


The Genetic Architecture of Diet-Induced Hepatic Fibrosis in Mice

Simon T. Hui,¹ Zeyneb Kurt,² Iina Tuominen,³ Frode Norheim,¹ Richard C. Davis,¹ Calvin Pan,¹ Darwin L. Dirks,¹ Clara E. Magyar,⁴ Samuel W. French,⁴ Karthickeyan Chella Krishnan,¹ Simon Sabir,¹ Francisco Campos-Pérez,⁵ Nahum Méndez-Sánchez ,⁶ Luis Macías-Kauffer,⁷ Paola León-Mimila,⁷ Samuel Canizales-Quinteros,⁷ Xia Yang,² Simon W. Beaven,³ Adriana Huertas-Vazquez,¹ and Aldons J. Lusis¹

We report the genetic analysis of a “humanized” hyperlipidemic mouse model for progressive nonalcoholic steatohepatitis (NASH) and fibrosis. Mice carrying transgenes for human apolipoprotein E*3-Leiden and cholesteryl ester transfer protein and fed a “Western” diet were studied on the genetic backgrounds of over 100 inbred mouse strains. The mice developed hepatic inflammation and fibrosis that was highly dependent on genetic background, with vast differences in the degree of fibrosis. Histological analysis showed features characteristic of human NASH, including macrovesicular steatosis, hepatocellular ballooning, inflammatory foci, and pericellular collagen deposition. Time course experiments indicated that while hepatic triglyceride levels increased steadily on the diet, hepatic fibrosis occurred at about 12 weeks. We found that the genetic variation predisposing to NASH and fibrosis differs markedly from that predisposing to simple steatosis, consistent with a multistep model in which distinct genetic factors are involved. Moreover, genome-wide association identified distinct genetic loci contributing to steatosis and NASH. Finally, we used hepatic expression data from the mouse panel and from 68 bariatric surgery patients with normal liver, steatosis, or NASH to identify enriched biological pathways. **Conclusion:** The pathways showed substantial overlap between our mouse model and the human disease. (HEPATOLOGY 2018;68:2182-2196).

SEE EDITORIAL ON PAGE 2059

Nonalcoholic fatty liver disease (NAFLD) is a multistage disease comprised of a spectrum of abnormalities ranging from simple steatosis to nonalcoholic steatohepatitis (NASH),

characterized by inflammation and liver cell damage. With sustained liver injury, individuals with NASH are at risk of developing progressive fibrosis, cirrhosis, and hepatocellular carcinoma.^(1,2) While simple steatosis appears to be benign in most cases, NASH is associated with increased overall morbidity and mortality.⁽²⁾ The etiology of

Abbreviations: ALT, alanine aminotransferase; APOE*3, apolipoprotein E*3; CETP, cholesteryl ester transfer protein; DAVID, Database for Annotation, Visualization and Integrated Discovery; ECM, extracellular matrix; ELISA, enzyme-linked immunosorbent assay; eQTL, expression quantitative trait locus; Fuca1, alpha-L-fucosidase 1; FDR, false discovery rate; GWAS, genome-wide association study; HMDP, Hybrid Mouse Diversity Panel; HOMA-IR, homeostasis model assessment of insulin resistance; LD, linkage disequilibrium; LDL, low-density lipoprotein; MCP-1, monocyte chemoattractant protein 1; MSEA, marker set enrichment analysis; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; Phactr, phosphatase and actin regulator 2; RNA-seq, RNA-sequencing; SNP, single-nucleotide polymorphism; TG, triglycerides.

Received December 5, 2017; accepted May 12, 2018.

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.30113/supinfo.

Supported by the National Institutes of Health (HL28481 and HL30568, to A.J.L.; DK104363 and NS103088, to X.Y.), an American Heart Association postdoctoral fellowship (17POST33670739, to Z.K.) and an Iris Cantor-UCLA Women's Health Center/CTSI fellowship (UL1TR001881, to Z.K.), a postdoctoral fellowship from the Sigrid Juselius Foundation and the Orion-Pharmos Research Foundation (to I.T.), a Research Council of Norway postdoctoral fellowship (240405/F20, to F.N.), the Medica Sur Clinic & Foundation (to N.M.-S.), and CONACyT SALUD (182801, to S.C.-Q.).

© 2018 The Authors. HEPATOLOGY published by Wiley Periodicals, Inc. on behalf of American Association for the Study of Liver Diseases. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

View this article online at wileyonlinelibrary.com.

DOI 10.1002/hep.30113

Potential conflict of interest: Nothing to report.

this disease is complex, involving multiple genetic and environmental factors and substantial overlap with obesity.⁽³⁾ Human genome-wide association studies (GWAS) have identified several loci and candidate genes for NAFLD, but these account for a small proportion of disease heritability.⁽⁴⁾ Among them, patatin-like phospholipase domain-containing 3 (PNPLA3) and transmembrane 6 superfamily member 2 (TM6SF2) have been validated in multiple studies.⁽⁵⁾

When fed a high-fat, high-sucrose diet, mice typically develop steatosis but do not progress to NASH.⁽⁶⁾ We now report that mice with a “humanized” lipoprotein profile due to transgenic expression of human apolipoprotein (APO) E*3-Leiden and human cholesteryl ester transfer protein (CETP) exhibit the entire spectrum of NAFLD when fed a “Western” diet. To understand the pathways underlying NASH/fibrosis, we introduced the transgenes onto a panel of over 100 different inbred strains of mice known as the Hybrid Mouse Diversity Panel (HMDP). We observed that the degree of fibrosis across the strains varied more than 500-fold, indicating a strong genetic component. We examined a number of molecular phenotypes across the strains, including global gene expression in liver, cytokine levels in plasma, and a number of relevant plasma metabolites. A time course study revealed key features of NASH progression. Our results indicate that the genetic factors contributing to steatosis and to NASH are largely distinct, in accordance with a two-step model. Finally, using

genetic and RNA-sequencing (RNA-seq) analysis of human liver biopsies, we provide evidence that many of the same pathways are shared by our mouse model and by humans.

