-1094, doi:10.1016/j.jhep.2011.01.048 (2011).

34 Wieckowska, A. *et al.* Increased hepatic and circulating interleukin-6 levels in human nonalcoholic steatohepatitis. *Am J Gastroenterol* **103**, 1372-1379, doi:10.1111/j.1572-0241.2007.01774.x (2008).

35 Baeck, C. *et al.* Pharmacological inhibition of the chemokine CCL2 (MCP-1) diminishes liver macrophage infiltration and steatohepatitis in chronic hepatic injury. *Gut* **61**, 416-426, doi:10.1136/gutjnl-2011-300304 (2012).

36 Mandrekar, P., Ambade, A., Lim, A., Szabo, G. & Catalano, D. An essential role for monocyte chemoattractant protein-1 in alcoholic liver injury: regulation of proinflammatory cytokines and hepatic steatosis in mice. *Hepatology* **54**, 2185-2197, doi:10.1002/hep.24599 (2011).

37 Wang, X. *et al.* Macrophage-Specific Hypoxia-Inducible Factor-1alpha Contributes to Impaired Autophagic Flux in Nonalcoholic Steatohepatitis. *Hepatology* **69**, 545-563, doi:10.1002/hep.30215 (2019).

38 Rao, S. *et al.* Inhibition of TREM-1 attenuates inflammation and lipid accumulation in diet-induced nonalcoholic fatty liver disease. *J Cell Biochem*, doi:10.1002/jcb.28468 (2019).

39 Ouyang, X. *et al.* Digoxin Suppresses Pyruvate Kinase M2-Promoted HIF-1alpha Transactivation in Steatohepatitis. *Cell Metab* **27**, 339-350 e333, doi:10.1016/j.cmet.2018.01.007 (2018).

40 Ye, D. *et al.* Lipocalin-2 mediates non-alcoholic steatohepatitis by promoting neutrophil-macrophage crosstalk via the induction of CXCR2. *J Hepatol* **65**, 988-997, doi:10.1016/j.jhep.2016.05.041 (2016).

41 Wang, J. N. *et al.* Emerging role and therapeutic implication of Wnt signaling pathways in liver fibrosis. *Gene* **674**, 57-69, doi:10.1016/j.gene.2018.06.053 (2018).

42 Yoneda, M. *et al.* High-sensitivity C-reactive protein is an independent clinical feature of nonalcoholic steatohepatitis (NASH) and also of the severity of fibrosis in NASH. *J Gastroenterol* **42**, 573-582, doi:10.1007/s00535-007-2060-x (2007).

43 Xu, Y. *et al.* Hepatocyte Nuclear Factor 4alpha Prevents the Steatosis-to-NASH Progression by Regulating p53 and Bile Acid Signaling (in mice). *Hepatology* **73**, 2251-2265, doi:10.1002/hep.31604 (2021).

44 Patouraux, S. *et al.* CD44 is a key player in non-alcoholic steatohepatitis. *J Hepatol* **67**, 328-338, doi:10.1016/j.jhep.2017.03.003 (2017).

45 Chen, G. *et al.* CYP2J2 overexpression attenuates nonalcoholic fatty liver disease induced by high-fat diet in mice. *Am J Physiol Endocrinol Metab* **308**, E97-E110, doi:10.1152/ajpendo.00366.2014 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

46 Lefebvre, P. *et al.* Interspecies NASH disease activity whole-genome profiling identifies a fibrogenic role of PPARalpha-regulated dermatopontin. *JCI Insight* **2**, doi:10.1172/jci.insight.92264 (2017).

47 Xiao, W., Wang, X., Wang, T., Chen, B. & Xing, J. HAO2 inhibits malignancy of clear cell renal cell carcinoma by promoting lipid catabolic process. *J Cell Physiol* **234**, 23005-23016, doi:10.1002/jcp.28861 (2019).

48 Li, Z. *et al.* Methionine metabolism in chronic liver diseases: an update on molecular mechanism and therapeutic implication. *Signal Transduct Target Ther* **5**, 280, doi:10.1038/s41392-020-00349-7 (2020).

49 Rom, O. *et al.* Glycine-based treatment ameliorates NAFLD by modulating fatty acid oxidation, glutathione synthesis, and the gut microbiome. *Sci Transl Med* **12**, doi:10.1126/scitranslmed.aaz2841 (2020).

50 Krstic, J., Galhuber, M., Schulz, T. J., Schupp, M. & Prokesch, A. p53 as a Dichotomous Regulator of Liver Disease: The Dose Makes the Medicine. *Int J Mol Sci* **19**, doi:10.3390/ijms19030921 (2018).

51 Farrell, G. C. *et al.* Apoptosis in experimental NASH is associated with p53 activation and TRAIL receptor expression. *J Gastroenterol Hepatol* **24**, 443-452, doi:10.1111/j.1440-1746.2009.05785.x (2009).

52 Amanatidou, A. I. & Dedoussis, G. V. Construction and analysis of protein-protein interaction network of non-alcoholic fatty liver disease. *Comput Biol Med* **131**, 104243, doi:10.1016/j.compbiomed.2021.104243 (2021).

53 Hasenfuss, S. C. *et al.* Regulation of steatohepatitis and PPARgamma signaling by distinct AP-1 dimers. *Cell Metab* **19**, 84-95, doi:10.1016/j.cmet.2013.11.018 (2014).

54 Zhu, C. *et al.* Hepatocyte Notch activation induces liver fibrosis in nonalcoholic steatohepatitis. *Sci Transl Med* **10**, doi:10.1126/scitranslmed.aat0344 (2018).

55 Ramachandran, P. *et al.* Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* **575**, 512-518, doi:10.1038/s41586-019-1631-3 (2019).

56 Eferl, R. & Wagner, E. F. AP-1: a double-edged sword in tumorigenesis. *Nat Rev Cancer* **3**, 859-868, doi:10.1038/nrc1209 (2003).

57 Katsuoka, F. & Yamamoto, M. Small Maf proteins (MafF, MafG, MafK): History, structure and function. *Gene* **586**, 197-205, doi:10.1016/j.gene.2016.03.058 (2016).

58 Wang, X., Cornelis, F. M. F., Lories, R. J. & Monteagudo, S. Exostosin-1 enhances canonical Wnt signaling activity during chondrogenic differentiation. *Osteoarthritis Cartilage* **27**, 1702-1710, doi:10.1016/j.joca.2019.07.007 (2019).

59 Simoes-Costa, M., Stone, M. & Bronner, M. E. Axud1 Integrates Wnt Signaling and Transcriptional Inputs to Drive Neural Crest Formation. *Dev Cell* **34**, 544-554, doi:10.1016/j.devcel.2015.06.024 (2015).

60 Ye, X. *et al.* Axin1 up-regulated 1 accelerates stress-induced cardiomyocytes apoptosis through activating Wnt/beta-catenin signaling. *Exp Cell Res* **359**, 441-448, doi:10.1016/j.yexcr.2017.08.027 (2017).

61 Yang, M. *et al.* Hepatic E4BP4 induction promotes lipid accumulation by suppressing AMPK signaling in response to chemical or diet-induced ER stress. *FASEB J* **34**, 13533-13547, doi:10.1096/fj.201903292RR (2020).

62 Ceccarelli, S. *et al.* LPS-induced TNF-alpha factor mediates pro-inflammatory and pro-fibrogenic pattern in non-alcoholic fatty liver disease. *Oncotarget* **6**, 41434-41452, doi:10.18632/oncotarget.5163 (2015).

63 De Santa, F. *et al.* The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* **130**, 1083-1094, doi:10.1016/j.cell.2007.08.019 (2007).

64 De Santa, F. *et al.* Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. *EMBO J* **28**, 3341-3352, doi:10.1038/emboj.2009.271 (2009).

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

65 White, P. J. *et al.* Muscle-Liver Trafficking of BCAA-Derived Nitrogen Underlies Obesity-Related Glycine Depletion. *Cell Rep* **33**, 108375, doi:10.1016/j.celrep.2020.108375 (2020).
66 Hakkola, J., Rysa, J. & Hukkanen, J. Regulation of hepatic energy metabolism by the nuclear receptor PXR. *Biochim Biophys Acta* **1859**, 1072-1082, doi:10.1016/j.bbagrm.2016.03.012 (2016).
67 Sun, M., Cui, W., Woody, S. K. & Staudinger, J. L. Pregnane X receptor modulates the inflammatory response in primary cultures of hepatocytes. *Drug Metab Dispos* **43**, 335-343, doi:10.1124/dmd.114.062307 (2015).
68 Kotani, K., Watanabe, J., Miura, K. & Gugliucci, A. Paraoxonase 1 and Non-Alcoholic Fatty Liver Disease: A Meta-Analysis. *Molecules* **26**, doi:10.3390/molecules26082323 (2021).
69 Polyzos, S. A. *et al.* Selenoprotein P in Patients with Nonalcoholic Fatty Liver Disease. *Exp Clin Endocrinol Diabetes* **127**, 598-602, doi:10.1055/a-0811-9136 (2019).
70 Nikolaou, N. *et al.* AKR1D1 is a novel regulator of metabolic phenotype in human hepatocytes and is dysregulated in non-alcoholic fatty liver disease. *Metabolism* **99**, 67-80, doi:10.1016/j.metabol.2019.153947 (2019).
71 Shin, S. K. *et al.* Ablation of catalase promotes non-alcoholic fatty liver via oxidative stress and mitochondrial dysfunction in diet-induced obese mice. *Pflugers Arch* **471**, 829-843, doi:10.1007/s00424-018-02250-3 (2019).
72 Park, H. W. *et al.* Hepatoprotective role of Sestrin2 against chronic ER stress. *Nat Commun* **5**, 4233, doi:10.1038/ncomms5233 (2014).
73 Cordoba-Chacon, J. *et al.* Adult-Onset Hepatocyte GH Resistance Promotes NASH in Male Mice, Without Severe Systemic Metabolic Dysfunction. *Endocrinology* **159**, 3761-3774, doi:10.1210/en.2018-00669 (2018).
74 Lin, H. Y. *et al.* Increased hepatic steatosis and insulin resistance in mice lacking hepatic androgen receptor. *Hepatology* **47**, 1924-1935, doi:10.1002/hep.22252 (2008).
75 Lin, H. Y. *et al.* Insulin and leptin resistance with hyperleptinemia in mice lacking androgen receptor. *Diabetes* **54**, 1717-1725, doi:10.2337/diabetes.54.6.1717 (2005).
76 Gonzalez-Rodriguez, A. *et al.* Impaired autophagic flux is associated with increased endoplasmic reticulum stress during the development of NAFLD. *Cell Death Dis* **5**, e1179, doi:10.1038/cddis.2014.162 (2014).
77 Nagahara, R. *et al.* Gene Expression Analysis of the Activating Factor 3/Nuclear Protein 1 Axis in a Non-alcoholic Steatohepatitis Mouse Model. *Yonago Acta Med* **62**, 36-46 (2019).
78 Schulien, I. *et al.* The transcription factor c-Jun/AP-1 promotes liver fibrosis during non-alcoholic steatohepatitis by regulating Osteopontin expression. *Cell Death Differ* **26**, 1688-1699, doi:10.1038/s41418-018-0239-8 (2019).
79 Choi, E. *et al.* Expression patterns of STAT3, ERK and estrogen-receptor alpha are associated with development and histologic severity of hepatic steatosis: a retrospective study. *Diagn Pathol* **13**, 23, doi:10.1186/s13000-018-0698-8 (2018).
80 Grohmann, M. *et al.* Obesity Drives STAT-1-Dependent NASH and STAT-3-Dependent HCC. *Cell* **175**, 1289-1306 e1220, doi:10.1016/j.cell.2018.09.053 (2018).
81 Long, Z. *et al.* Inhibition of hepatocyte nuclear factor 1b induces hepatic steatosis through DPP4/NOX1-mediated regulation of superoxide. *Free Radic Biol Med* **113**, 71-83, doi:10.1016/j.freeradbiomed.2017.09.016 (2017).
82 Wolfrum, C. & Stoffel, M. Coactivation of Foxa2 through Pgc-1beta promotes liver fatty acid oxidation and triglyceride/VLDL secretion. *Cell Metab* **3**, 99-110, doi:10.1016/j.cmet.2006.01.001 (2006).
83 Mardinoglu, A. *et al.* Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nat Commun* **5**, 3083, doi:10.1038/ncomms4083 (2014).
84 White, P. J. *et al.* The BCKDH Kinase and Phosphatase Integrate BCAA and Lipid Metabolism via Regulation of ATP-Citrate Lyase. *Cell Metab* **27**, 1281-1293 e1287, doi:10.1016/j.cmet.2018.04.015 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

85 Lake, A. D. *et al.* Branched chain amino acid metabolism profiles in progressive human nonalcoholic fatty liver disease. *Amino Acids* **47**, 603-615, doi:10.1007/s00726-014-1894-9 (2015).
86 Walker, A. K. 1-Carbon Cycle Metabolites Methylate Their Way to Fatty Liver. *Trends Endocrinol Metab* **28**, 63-72, doi:10.1016/j.tem.2016.10.004 (2017).
87 Akazawa, Y. & Nakao, K. To die or not to die: death signaling in nonalcoholic fatty liver disease. *J Gastroenterol* **53**, 893-906, doi:10.1007/s00535-018-1451-5 (2018).
88 Aravinthan, A. *et al.* Gene polymorphisms of cellular senescence marker p21 and disease progression in non-alcohol-related fatty liver disease. *Cell Cycle* **13**, 1489-1494, doi:10.4161/cc.28471 (2014).
89 Min-DeBartolo, J. *et al.* Thrombospondin-I is a critical modulator in non-alcoholic steatohepatitis (NASH). *PLoS One* **14**, e0226854, doi:10.1371/journal.pone.0226854 (2019).
90 Tomita, K. *et al.* Tumour necrosis factor alpha signalling through activation of Kupffer cells plays an essential role in liver fibrosis of non-alcoholic steatohepatitis in mice. *Gut* **55**, 415-424, doi:10.1136/gut.2005.071118 (2006).
91 Crespo, J. *et al.* Gene expression of tumor necrosis factor alpha and TNF-receptors, p55 and p75, in nonalcoholic steatohepatitis patients. *Hepatology* **34**, 1158-1163, doi:10.1053/jhep.2001.29628 (2001).
92 Arif, E. *et al.* Targeting myosin 1c inhibits murine hepatic fibrogenesis. *Am J Physiol Gastrointest Liver Physiol* **320**, G1044-G1053, doi:10.1152/ajpgi.00105.2021 (2021).
93 Le Guilcher, C. *et al.* The P2X4 purinergic receptor regulates hepatic myofibroblast activation during liver fibrogenesis. *J Hepatol* **69**, 644-653, doi:10.1016/j.jhep.2018.05.020 (2018).
94 Flach, R. J., Qin, H., Zhang, L. & Bennett, A. M. Loss of mitogen-activated protein kinase phosphatase-1 protects from hepatic steatosis by repression of cell death-inducing DNA fragmentation factor A (DFFA)-like effector C (CIDEC)/fat-specific protein 27. *J Biol Chem* **286**, 22195-22202, doi:10.1074/jbc.M110.210237 (2011).
95 Liang, D. *et al.* Inhibition of EGFR attenuates fibrosis and stellate cell activation in diet-induced model of nonalcoholic fatty liver disease. *Biochim Biophys Acta Mol Basis Dis* **1864**, 133-142, doi:10.1016/j.bbadis.2017.10.016 (2018).
96 Means, A. L. *et al.* Overexpression of heparin-binding EGF-like growth factor in mouse pancreas results in fibrosis and epithelial metaplasia. *Gastroenterology* **124**, 1020-1036, doi:10.1053/gast.2003.50150 (2003).
97 El Ouaamari, A. *et al.* SerpinB1 Promotes Pancreatic beta Cell Proliferation. *Cell Metab* **23**, 194-205, doi:10.1016/j.cmet.2015.12.001 (2016).
98 Ahrens, M. *et al.* DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. *Cell Metab* **18**, 296-302, doi:10.1016/j.cmet.2013.07.004 (2013).
99 Xu, G. *et al.* Overexpression of the Lias gene attenuates hepatic steatosis in Leprdb/db mice. *J Endocrinol* **248**, 119-131, doi:10.1530/JOE-19-0606 (2021).
100 Thi Thanh Hai, N. *et al.* Selective overexpression of cytoglobin in stellate cells attenuates thioacetamide-induced liver fibrosis in mice. *Sci Rep* **8**, 17860, doi:10.1038/s41598-018-36215-4 (2018).
101 Wang, Y. *et al.* Adult Hepatocytes Are Hedgehog-Responsive Cells in the Setting of Liver Injury: Evidence for Smoothened-Mediated Activation of NF-kappaB/Epidermal Growth Factor Receptor/Akt in Hepatocytes that Counteract Fas-Induced Apoptosis. *Am J Pathol* **188**, 2605-2616, doi:10.1016/j.ajpath.2018.07.018 (2018).
102 Hasegawa, T. *et al.* Changed Amino Acids in NAFLD and Liver Fibrosis: A Large Cross-Sectional Study without Influence of Insulin Resistance. *Nutrients* **12**, doi:10.3390/nu12051450 (2020).
103 Mardinoglu, A. *et al.* Personal model-assisted identification of NAD(+) and glutathione metabolism as intervention target in NAFLD. *Mol Syst Biol* **13**, 916, doi:10.15252/msb.20167422 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