Materials and Methods

ANIMALS

Transgenic mice expressing human CETP were obtained from The Jackson Laboratory, and mice carrying the human APOE*3-Leiden variant were kindly provided by Dr. L. Havekes.^(7,8) These mice were interbred to create a strain carrying both transgenes in a C57BL/6J background and subsequently bred to females from about 100 common inbred strains (Supporting Table S1). Animals were maintained on a 12-hour light–dark cycle with *ad libitum* access to water. Male mice (8–10 weeks old) were fed a “Western” diet (33 kcal% fat from cocoa butter and 1% cholesterol; Research Diets; catalog no. D10042101) for 16 weeks. Body composition was measured by nuclear magnetic resonance (Bruker Biospin Corp., Billerica, MA). At the end of the study, mice were fasted for 4 hours beginning at 6:00 AM and sacrificed at 10:00 AM. To avoid sampling error, liver samples for histological and microarray analyses were taken from the same lobe in each mouse. All animal work was approved by the University of California, Los Angeles, Animal Research Committee, the institutional animal care and use committee.

ARTICLE INFORMATION:

From the ¹Department of Medicine, Division of Cardiology, David Geffen School of Medicine; ²Department of Integrative Biology and Physiology, University of California, Los Angeles, CA; ³Department of Medicine, Division of Digestive Diseases & Pflieger Liver Institute and Center for Obesity and Metabolic Health (COMET), David Geffen School of Medicine; ⁴Department of Pathology & Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA; ⁵Clínica Integral de Cirugía para la Obesidad y Enfermedades Metabólicas, Hospital General Dr. Rubén Lénero; ⁶Liver Research Unit, Medica Sur Clinic and Foundation; ⁷Facultad de Química, UNAM/Instituto Nacional de Medicina Genómica (INMEGEN), Unidad de Genómica de Poblaciones Aplicada a la Salud, Mexico City, Mexico.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO:

Aldons J. Lusis, Ph.D.
or
Simon T. Hui, Ph.D.
Department of Medicine/Division of Cardiology
University of California, Los Angeles

675 Charles E. Young Dr. S
Los Angeles, CA 90095
E-mail: jlusis@mednet.ucla.edu, sthui@mednet.ucla.edu
Tel: +1-310-825-1359, +1-310-825-2957

QUANTITATIVE ASSESSMENT OF STEATOSIS AND FIBROSIS IN THE LIVER

Liver lipids were quantitated by methods according to Hui et al.⁽⁶⁾ For histological examination, livers were fixed in 10% formalin, embedded in paraffin, sectioned at 5 μm , and stained with picrosirius red. Slides were scanned at $\times 20$ magnification (Aperio ScanScope XT; Leica Biosystems). Liver fibrosis was measured as a quantitative trait from liver sections using picrosirius red staining followed by a whole-slide digital imaging algorithm to quantify the fibrosis in each section as a percentage of the total liver area. The trained algorithm can correctly identify pathologic fibrosis and differentiate it from collagen staining of the vascular wall and liver capsule, which is excluded from the quantitation (Supporting Information and Fig. S1). This method was validated by showing that the fibrosed area correlated with both hepatic hydroxyproline content ($r = 0.54$, $P = 6.1 \times 10^{-7}$) and blinded fibrosis score assessed by a liver pathologist ($r = 0.80$, $P = 8.4 \times 10^{-18}$) (Supporting Fig. S2).

PLASMA LIPIDS, ENZYMES, CYTOKINES, AND METABOLITES

Plasma lipid profiles and alanine aminotransferase (ALT) activity were measured by colorimetric analysis as described.⁽⁶⁾ Quantification of plasma cytokines was carried out in a custom multiplexed cytokine enzyme-linked immunosorbent assay (ELISA) kit using an electrochemiluminescent detection method per the manufacturer's instructions (Meso Scale Discovery, Rockville, MD). Plasma insulin was measured using the mouse insulin ELISA kit (80-INSMS-E01) from Alpco (Salem, NH).⁽⁸⁾ Plasma metabolites in a subset of 124 mice from 52 strains (Supporting Table S2) were quantified using time-of-flight tandem mass spectrometry as described.⁽⁹⁾

RNA ISOLATION AND GENE EXPRESSION ANALYSIS

Liver RNA was isolated from a subset of 249 livers from 102 strains (Supporting Table S3) and analyzed for global gene expression using Affymetrix HT-MG_430 PM Microarrays as described.⁽¹⁰⁾ Expression quantitative trait loci (eQTL) in liver were

mapped using the FaST-LMM algorithm with correction for population structure.⁽¹¹⁾ Correlations were calculated using the bicor function from the WGCNA R package.⁽¹²⁾ Gene Ontology (GO) enrichment analysis was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID, v6.7) program.⁽¹³⁾

GENOME-WIDE ASSOCIATION ANALYSIS

High-density genotypes for inbred strains of mice were generated by the Mouse Diversity Array.⁽¹⁴⁾ Genome-wide association mapping of hepatic fibrosis was performed using FaST-LMM.⁽⁸⁾ Deleterious missense mutations were identified using the PROVEAN prediction algorithm.⁽¹⁵⁾

RNA-SEQ OF HUMAN NAFLD LIVER BIOPSIES

A total of 68 male and female nonrelated Mexican mestizo subjects who underwent bariatric surgery for morbid obesity with a body mass index (BMI) $\geq 40 \text{ kg/m}^2$ were included in the study. The mean age was 39.1 ± 9.4 years. Detailed characteristics of the study have been reported.⁽¹⁶⁾ Anthropometric and biochemical parameters of the study population can be found in Supporting Table S4. Liver biopsies were collected in RNAlater (Sigma, St. Louis, MO) and stored at -80°C . Genotyping was performed using the Multi-Ethnic Genotyping Array (Illumina, San Diego, CA). RNA quality was assessed using the Bioanalyzer RNA chip analysis to ensure that the RNA integrity number was >7 . Complementary DNA libraries were prepared using the TruSeq RNA Stranded Total RNA Library Preparation kit (Illumina) and sequenced using an Illumina HiSeq2500 instrument, generating approximately 50 million reads/sample. After data quality control, sequencing reads were mapped to the human reference genome using TopHat and quantified using Cufflinks. This study was approved by the institutional review boards of the National Institute of Genomic Medicine (Mexico) and the Hospital Dr. General Ruben Leñero. All participants provided written informed consent prior to their inclusion. The study was performed according to the principles of the Declaration of Helsinki.

COMPARISON OF BIOLOGICAL PATHWAYS ASSOCIATED WITH NASH BETWEEN MICE AND HUMANS

Marker set enrichment analysis (MSEA) was employed to identify causal biological pathways of NASH/fibrosis in our mouse model and a human bariatric surgery cohort.^(17,18) Genes in each canonical pathway or coexpression module were mapped to single-nucleotide polymorphisms (SNPs) using liver eQTLs and then integrated with GWAS of mouse liver fibrosis or human NASH to capture strong, moderate, and subtle genetic signals. Genotyping of the obese patients was performed using the Multi-Ethnic Genotyping Array. A total of 507,122 SNPs remained after removal of SNPs with a minor allele frequency <5% and Hardy-Weinberg equilibrium $P < 1 \times 10^{-5}$. Additional SNP imputation with Beagle resulted in 865,847 SNPs analyzed.⁽¹⁹⁾ Genotypic associations for NASH were estimated in a mixed linear model with, sex, age, and BMI as fixed effects and the genetic relatedness matrix as a random effect. Liver tissue eQTLs in humans were calculated using the FaST-LMM algorithm.⁽¹¹⁾ Additional human liver tissue eQTLs collected from multiple resources including Massachusetts General Hospital⁽²⁰⁾ and the Genotype-Tissue Expression project⁽²¹⁾ were pooled together and used in the MSEA. Cis-eQTLs (or “local” eQTLs) were defined as those within 1 Mb of the gene. In total 315,652 liver tissue cis-eQTL associations (false discovery rate [FDR] < 0.05) were included in the analysis. Pathways were considered to be shared between mice and humans if >50% of genes were shared. Associations between SNPs affecting expression of genes in a pathway were tested for association with NASH/fibrosis, and the set of P values was compared with random genes. An FDR < 0.05 cutoff was used.

Additional details about the methodology are provided in Supporting Information.

Results

A “HUMANIZED” HYPERLIPIDEMIA MOUSE MODEL FOR PROGRESSIVE NASH AND FIBROSIS

Mice transgenic for APOE*3-Leiden and CETP and fed a high-fat, high-cholesterol diet are a model

of metabolic syndrome and atherosclerosis.^(22,23) During a study of atherosclerosis,⁽⁸⁾ we observed that transgenic mice on a pure C57BL/6J background develop features characteristic of steatosis, NASH, and fibrosis. Unlike mice fed with the commonly used methionine/choline-deficient diet to induce fibrosis, the CETP and APOE*3-Leiden transgenic mice fed a Western diet containing 1% cholesterol developed insulin resistance along with a steady increase in body weight, adiposity, and insulin resistance (Supporting Fig. S3A-C). Liver triglycerides (TG) accumulated progressively, reaching a plateau of ~6-fold at 12 weeks (Fig. 1,3A). Hepatic cholesterol levels began increasing within 2 weeks of the beginning of diet feeding (Fig. 1B), while unesterified cholesterol did not show a sharp increase until week 8 (Fig. 1C). Hepatic phospholipid levels were not altered (data not shown). Histological analysis showed that steatosis and hepatocellular ballooning were apparent at 8 weeks (Fig. 1D). Subsequently, advanced steatosis developed, as evidenced by the presence of macrovesicular steatosis (Fig. 1D). Collagen deposition was observed around week 12, and overt hepatic fibrosis developed between weeks 12 and 16 (Fig. 1E). The severity of histological changes paralleled the increasing levels of proinflammatory cytokines (including tumor necrosis factor alpha [TNF- α], interleukin 6 [IL-6], and monocyte chemoattractant protein 1 [MCP-1]) in the plasma (Fig. 2A-C), a characteristic of NASH patients. Furthermore, these changes were accompanied by progressive increases in the expression of gene markers for immune cell infiltration (F4/80, MCP-1; Fig. 2D,E) and fibrogenesis (collagen type 1 alpha 1 [*Col1a1*]; Fig. 2F).

INDEPENDENT GENETIC REGULATION OF HEPATIC STEATOSIS AND FIBROSIS

To assess the impact of genetic variation on NASH susceptibility, each HMDP strain was crossed to a C57BL/6J male mouse hemizygous for human APOE*3-Leiden and homozygous for CETP transgenes. F1 progeny hemizygous for both APOE*3-Leiden and CETP transgenes were selected and fed a “Western” high-fat, high-cholesterol diet for 16 weeks.⁽⁸⁾ Overall, 619 male mice from 102 strains were surveyed (Supporting Table S1). A wide spectrum of fibrosis, predominantly pericellular fibrosis,