104 Deminice, R., de Castro, G. S., Brosnan, M. E. & Brosnan, J. T. Creatine supplementation as a possible new therapeutic approach for fatty liver disease: early findings. *Amino Acids* **48**, 1983-1991, doi:10.1007/s00726-016-2183-6 (2016).
105 Mechanick, J. I. *et al.* Clinical Practice Guidelines for the Perioperative Nutrition, Metabolic, and Nonsurgical Support of Patients Undergoing Bariatric Procedures - 2019 Update: Cosponsored by American Association of Clinical Endocrinologists/American College of Endocrinology, the Obesity Society, American Society for Metabolic & Bariatric Surgery, Obesity Medicine Association, and American Society of Anesthesiologists - Executive Summary. *Endocr Pract* **25**, 1346-1359, doi:10.4158/GL-2019-0406 (2019).
106 Brunt, E. M., Janney, C. G., Di Bisceglie, A. M., Neuschwander-Tetri, B. A. & Bacon, B. R. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol* **94**, 2467-2474, doi:10.1111/j.1572-0241.1999.01377.x (1999).
107 White, P. J. *et al.* Branched-chain amino acid restriction in Zucker-fatty rats improves muscle insulin sensitivity by enhancing efficiency of fatty acid oxidation and acyl-glycine export. *Mol Metab* **5**, 538-551, doi:10.1016/j.molmet.2016.04.006 (2016).
108 Stockli, J. *et al.* Metabolomic analysis of insulin resistance across different mouse strains and diets. *J Biol Chem* **292**, 19135-19145, doi:10.1074/jbc.M117.818351 (2017).
109 Chu, T., Wang, Z., Chou, S. P. & Danko, C. G. Discovering Transcriptional Regulatory Elements From Run-On and Sequencing Data Using the Web-Based dREG Gateway. *Curr Protoc Bioinformatics* **66**, e70, doi:10.1002/cpbi.70 (2019).
110 Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863-864, doi:10.1093/bioinformatics/btr026 (2011).
111 Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595, doi:10.1093/bioinformatics/btp698 (2010).
112 Hahne, F. & Ivanek, R. Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol* **1418**, 335-351, doi:10.1007/978-1-4939-3578-9_16 (2016).
113 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
114 Wang, Z., Chu, T., Choate, L. A. & Danko, C. G. Identification of regulatory elements from nascent transcription using dREG. *Genome Res* **29**, 293-303, doi:10.1101/gr.238279.118 (2019).
115 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).
116 Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Curr Protoc* **1**, e90, doi:10.1002/cpz1.90 (2021).
117 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97, doi:10.1093/nar/gkw377 (2016).
118 Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128, doi:10.1186/1471-2105-14-128 (2013).
119 Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30, doi:10.1093/nar/28.1.27 (2000).
120 Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523, doi:10.1038/s41467-019-09234-6 (2019).

Figure legends

Figure 1. Distinct sets of genes are activated at the transcriptional level in human NASH-resistant (NR) and NASH-prone (NP) livers. (A) Clinic characterizations of human subjects with NR and NP phenotype. In BMI panel, BMI above 40 (dashed line) is a general criterion for eligibility of bariatric surgery operation. In fasting glucose panel, values above 5.6 mM (dashed line) indicates

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

hyperglycemia and risk for T2D. In HbA1c panel, values above 6.5 percent (dashed line) indicates poor glycemic control or T2D. * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001 by unpaired two sample T-test. See additional information in Supplementary Table 1. (B) Left: genotype of rs738409 at *PNPLA3* locus (CG, risk allele; CC, reference allele). Right: glycemic status of human participants with NR or NP phenotype. T2D, type 2 diabetes; IGT, impaired glucose tolerant; NGT, normal glucose tolerant. (C) Experimental scheme of chromatin run-on sequencing (ChRO-seq) profiling gene regulatory landscapes of primary liver biopsy samples from NR and NP subjects. (D) PCA plot of ChRO-seq gene transcription profile of NP and NR livers. Color denotes NP or NR phenotype. Shape denotes genotype of rs738409 at *PNPLA3* locus. (E) Volcano plot showing differentially transcribed genes between NP and NR livers. Gene with P < 0.05, adjusted P < 0.2, log2 foldchange > 0 (or < 0), average normalized counts > 100 were shown (Wald test; DESeq2). See complete gene lists in Supplementary Table 3 and 4. (F-G) Pathway enrichment analysis of NP-activated genes (n = 879) and NR-activated genes (n = 651) based on Gene Ontology (GO) Biological Process 2018. (H) Histogram showing foldchanges of transcription activity for selected genes in NR compared to NP livers. Asterisks highlight genes with P < 0.05, adjusted P < 0.2, log2 foldchange > 0 in NR compared to NP (Wald test; DESeq2). NASH-resistant (NR), n = 6; NASH-prone (NP), n = 8.

Figure 2. Define distinct active enhancers and the linked genes in human NASH-resistant (NR) and NASH-prone (NP) livers. (A) Hierarchical clustering analysis of active enhancers (n = 87010) defined by ChRO-seq. (B) Hierarchical clustering analysis of active promoters (n = 49853) defined by ChRO-seq. In (A-B), color denotes NP or NR phenotype. Shape denotes genotype of rs738409 at PNPLA3 locus (CG, risk allele; CC, reference allele). (C) Left: Ranked plot of all transcriptional regulatory elements (TREs), including both promoters and enhancers, based on foldchange of TRE activity in NR compared to NP livers. Dashed line denotes log2 foldchange at 1 and -1. Right: Donut plots showing numbers of differentially transcribed TREs between NP and NR livers. 1447 TREs (694 enhancers and 753 promoters) are significantly active in NP livers. 1833 TREs (1407 enhancers and 426 promoters) are significantly activated in NR livers. P < 0.05, adjusted P < 0.2, log2 foldchange > 1 or < -1 in NR compared to NP by Wald test; DESeq2. (D-E) Ridgeline plot of genes (vertical lines) associated with one or more NP-specific (D) or NR-specific enhancers (E). Plots were generated for comparison across different enhancer-density groups. (F) Dot plot of genes associated with NP enhancers. Size of the dot denotes enhancer density. Genes highlighted in blue are NP-activated genes (defined in Figure 1E). Also see Supplementary Table 3. (G) Dot plot of genes associated with NR enhancers. Size of the dot denotes enhancer density. Genes highlighted in purple are NR-activated genes (defined in Figure 1E). Also see Supplementary Table 4. In (D-G), only genes with average normalized counts > 100 across all samples were shown. (H) Transcription activity (normalized ChRO-seq signal), all TREs (regardless activity) and NP-specific enhancers around locus of *THBS1* are shown. (I) Transcription activity (normalized ChRO-seq signal), all TREs (regardless activity) and NR-specific enhancers around loci of *BHMT and BHMT2* are shown. NASH-resistant (NR), n = 6; NASH-prone (NP), n = 8.

Figure 3. Define NASH resistant (NR) or NASH prone (NP) specific-enhancer hotspots in human livers. (A) Stitched enhancers are ranked based on the sum of transcriptional activity from each constituent NP-specific enhancer. Blue dots represent NP-specific enhancer hotspots (n = 56) and gray dots are non-enhancer hotspots (n = 464). (B) Transcription fold change of genes associated with either NP-specific enhancer hotspots or non-enhancer hotspots. Blue dots highlight NP-activated genes (defined in Figure 1D). (C) NP-activated genes that are associated with NP-specific enhancer hotspots are shown. Lollipop plot height indicates the distance between NP enhancer hotspots and the associated transcription start sites (TSSs). Dot color denotes distance category and dot size denotes the activity of NP-specific enhancer hotspots. Also see Supplementary Table 5. (D-E) Transcription activity (normalized ChRO-seq signal), all TREs (regardless activity) and NP-specific enhancer hotspots (EH) around the loci of *GADD45B* (D) and *HES1* (E) are shown. (F) Stitched enhancers are ranked based on the sum of transcriptional activity from each constituent NR-

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

specific enhancer. Purple dots represent NR-specific enhancer hotspots (n = 95) and gray dots are non-enhancer hotspots (n = 951). (G) Transcription fold change of genes associated with either NR-specific enhancer hotspots or non-enhancer hotspots. Purple dots highlight NR-activated genes (defined in Figure 1D). (G) NR-activated genes that are associated with NR-specific enhancer hotspots are shown. Lollipop plot height indicates the distance between NR-specific enhancer hotspots and the associated TSSs. Dot color denotes distance category and dot size denotes transcriptional activity of NR-specific enhancer hotspots. Also see Supplementary Table 6. (H) Transcription activity (normalized ChRO-seq signal), all TREs (regardless activity) and NR-specific enhancer hotspots (EH) around the loci of *GLYAT* (I) and *GATM* (J) are shown. In (B, G), **** P < 0.0001 by unpaired two-sample Wilcoxon test. NASH-resistant (NR), n = 6; NASH-prone (NP), n = 8.