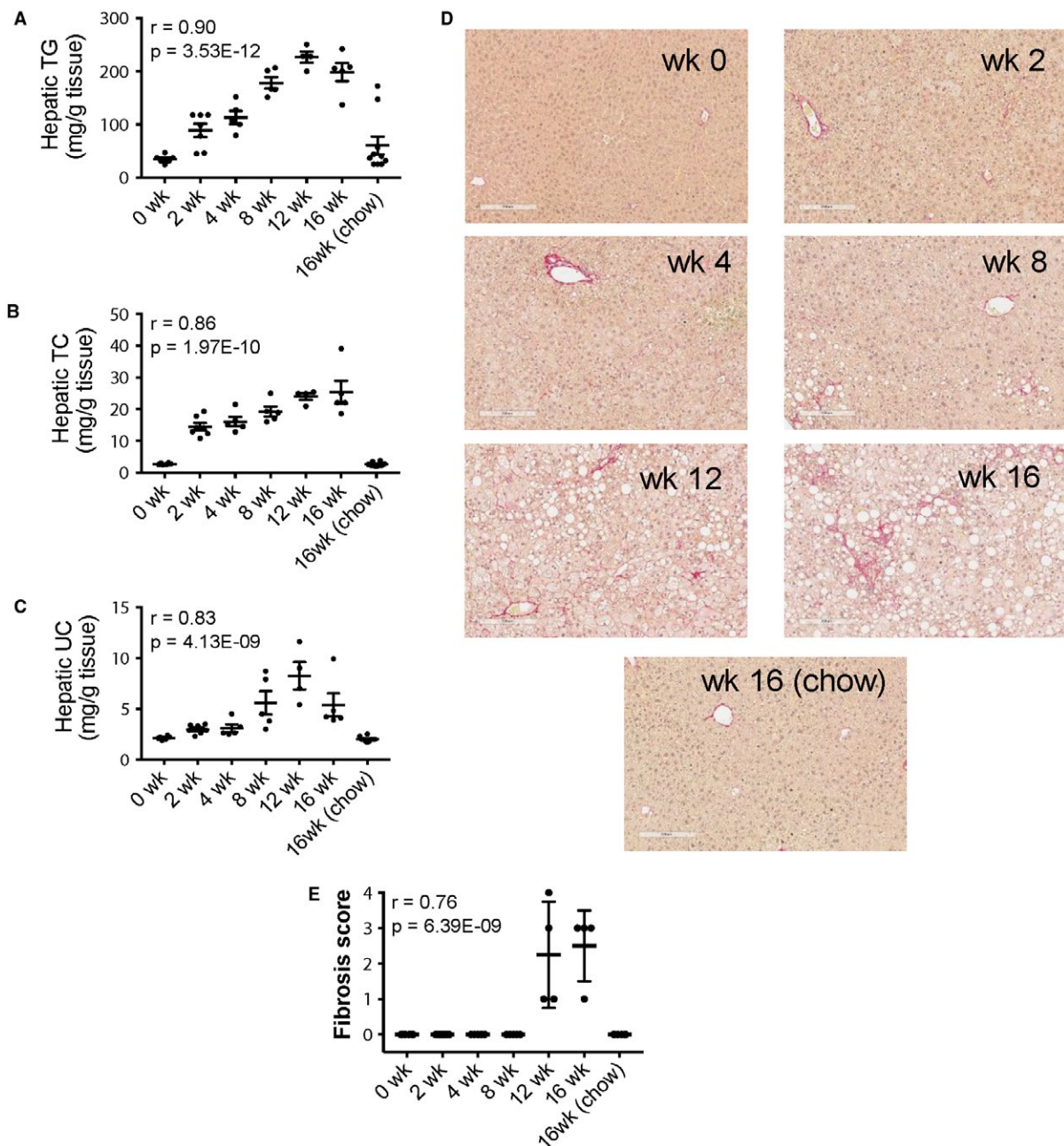


FIG. 1. Change in hepatic histology and lipid levels during progression of NASH. Male C57BL/6J mice carrying human APOE*3-Leiden and CETP transgenes were fed a “Western” high-fat, high-cholesterol diet for 0-16 weeks. Hepatic levels of TG (A), total cholesterol (B), and unesterified cholesterol (C) are presented as mean \pm SEM from 4-10 mice per group. Control mice were fed 16 weeks on chow. (D) Liver sections were stained for collagen with picosirius red at the indicated time after diet, and the fibrosis score was determined by a pathologist blinded to the study (E). Abbreviations: TC, total cholesterol; UC, unesterified cholesterol.

was observed among the strains (Fig. 3A). Female mice generally showed a similar strain distribution for fibrosis, although there was evidence of sexual dimorphism as well (Supporting Fig. S4). While some strains are resistant to fibrosis (Fig. 3B), livers from

susceptible strains displayed signature histological features of human NASH livers, including macrovascular steatosis, inflammatory foci, hepatocellular ballooning, and a pericellular (“chicken wire”) pattern of collagen deposition (Fig. 3C,D). The degree of fibrosis was

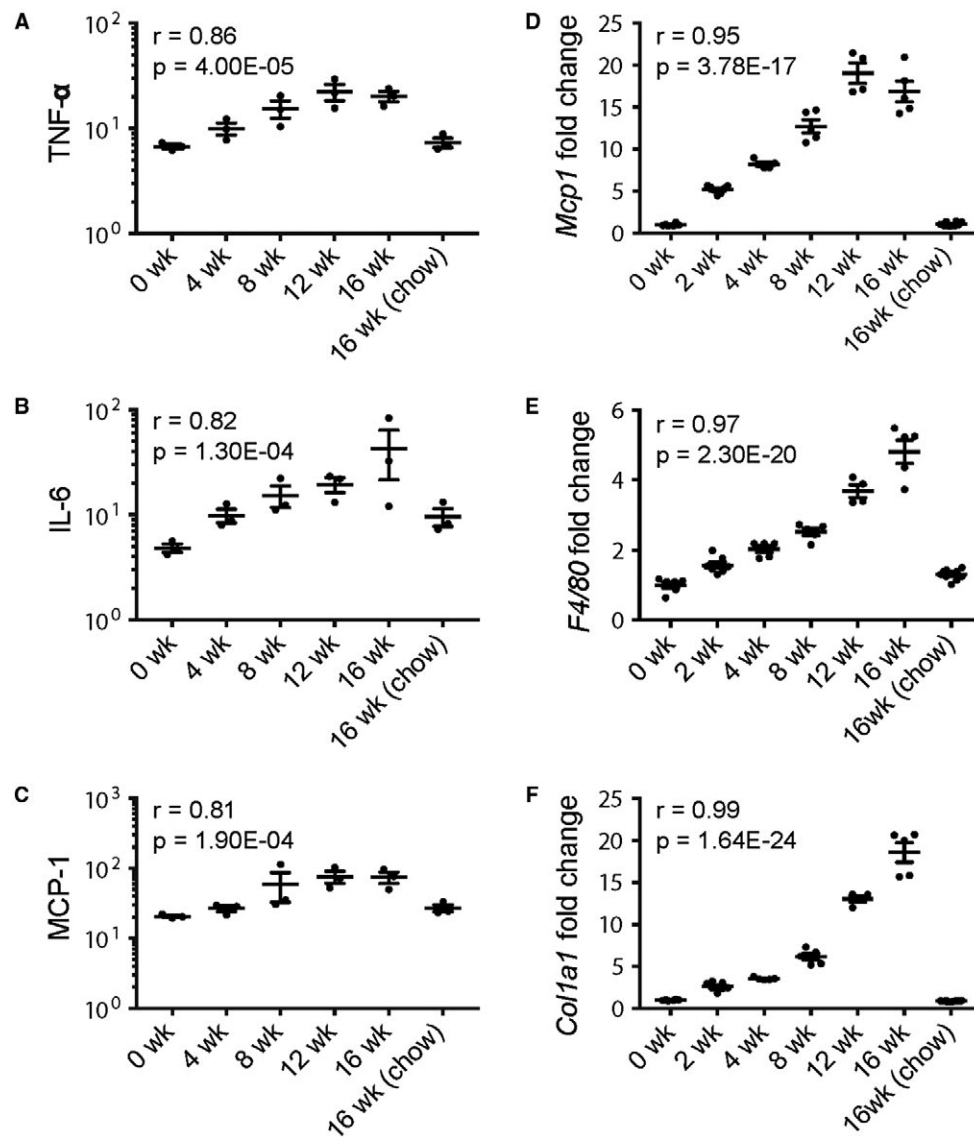


FIG. 2. Change in plasma cytokines and hepatic gene expression during progression of NASH. Male C57BL/6J mice carrying human APOE*3-Leiden and CETP transgenes were fed a “Western” diet for 0–16 weeks. Control mice were fed for 16 weeks on chow. Plasma levels (picograms per milliliter) of TNF- α (A), IL-6 (B), and MCP-1 (C) were measured by ELISA. Data are mean \pm SEM from three mice per group. Liver expression of F4/80 (D), MCP-1 (E), and *Col1a1* (F) was measured by quantitative PCR and normalized to ribosomal protein L4. Fold change was compared to the 0 week diet group. Data are presented as mean \pm SEM from 4–10 mice per group.

positively correlated with liver injury, as shown by significant correlation between hepatic fibrosis and ALT activity in the plasma (Fig. 3E).