Figure 4. Define active gene networks of transcription factors (TFs) in NASH prone (NP) livers. (A) Motif enrichment analysis of NP-specific enhancers against the defined background enhancers. (B) TF motifs significantly enriched in NP-specific enhancers (p < 0.05, q < 0.05, enrichment foldchange > 1.5 by HOMER). Only TFs with high transcription levels in NP livers (normalized counts > 500 in NP livers) are shown. Also see Supplementary Figure 3. (C) Transcription foldchange of TFs in NP relative to NR livers in ChRO-seq data. Asterisks highlight genes with P < 0.05, adjusted P < 0.2, log2 foldchange < 0 in NR compared to NP (Wald test; DESeq2). (D) Upset plot showing NP-activated genes shared across multiple TF networks or uniquely associated with a specific TF in NP livers: ATF3, JUNB, STAT3 and MAFK. Genes in blue represent those associated with NP-specific enhancer hotspots. Also see Supplementary Table 7. (E-G) Transcription activity (normalized ChRO-seq signal), All TREs (regardless activity) and NP-specific enhancers/enhancer hotspots (EH) around loci of *CDKN1A* (E), *CSRNP1* (F) and *FOS* (G) are shown. (H) Gene Ontology (GO) pathway enrichment analysis of gene network for each of the selected NP associated TFs. Only top-three ranked pathways for each TF network are shown in the heatmap. Color shade indicates ranks of pathways for each TF network. (I) A proposed NP-associated TF subnetwork based on Figure 3D and Figure 4D. JUNB, ATF3, STAT3 and MAFK (arrows indicated by different colors) coordinately drive activation of NP-activated genes by acting on nearby NP-specific enhancer/enhancer hotspots. NASH-resistant (NR), n = 6; NASH-prone (NP), n = 8.

Figure 5. Define active gene networks of transcription factors (TFs) in NASH-resistant (NR) livers. (A) Motif enrichment analysis of NR-specific enhancers against the defined background enhancers. (B) TF motifs significantly enriched in NR-specific enhancers (p < 0.05, q < 0.05, enrichment foldchange > 1.5 by HOMER). Only TFs with high transcription levels in NR livers (normalized counts > 500 in NR livers) are shown. (C) Transcription foldchange of TFs in NP relative to NR livers in ChRO-seq data. Asterisks highlight genes with P < 0.05, adjusted P < 0.2, log2 foldchange > 0 in NR compared to NP (Wald test; DESeq2). (D) Upset plot showing NR-activated genes shared across multiple TF networks or uniquely associated with a specific TF in NR livers: HNF4A, HNF1B, FOXA2 and AR. Genes in purple represent those associated with NR-specific enhancer hotspots. Also see Supplementary Table 8. (E) Gene Ontology (GO) pathway enrichment analysis of gene network for each of the selected NR associated TFs. Only top three ranked pathways for each TF network are shown in the heatmap. Color shade indicates ranks of pathways for each TF network. (F) Protein-protein interaction (PPI) enrichment analysis of all NR-activated genes under one or more networks of TFs (HNF4A, HNF1B, FOXA2 or AR). The resulting key gene modules and the associated biological processes are shown. (G-H) Transcription activity (normalized ChRO-seq signal), All TREs (regardless activity) and NR-specific enhancers around loci of *SARDH* (G) and *SHMT1* (H) are shown. NASH resistant (NR), n = 6; NASH-prone (NP), n = 8.

Figure 6. Human NASH-resistant (NR) livers are featured with active enhancers and gene transcription that promote serine and glycine metabolism. (A) Circulating amino acid profiling of NR compared with NP subjects. Color shade indicates fold change in NR relative to NP subjects. Significantly altered amino acids are denoted by P values (unpaired two sample T-test). (B) Genes

bioRxiv preprint doi: https://doi.org/10.1101/2021.08.20.457162; this version posted August 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

involved in serine/glycine biosynthesis and utilization are significantly activated in NP livers, associated with NR-specific enhancers/enhancer hotspots, and/or positively correlated with circulating serine. NASH-resistant (NR), n = 6; NASH-prone (NP), n = 8.

Supplementary Figure 1. (A) Genomic location of active transcriptional regulatory elements (TREs, n = 136863) identified across all samples by ChRO-seq. (B) Size distribution of TREs. (C) Number of genes that are associated with both NR and NP enhancers, NR-specific enhancer(s) only, NP-specific enhancer(s) only. (D) Transcription activity (normalized ChRO-seq signal), all TREs (regardless activity) and NP-specific enhancers around locus of *CD44* are shown. (E) Transcription activity (normalized ChRO-seq signal), all TREs (regardless activity) and NR-specific enhancers around locus *HAO2* are shown. NASH-resistant (NR), n = 6; NASH-prone (NP), n = 8.

Supplementary Figure 2. (A) Summary of defining NP-specific enhancer hotspots and the associated genes. (B) Summary of defining NR-specific enhancer hotspots and the associated genes.

Supplementary Figure 3. Complete list of TF motifs significantly enriched in NP-specific enhancers ($p < 0.05$, $q < 0.05$, enrichment foldchange > 1.5 by HOMER). Only TFs with high transcription levels in NP livers (normalized counts > 500 in NP livers) are shown.

Supplementary Table 1. Clinical characterization of obese individuals with livers stratified into either NP or NP phenotype.

Supplementary Table 2. Summary of ChRO-seq mapping statistics and identification of active transcription regulatory elements (TREs).

Supplementary Table 3. List of NP-activated genes and the number of associated NP enhancers.

Supplementary Table 4. List of NR-activated genes and the number of associated NR enhancers.

Supplementary Table 5. NP-specific enhancer hotspots.

Supplementary Table 6. NR-specific enhancer hotspots.

Supplementary Table 7. Active gene networks of selected NP associated TFs (ATF3, JUNB, STAT3 and MAFK). Numbers of total motif sequences present in NP-specific enhancers nearby a given gene are provided.

Supplementary Table 8. Active gene networks of selected NR associated TFs (HNF4A, HNF1B, FOXA2 and AR). Numbers of total motif sequences present in NR-specific enhancers nearby a given gene are provided.