We previously studied the HMDP inbred mouse strains on a high-fat, high-sucrose diet, which results in dramatic steatosis but little or no fibrosis.⁽⁶⁾ We examined whether the same genetic background

would confer susceptibility equally to steatosis and fibrosis. No significant correlation was observed between the strain susceptibility to steatosis in the previous cohort and that to NASH/fibrosis in the present study (Fig. 3F), indicating that these two processes were regulated largely by independent genetic determinants.

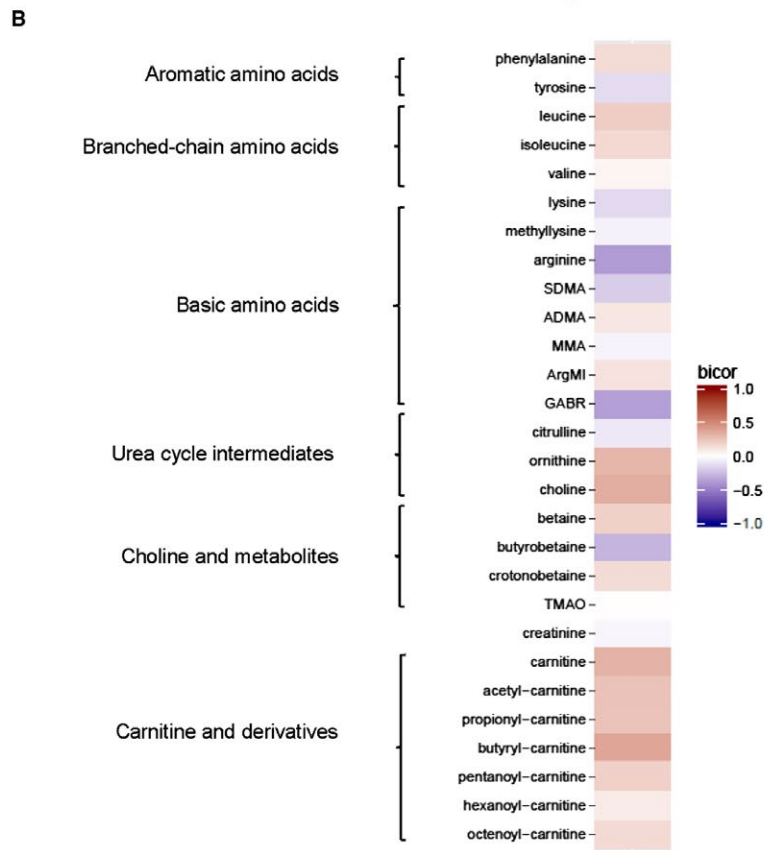
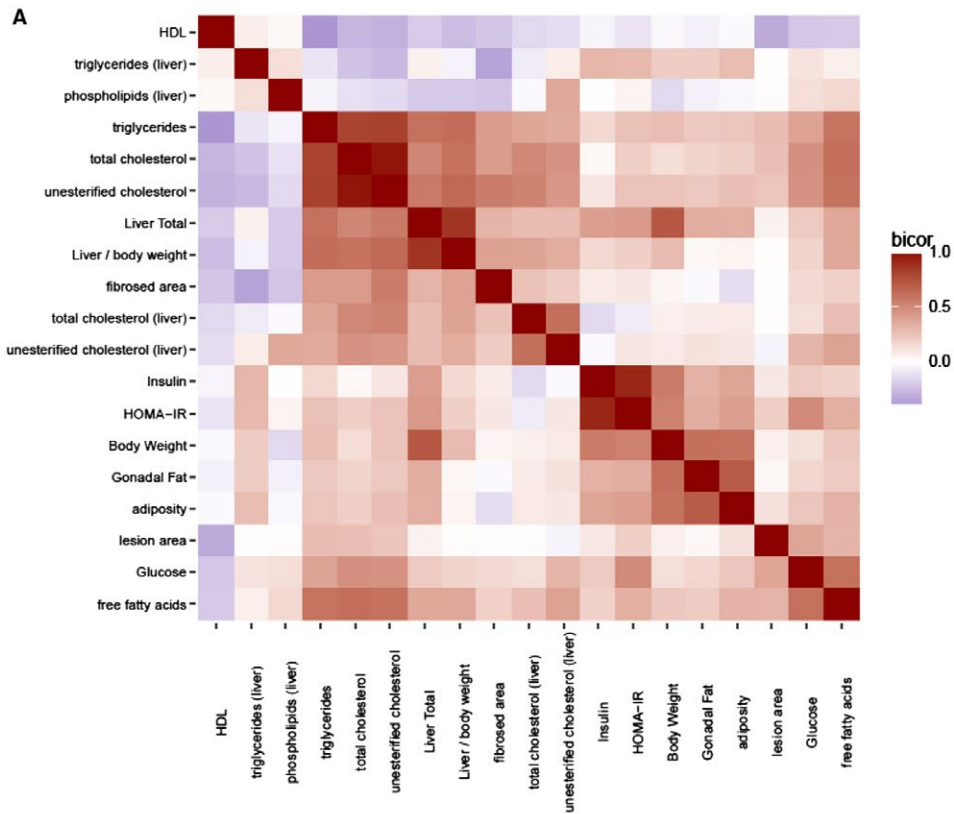


FIG. 4. Correlation of fibrosis with clinical traits and plasma metabolites. Results are represented as heatmaps. Correlations of liver fibrosis and TG with clinical traits and atherosclerotic lesion area are shown in (A), and correlations between fibrosis and plasma metabolites are shown in (B). Abbreviations: ADMA, asymmetric dimethylarginine; bicor, biweight midcorrelation; GABR, global arginine bioavailability ratio; HDL, high-density lipoprotein cholesterol; MMA, N-mono-methylarginine; SDMA, symmetric dimethylarginine; TMAO, trimethylamine N-oxide.

($r = -0.35$) and phospholipid ($r = -0.22$) levels were negatively correlated with fibrosis at later stages, the latter perhaps reflecting larger lipid droplets as macrovesicular steatosis. As in human NASH, hepatic fibrosis was associated with hyperlipidemia, including significant correlations with plasma TG ($r = 0.42$), free fatty acids ($r = 0.20$), total cholesterol ($r = 0.42$), unesterified cholesterol ($r = 0.55$), and very low-density lipoprotein + low-density lipoprotein (LDL) cholesterol ($r = 0.36$) (Fig. 4A; Supporting Table S5). While the homeostasis model assessment of insulin resistance (HOMA-IR) was positively correlated with body weight, fat mass, and insulin levels, it showed only a modest correlation with hepatic fibrosis, in contrast to the strong correlation with steatosis in mice fed a high-fat, high-sucrose diet.⁽⁶⁾ We did not observe significant correlations between fibrosis and body weight, gonadal fat mass, or atherosclerotic lesions (Fig. 4A).

ALTERED AMINO ACID AND FATTY ACID METABOLISM IN HEPATIC FIBROSIS

Previous metabolomics studies have identified amino acid abnormalities associated with NAFLD in humans.⁽²⁵⁾ Therefore, we performed targeted metabolomic analysis of polar metabolites in the plasma of a subset of our mice (Fig 4B; Supporting Table S6). Among the metabolites measured, plasma arginine exhibited the strongest correlation with hepatic fibrosis ($r = -0.38$). On the other hand, plasma levels of choline ($r = 0.34$), betaine ($r = 0.19$), and short chain acyl-carnitines were positively correlated with hepatic fibrosis.

GLOBAL TRANSCRIPTOMIC PROFILING REVEALS DISTINCT PATHWAYS UNDERLYING STEATOSIS AND FIBROSIS

To gain insights into the fibrotic process, we quantitated global gene expression in the livers across all

102 strains by microarrays (Supporting Table S3). The top 1,000 genes correlated with hepatic fibrosis ($P < 1.3 \times 10^{-9}$) were analyzed using DAVID to test for enrichment of GO categories (Fig. 5A). We found 10 enriched pathways which are broadly related to (1) collagen metabolism (collagen fibril organization, extracellular matrix [ECM] structural constituent, ECM-receptor interaction, focal adhesion, and protein digestion and absorption), (2) stellate cell activation (platelet-derived growth factor binding), (3) lipid metabolism (fatty acid metabolic process, oxidation-reduction process), (4) flavonoid metabolism, and (5) insulin signaling (phosphoinositide 3-kinase-Akt signaling pathway). Most of these pathways are consistent with pathways known to be important in maintaining hepatic lipid balance and response to hepatic injury.

We also examined how individual gene expression varied with fibrosis-related clinical traits (LDL cholesterol, plasma arginine, hepatic TG, adiposity, and HOMA-IR) (Fig. 5B). Unsupervised hierarchical clustering showed that there was parallel expression between many of the genes correlated with hepatic fibrosis and LDL cholesterol (Fig. 5B, clusters 1-3). In contrast, the expression pattern of genes associated with plasma arginine and, to a lesser extent, hepatic TG showed a reciprocal pattern compared to fibrosis (Fig. 5B, clusters 1-3). Genes correlated with adiposity and HOMA-IR showed only weak correlation with hepatic fibrosis. This is in sharp contrast to the gene expression pattern in the high-fat/high-sucrose diet-induced steatosis model,⁽⁶⁾ in which genes associated with adiposity and HOMA-IR paralleled those associated with hepatic steatosis (Fig. 5C, clusters 4-6). The distinct gene expression signatures associated with hepatic steatosis and fibrosis suggest that the genetic regulation of these two stages of NAFLD is different.

IDENTIFICATION OF FIBROSIS CANDIDATE GENES BY GWAS

To identify genetic loci contributing to hepatic fibrosis, we performed genome-wide association