Figure 1

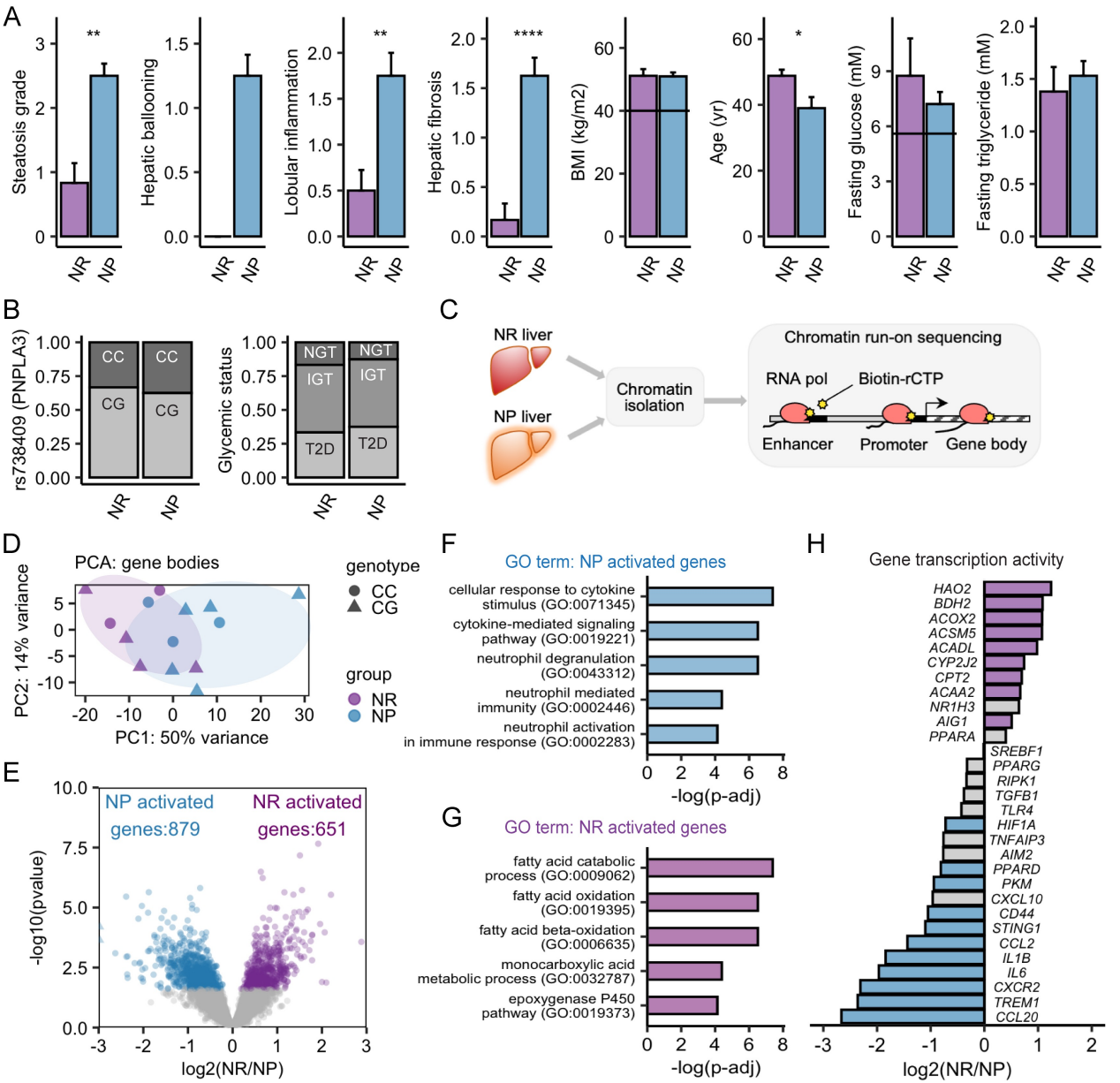

A
Steatosis grade
Hepatic ballooning
Lobular inflammation
Hepatic fibrosis
BMI (kg/m2)
Age (yr)
Fasting glucose (mM)
Fasting triglyceride (mM)
NR
NP
B
rs738409 (PNPLA3)
CC
CG
Glycemic status
NGT
IGT
T2D
C
NR liver
NP liver
Chromatin isolation
Chromatin run-on sequencing
RNA pol
Biotin-rCTP
Enhancer
Promoter
Gene body
D
PCA: gene bodies
PC2: 14% variance
PC1: 50% variance
genotype
group
E
NP activated genes:879
NR activated genes:651
-log10(pvalue)
log2(NR/NP)
F
GO term: NP activated genes
cellular response to cytokine stimulus (GO:0071345)
cytokine-mediated signaling pathway (GO:0019221)
neutrophil degranulation (GO:0043312)
neutrophil mediated immunity (GO:0002446)
neutrophil activation in immune response (GO:0002283)
-log(p-adj)
G
GO term: NR activated genes
fatty acid catabolic process (GO:0009062)
fatty acid oxidation (GO:0019395)
fatty acid beta-oxidation (GO:0006635)
monocarboxylic acid metabolic process (GO:0032787)
epoxygenase P450 pathway (GO:0019373)
H
Gene transcription activity
HAO2
BDH2
ACOX2
ACSM5
ACADL
CYP2J2
CPT2
ACAA2
NR1H3
AIG1
PPARA
SREBF1
PPARG
RIPK1
TGFB1
TLR4
HIF1A
TNFAIP3
AIM2
PPARD
PKM
CXCL10
CD44
STING1
CCL2
IL1B
IL6
CXCR2
TREM1
CCL20
log2(NR/NP)

Figure 2

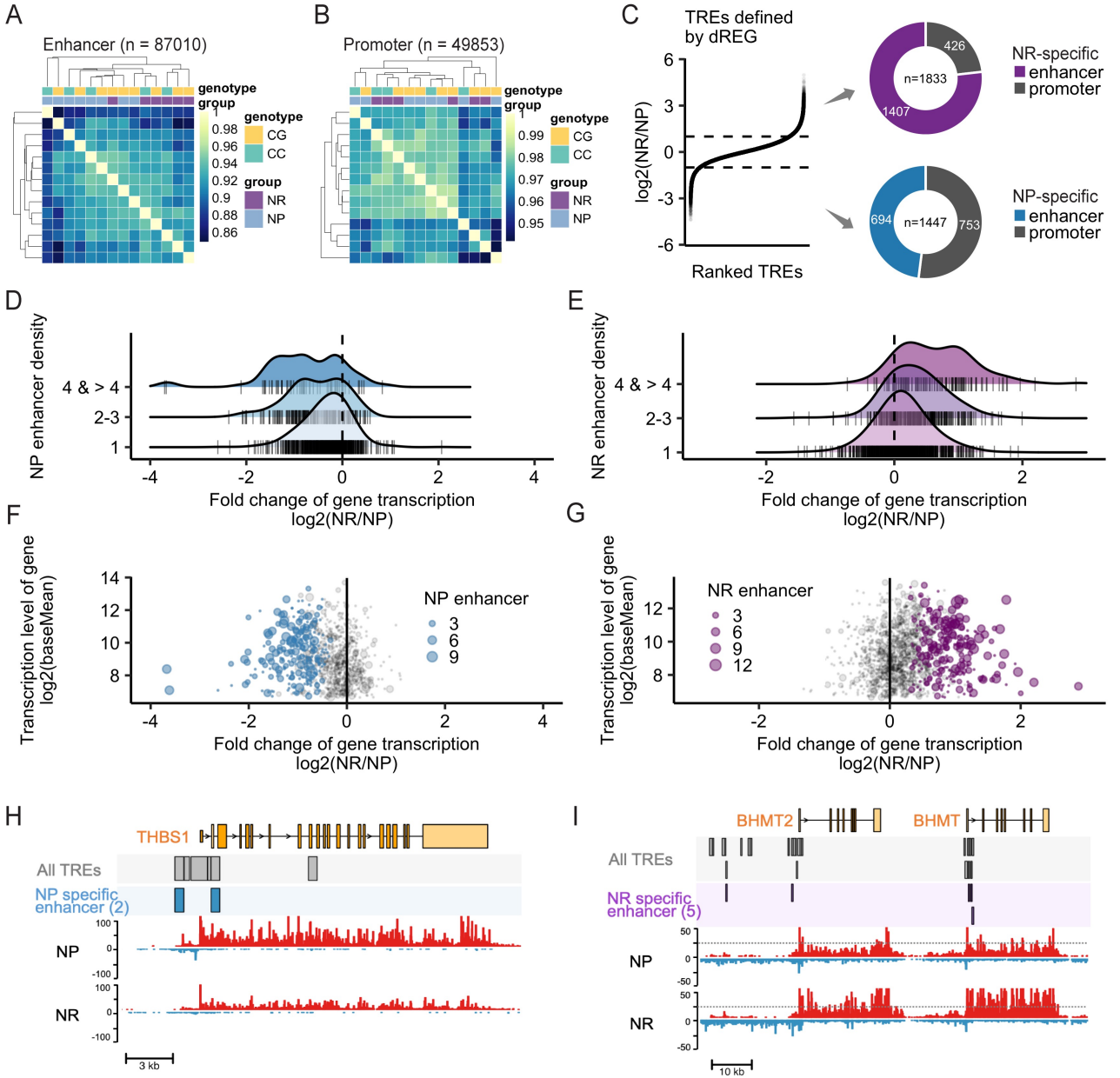

A
Enhancer (n = 87010)
genotype
group
genotype
CG
CC
group
NR
NP
B
Promoter (n = 49853)
C
TREs defined by dREG
log2(NR/NP)
Ranked TREs
n=1833
426
1407
NR-specific
enhancer
promoter
n=1447
694
753
NP-specific
D
NP enhancer density
4 & > 4
2-3
1
Fold change of gene transcription log2(NR/NP)
E
NR enhancer density
F
Transcription level of gene log2(baseMean)
NP enhancer
3
6
9
G
NR enhancer
12
H
THBS1
All TREs
NP specific enhancer (2)
NP
NR
3 kb
I
BHMT2
BHMT
NR specific enhancer (5)
10 kb

Figure 3

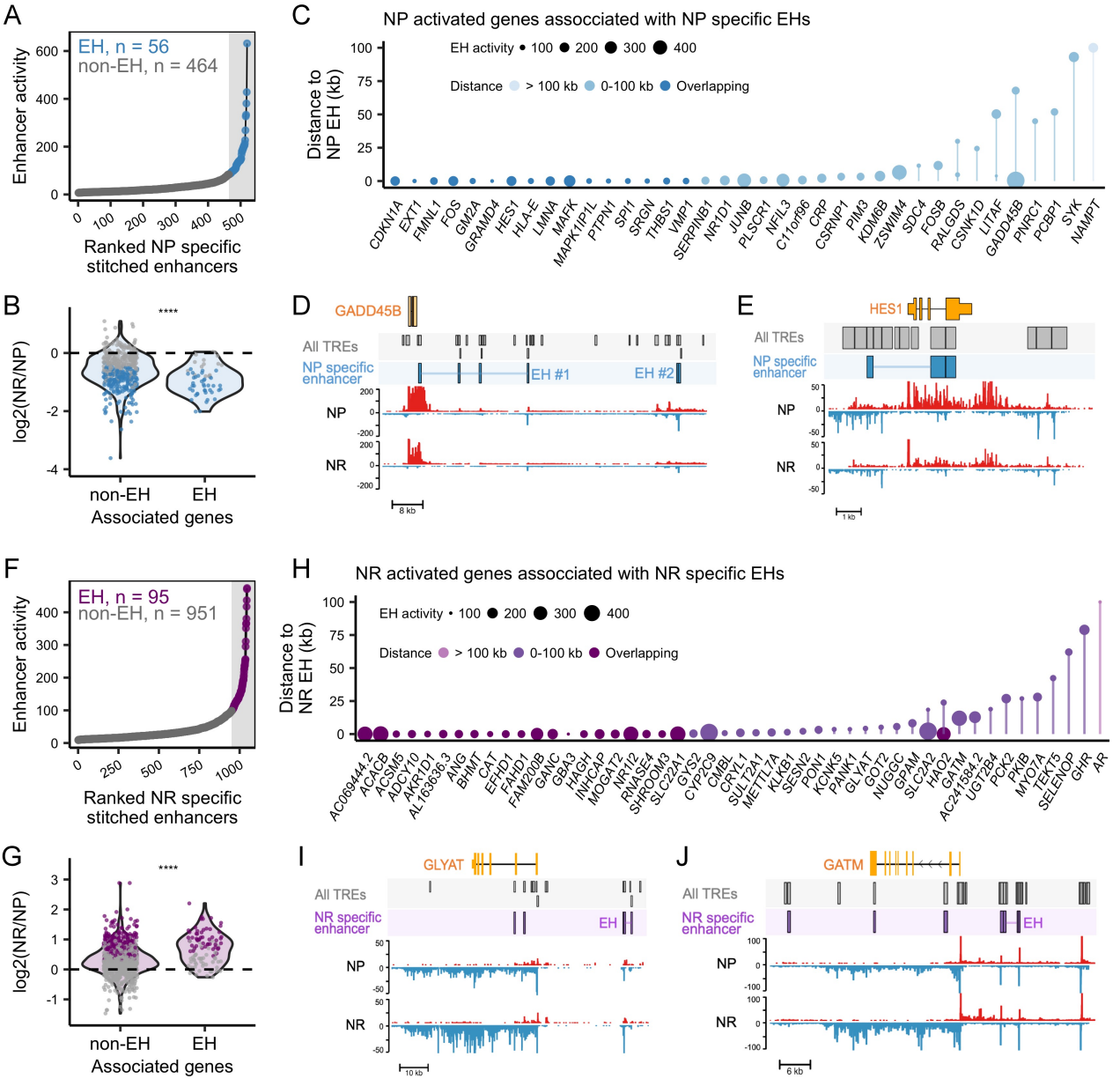

A
EH, n = 56
non-EH, n = 464
Enhancer activity
Ranked NP specific stitched enhancers
B
log2(NR/NP)
****
non-EH
EH
Associated genes
C
NP activated genes assocciated with NP specific EHs
EH activity 100 200 300 400
Distance > 100 kb 0-100 kb Overlapping
Distance to NP EH (kb)
CDKN1A EXT1 FMNL1 FOS GM2A GRAMD4 HES1 HLA-E LMNA MAFK MAPK1IP1L PTPN1 SPI1 SRGN THBS1 VMP1 SERPINB1 NR1D1 JUNB PLSCR1 NFIL3 C11orf96 CRP CSRNP1 PIM3 KDM6B ZSWIM4 SDC4 FOSB RALGDS CSNK1D LITAF GADD45B PNRC1 PCBP1 SYK NAMPT
D
GADD45B
All TREs
NP specific enhancer
EH #1
EH #2
NP
NR
8 kb
E
HES1
All TREs
NP specific enhancer
NP
NR
1 kb
F
EH, n = 95
non-EH, n = 951
Enhancer activity
Ranked NR specific stitched enhancers
G
log2(NR/NP)
****
non-EH
EH
Associated genes
H
NR activated genes assocciated with NR specific EHs
EH activity 100 200 300 400
Distance > 100 kb 0-100 kb Overlapping
Distance to NR EH (kb)
AC089444.2 ACACB ACSM5 ADCY10 AKR1D1 AL163536.3 ANG BHMT CAT EFHD1 FAHD1 FAM200B GANC GBA3 HAGH INHCAP MOGAT2 NR1I2 RNASE4 SHROOM3 SLC22A1 GYS2 CYP2C9 CMBL CRYL1 SULT2A1 METTL7A KLKB1 SESN2 PON1 KCNK5 PANK1 GLYAT GOT2 NUGGC GPAM SLC2A2 HAO2 GATM AC241584.2 UGT2B4 PCK2 PKIB MYOTA TEKT5 SELENOP GHR AR
I
GLYAT
All TREs
NR specific enhancer
EH
NP
NR
10 kb
J
GATM
All TREs
NR specific enhancer
EH
NP
NR
6 kb

Figure 4

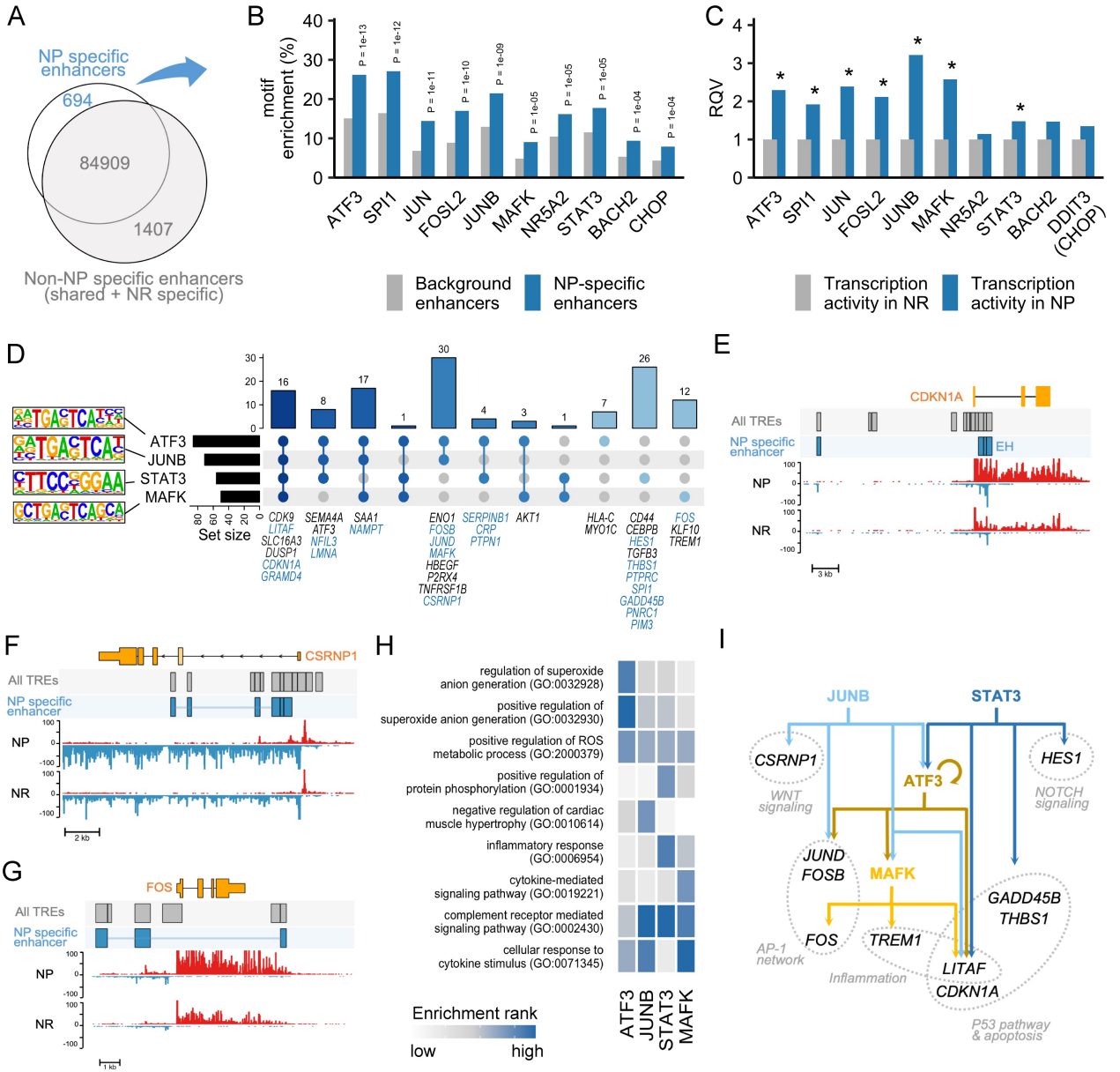

A
NP specific enhancers
694
84909
1407
Non-NP specific enhancers (shared + NR specific)
B
motif enrichment (%)
P = 1e-13
P = 1e-12
P = 1e-11
P = 1e-10
P = 1e-09
P = 1e-05
P = 1e-05
P = 1e-05
P = 1e-04
P = 1e-04
ATF3 SPI1 JUN FOSL2 JUNB MAFK NR5A2 STAT3 BACH2 CHOP
Background enhancers
NP-specific enhancers
C
RQV
ATF3 SPI1 JUN FOSL2 JUNB MAFK NR5A2 STAT3 BACH2 DDIT3 (CHOP)
Transcription activity in NR
Transcription activity in NP
D
ATF3
JUNB
STAT3
MAFK
Set size
16 8 17 1 30 4 3 1 7 26 12
CDK9 LITAF SLC16A3 DUSP1 CDKN1A GRAMD4
SEMA4A ATF3 NFIL3 LMNA
SAA1 NAMPT
ENO1 FOSB JUND MAFK HBEGF P2RX4 TNFRSF1B CSRNP1
SERPINB1 CRP PTPN1
AKT1
HLA-C MYO1C
CD44 CEBPB HES1 TGFB3 THBS1 PTPRC SPI1 GADD45B PNRC1 PIM3
FOS KLF10 TREM1
E
CDKN1A
All TREs
NP specific enhancer
EH
NP
NR
3 kb
F
CSRNP1
All TREs
NP specific enhancer
NP
NR
2 kb
G
FOS
All TREs
NP specific enhancer
NP
NR
1 kb
H
regulation of superoxide anion generation (GO:0032928)
positive regulation of superoxide anion generation (GO:0032930)
positive regulation of ROS metabolic process (GO:2000379)
positive regulation of protein phosphorylation (GO:0001934)
negative regulation of cardiac muscle hypertrophy (GO:0010614)
inflammatory response (GO:0006954)
cytokine-mediated signaling pathway (GO:0019221)
complement receptor mediated signaling pathway (GO:0002430)
cellular response to cytokine stimulus (GO:0071345)
ATF3 JUNB STAT3 MAFK
Enrichment rank
low
high
I
JUNB
STAT3
CSRNP1
WNT signaling
ATF3
HES1
NOTCH signaling
JUND FOSB
MAFK
GADD45B THBS1
FOS
TREM1
AP-1 network
Inflammation
LITAF CDKN1A
P53 pathway & apoptosis