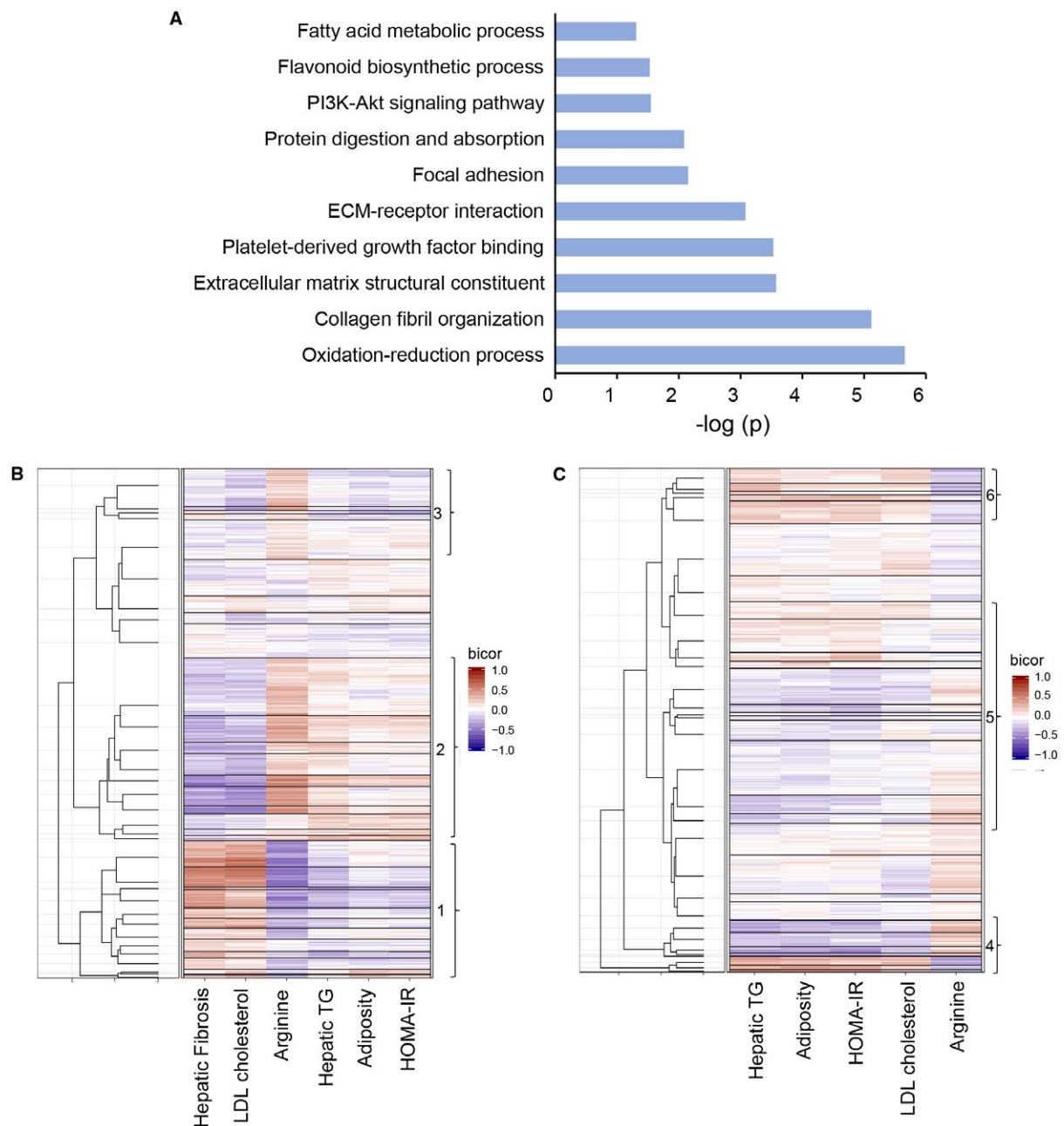


FIG. 5. Transcriptomic analysis of liver gene expression. (A) Enrichment pathways from top 1,000 genes correlated with hepatic fibrosis identified by DAVID. Benjamini-corrected P values are shown as $-\log(p)$ on the x -axis. (B,C) Full-length view of the cluster diagram with individual gene probes along the vertical axis and selected clinical traits along the horizontal axis. Positive correlations between transcript level and clinical trait are indicated in red and negative correlations in blue. Results from CETP and APOE*3-Leiden mice on a high-fat, high-cholesterol diet for 16 weeks are shown in (B), and results from HMDP mice fed a high-fat, high-sucrose diet for 8 weeks are shown in (C). Abbreviations: bicor, biweight midcorrelation; PI3K, phosphoinositide 3-kinase.

analyses. The threshold for genome-wide significance was based on simulation and permutation, as described.⁽²⁶⁾ Two genome-wide significant loci for hepatic fibrosis were identified in male mice (Fig. 6A)

on chromosomes 10 and 15. The peak SNP (rs50309490, $P = 1.2 \times 10^{-6}$) for the chromosome 10 locus falls within a linkage disequilibrium (LD) block ~2.6 Mb containing 16 genes (Fig. 6B). To