Figure 5

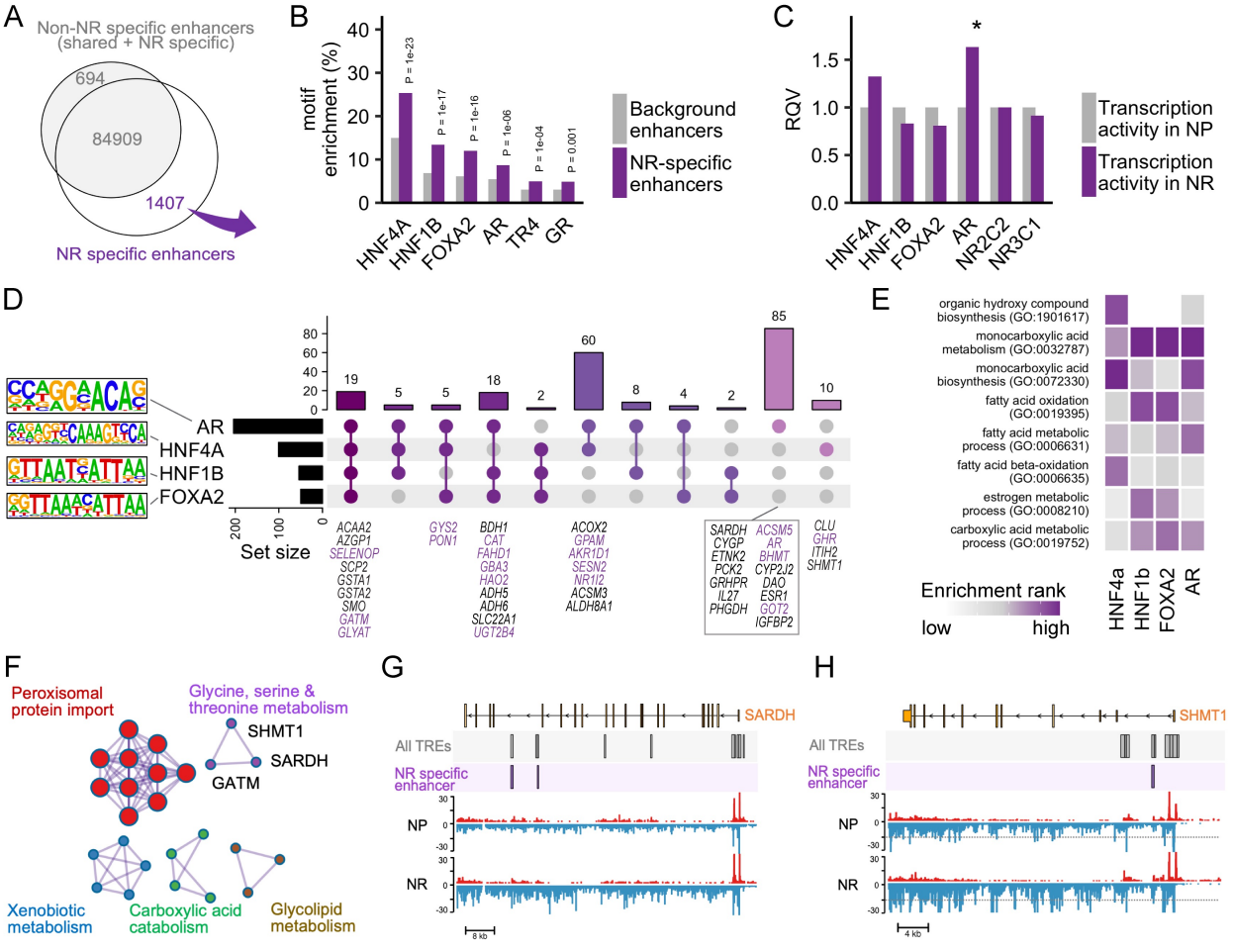
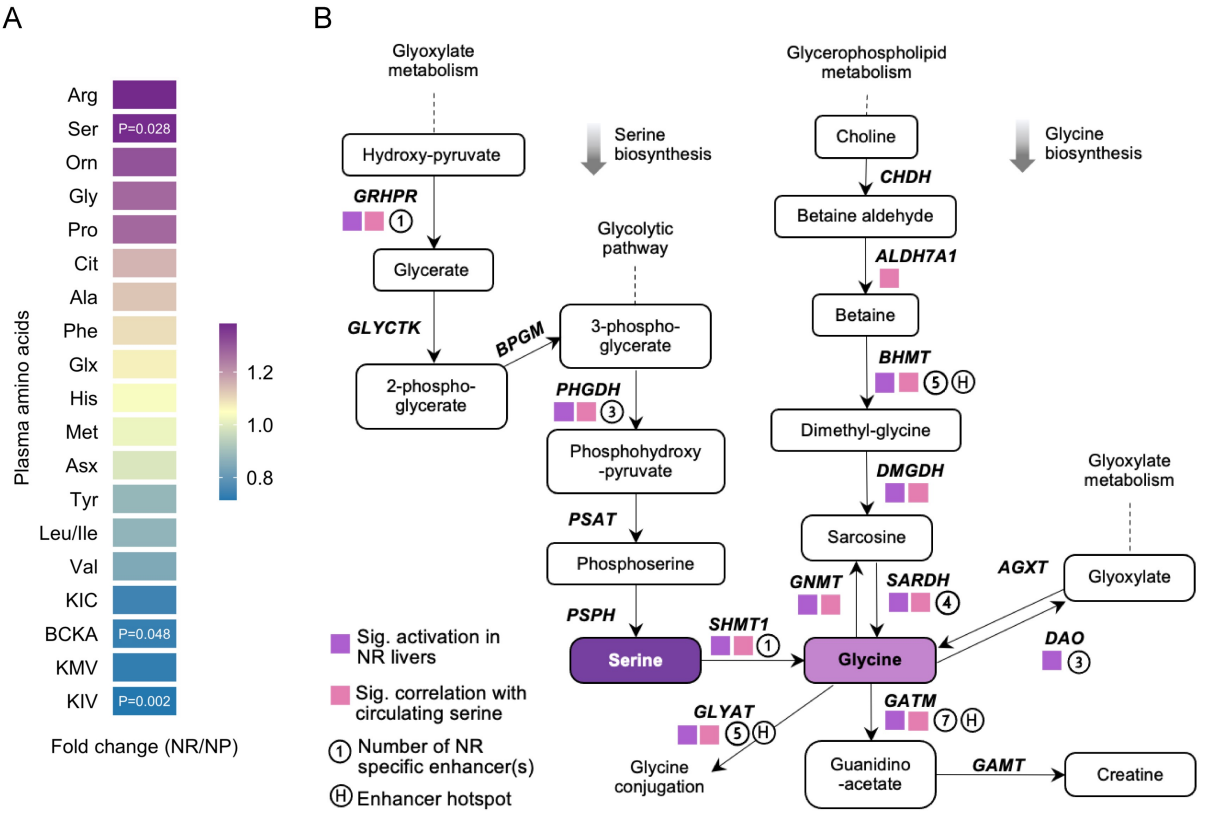

A
Non-NR specific enhancers
(shared + NR specific)
694
84909
1407
NR specific enhancers
B
motif enrichment (%)
P = 1e-23
P = 1e-17
P = 1e-16
P = 1e-06
P = 1e-04
P = 0.001
HNF4A
HNF1B
FOXA2
AR
TR4
GR
Background enhancers
NR-specific enhancers
C
RQV
*
HNF4A
HNF1B
FOXA2
AR
NR2C2
NR3C1
Transcription activity in NP
Transcription activity in NR
D
AR
HNF4A
HNF1B
FOXA2
Set size
ACAA2 AZGP1 SELENOP SCP2 GSTA1 GSTA2 SMO GATM GLYAT
GYS2 PON1
BDH1 CAT FAHD1 GBA3 HAO2 ADH5 ADH6 SLC22A1 UGT2B4
ACOX2 GPAM AKR1D1 SESN2 NR1I2 ACSM3 ALDH8A1
SARDH CYGP ETNK2 PCK2 GRHPR IL27 PHGDH
ACSM5 AR BHMT CYP2J2 DAO ESR1 GOT2 IGFBP2
CLU GHR ITIH2 SHMT1
E
organic hydroxy compound biosynthesis (GO:1901617)
monocarboxylic acid metabolism (GO:0032787)
monocarboxylic acid biosynthesis (GO:0072330)
fatty acid oxidation (GO:0019395)
fatty acid metabolic process (GO:0006631)
fatty acid beta-oxidation (GO:0006635)
estrogen metabolic process (GO:0008210)
carboxylic acid metabolic process (GO:0019752)
Enrichment rank
low
high
HNF4a
HNF1b
FOXA2
AR
F
Peroxisomal protein import
Glycine, serine & threonine metabolism
SHMT1
SARDH
GATM
Xenobiotic metabolism
Carboxylic acid catabolism
Glycolipid metabolism
G
SARDH
All TREs
NR specific enhancer
NP
NR
8 kb
H
SHMT1
All TREs
NR specific enhancer
NP
NR
4 kb

Figure 6

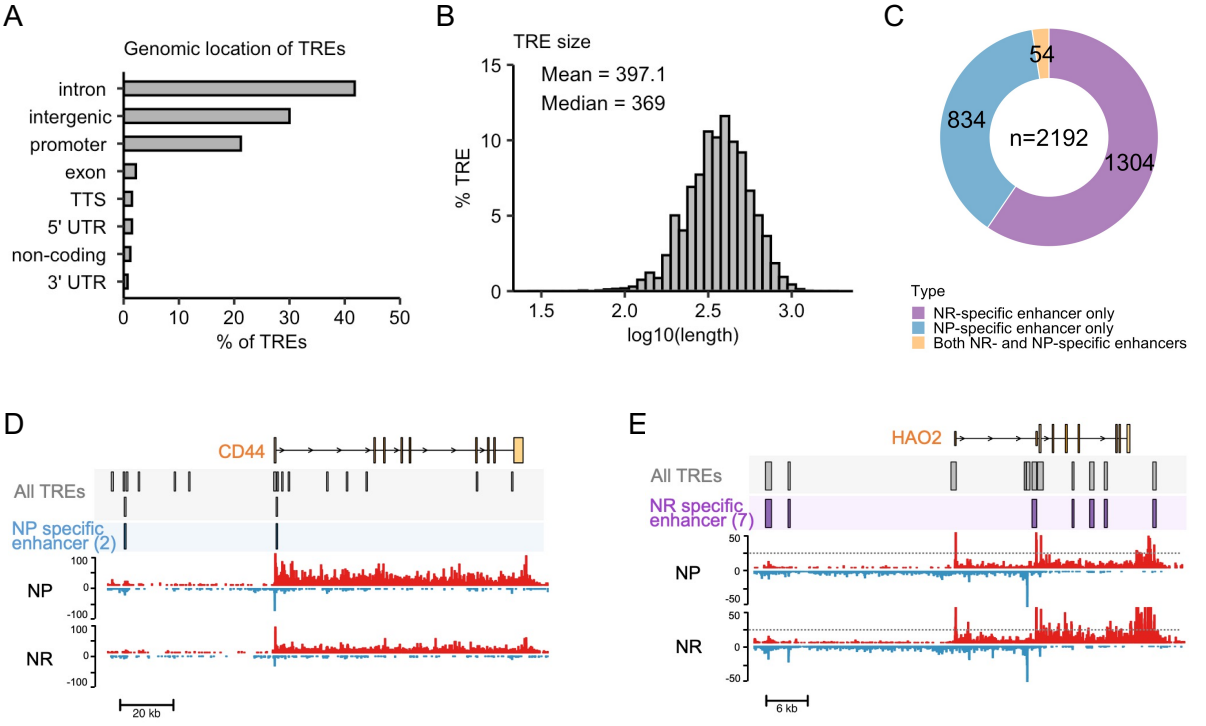

A
Arg
Ser
P=0.028
Orn
Gly
Pro
Cit
Ala
Phe
Glx
His
Met
Asx
Tyr
Leu/Ile
Val
KIC
BCKA
P=0.048
KMV
KIV
P=0.002
Plasma amino acids
1.2
1.0
0.8
Fold change (NR/NP)
B
Glyoxylate metabolism
Hydroxy-pyruvate
GRHPR
Glycerate
GLYCTK
2-phospho-glycerate
BPGM
Glycolytic pathway
3-phospho-glycerate
PHGDH
Phosphohydroxy-pyruvate
PSAT
Phosphoserine
PSPH
Serine
Serine biosynthesis
Glycerophospholipid metabolism
Choline
CHDH
Betaine aldehyde
ALDH7A1
Betaine
BHMT
Dimethyl-glycine
DMGDH
Sarcosine
GNMT
SARDH
SHMT1
Glycine
Glycine biosynthesis
Glyoxylate metabolism
AGXT
Glyoxylate
DAO
GLYAT
Glycine conjugation
GATM
Guanidino-acetate
GAMT
Creatine
Sig. activation in NR livers
Sig. correlation with circulating serine
Number of NR specific enhancer(s)
Enhancer hotspot

Supplementary Table 1

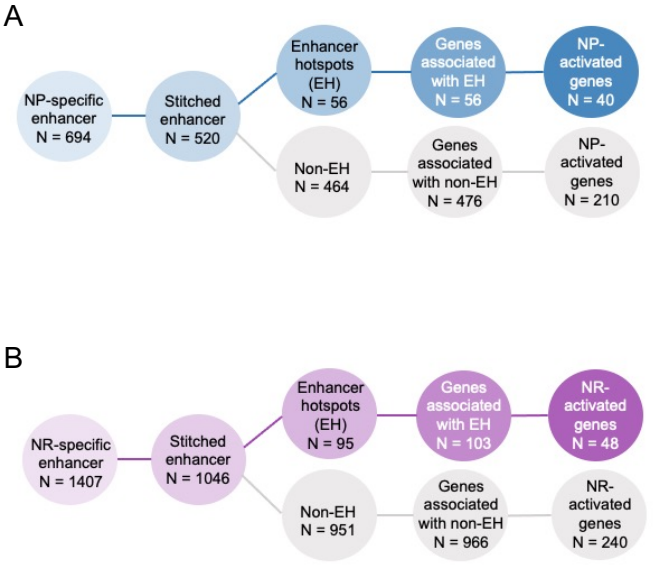
	NR	NP	P-value
Glycemic status	1 NG / 3 IGT / 2 T2D	1 NG/ 4 IGT / 3 T2D	NA
rs738409	2 CC / 4 CG	3 CC / 5 CG	NA
Gender (Female; Male)	6 F / 0 M	8 F / 0 M	NA
age	48.83 ± 4.54	39 ± 9.49	0.027
Height (cm)	156.83 ± 8.70	159.44 ± 3.98	0.519
Weight (Kg)	125.95 ± 17.56	129.66 ± 11.65	0.665
BMI	51.13 ± 5.22	50.95 ± 3.54	0.943
Waist Cir. (cm)	138.4 ± 15.34	137.75 ± 14.46	0.941
Hip Cir. (cm)	149.4 ± 13.99	148.13 ± 8.46	0.860
Waist-to-hip ratio	0.93 ± 0.08	0.93 ± 0.09	0.980
Hemoglobin A1C	0.069 ± 0.02	0.07 ± 0.01	0.812
Fasting plasma glucose (FPG; mM)	8.75 ± 4.99	7.21 ± 1.85	0.499
Total Cholesterol (mM)	4.04 ± 0.84	5.06 ± 0.74	0.040
HDL Cholesterol (mM)	1.27 ± 0.15	1.27 ± 0.21	0.956
LDL Cholesterol (mM)	2.15 ± 0.76	3.09 ± 0.81	0.047
Triglyceride (mM)	1.38 ± 0.57	1.53 ± 0.40	0.596
Total : HDL cholesterol	3.24 ± 0.87	4.08 ± 0.97	0.114
Smoking status (1 = smoker / 2 = non-smoker / 3 = NA)	0/5/1	0/8/0	0.363
% steatosis	13.5 ± 14.54	63.13 ± 16.46	0.0001
Steatosis grade (0/1/2/3)	2/3/1/0	0/0/4/4	0.0014
Ballooning (0/1/2/3)	6/0/0/0	0/6/2/0	0.0001
Lobular inflammation (0/1/2/3/4)	3/3/0/0/0	0/3/4/1/0	0.0029
Portal inflammation (0/1/2/3/4)	4/2/0/0/0	0/7/1/0/0	0.0113
Hepatic fibrosis (0/1/2/3/4)	5/1/0/0/0	0/3/5/0/0	0.0001
NAFLD activity score (/11)	1.5 ± 1.38	7 ± 1.51	0.000016
Alanine aminotransferase (ALT; U/L)	24.33 ± 8.29	44.88 ± 20.77	0.0300
Aspartate aminotransferase (AST; U/L)	19.67 ± 2.42	30.75 ± 7.27	0.0030
AST/ALT	0.86 ± 0.19	0.78 ± 0.28	0.5148
gamma-glutamyl transferase (GGT; U/L)	26 ± 7.07	43.13 ± 22.77	0.0773
Bilirubin total (umol/L)	6 ± 1.79	9.13 ± 3.44	0.0500
Bilirubin direct (umol/L)	1.83 ± 0.41	2.13 ± 0.99	0.4699
Alkaline phosphatase (ALP; U/L)	100 ± 20.81	74.38 ± 12.26	0.0290

Means ± STD were reported.

Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3