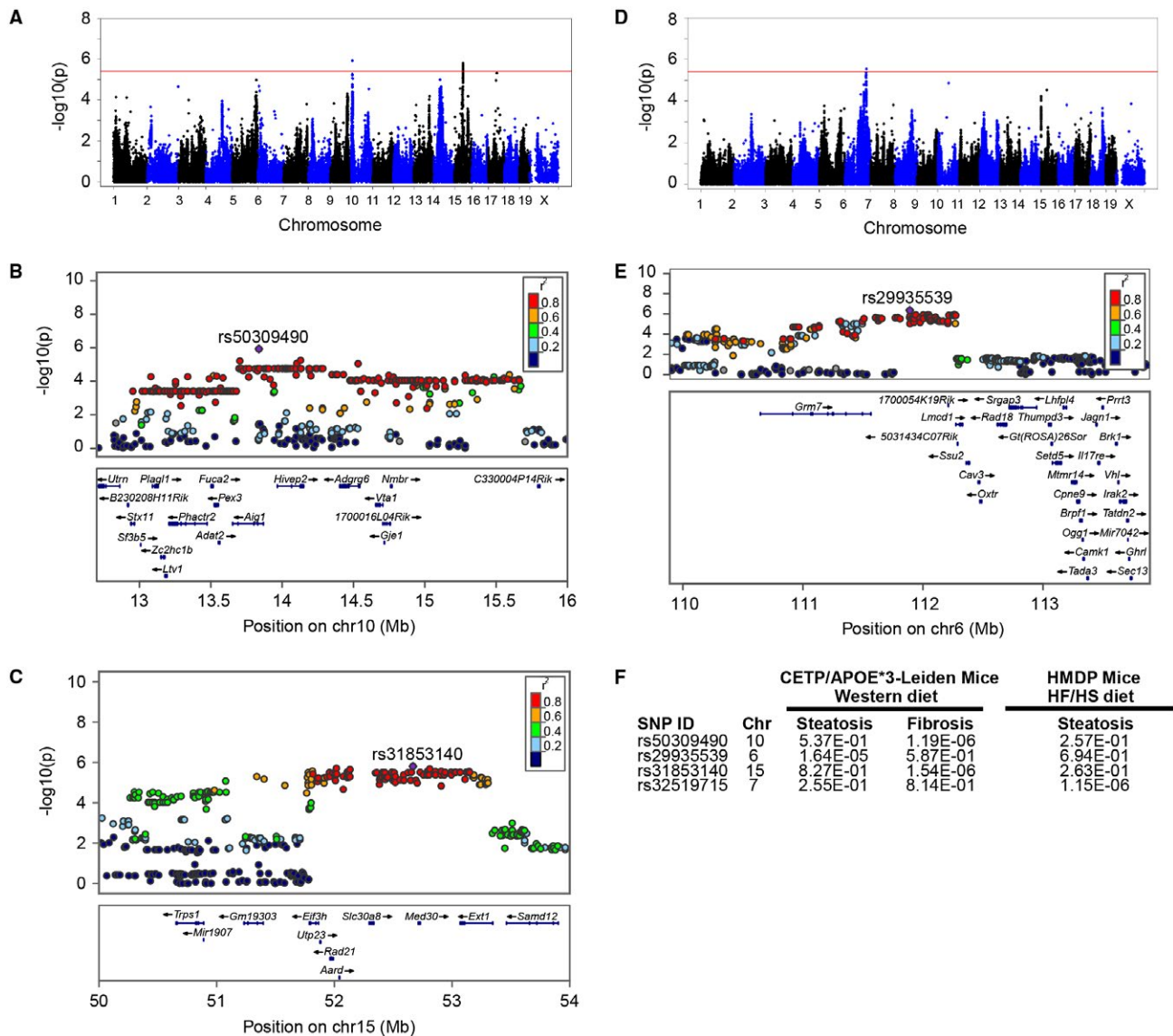


FIG. 6. Genome-wide association mapping of hepatic fibrosis and steatosis. (A) Manhattan plot showing the association of SNP genotype with hepatic fibrosis (A) and hepatic TG (D). Red line represents the threshold for genome-wide significance. LocusZoom plots are shown for genome-wide significant loci on chromosome 10 (B), chromosome 15 (C), and chromosome 6 (E). Association *P* values for steatosis and hepatic fibrosis for the peak SNPs in genome-wide significant loci for hepatic fibrosis or steatosis are shown in (F). Abbreviation: HF/HS, high-fat, high-sucrose.

prioritize candidate genes in this locus, we identified genes whose expression was controlled locally or that exhibited missense mutations leading to predicted deleterious effects. Significant local or *cis*-eQTLs are present in alpha-L-fucosidase 1 (*Fuca1*; rs29329055, $P = 4.9 \times 10^{-6}$) and phosphatase and actin regulator 2 (*Phactr2*; rs50482796, $P = 1.0 \times 10^{-11}$). In addition, *Phactr2* has a potential deleterious missense mutation (S217Y) (Supporting Table S7). The peak SNP of the

chromosome 15 locus (rs31853140, $P = 1.5 \times 10^{-6}$) is within an LD block of ~1.6 Mb, containing seven candidate genes (Fig. 6C). Eukaryotic translation initiation factor 3 subunit H (*Eif3h*) is the only gene whose expression correlated with liver fibrosis and exhibits a *cis*-eQTL (rs38654821, $P = 1.3 \times 10^{-6}$). Association mapping for hepatic TG yielded a completely different associated pattern compared to fibrosis (Fig. 6D). The peak SNP (rs29935539, $P = 4.4 \times 10^{-7}$) lies within an

LD block of ~1 Mb on chromosome 6 and contains only four genes (Fig. 6E). Previously, we identified a locus on chromosome 7 for steatosis using HMDP strains on a high-fat/high-sucrose diet.⁽⁶⁾ The distribution of steatosis susceptibility across the HMDP in that model was only modestly correlated with the fibrosis reported here (Fig. 3F), and the GWAS loci for steatosis and fibrosis differed (Fig. 6F). These data suggest that the genetic factors affecting susceptibility to the two traits differ.

MOUSE AND HUMAN NASH/FIBROSIS PATHWAYS EXHIBIT SUBSTANTIAL OVERLAP

To evaluate the relevance of our mouse model to that of human pathophysiology, we examined the transcriptomic context of our mouse model in relation to human NAFLD. Significant correlations were found between liver fibrosis and genes implicated in human GWAS for NAFLD, suggesting that the hepatic transcriptomic context in our mice resembles that in human NASH (Table 1). We further compared our mouse results to data from 68 bariatric surgery patients exhibiting a range of NAFLD phenotypes based on histological assessment using liver biopsies. The human liver biopsies were also subjected to RNA-seq analysis to allow modeling biologic networks. We identified pathways associated with NASH/fibrosis in both species using the MSEA pipeline as described under Materials and Methods. Briefly, this method tests for overrepresentation of disease-related GWAS signals among the eQTLs mapped to individual pathways compared to

random sets of genes.⁽¹⁷⁾ At an FDR of <5%, 60% of the mouse fibrosis causal mechanisms were identified in causal human NASH subnetworks (Fig. 7). Nine pathways were found to be shared between mouse and human livers: innate and adaptive immune system, cell cycle, notch signaling, TGF β signaling, Wnt signaling, signal transduction, neurotrophin signaling, peroxisome proliferator-activated receptor signaling and peroxisome metabolism, and viral human immunodeficiency virus infection. Six gene sets were mouse-specific (fatty acid, triacylglycerol, and ketone body metabolism, mitogen-activated protein kinase signaling, mRNA processing, and transcription pathways), whereas 12 gene sets were human-specific (cytokine signaling, chemokine receptors, ECM and matrisome terms, hedgehog and Janus kinase-signal transducer and activator of transcription signaling, apoptosis, adherens junctions, circadian clock terms) (Fig. 7).

Discussion

We report a systems genetics analysis of NASH/fibrosis using hyperlipidemic human CETP and APOE*3-Leiden transgenic mice that develop many features and molecular signatures characteristic of human NASH pathophysiology. Several conclusions have emerged. First, we demonstrate that hepatic fibrosis in mice is strongly dependent on genetic background, consistent with heritability estimates in humans.⁽³⁾ Second, our studies indicate that hepatic steatosis and NASH/fibrosis are mediated by distinct genetic factors, consistent with a multistep model of NAFLD. Third, our multi-omics analysis

TABLE 1. Correlation Between Expression of Human GWAS Candidate Genes for NAFLD/NASH and Liver Fibrosis in Mice

Gene	Name	<i>r</i>	<i>P</i>
<i>Ppp1r3b</i>	Protein phosphatase 1, regulatory (inhibitor) subunit 3B	-0.251	0.011
<i>Tm6sf2</i>	Transmembrane 6 superfamily member 2	-0.087	0.384
<i>Gckr</i>	Glucokinase regulatory protein	-0.220	0.026
<i>Col13a1</i>	Collagen, type XIII, alpha 1	-0.359	2.0E-04
<i>Pnpla3</i>	Patatin-like phospholipase domain containing 3	-0.177	0.076
<i>Pzp</i>	Pregnancy zone protein	0.112	0.261
<i>Mboat7</i>	Membrane bound O-acyltransferase domain containing 7	-0.405	2.5E-05
<i>Lyplal1</i>	Lysophospholipase-like 1	-0.233	0.018
<i>Fdft1</i>	Farnesyl diphosphate farnesyl transferase 1	-0.238	0.016
<i>Trib1</i>	Tribbles homolog 1 (<i>Drosophila</i>)	-0.208	0.036
<i>Ncan</i>	Neurocan	-0.251	0.011

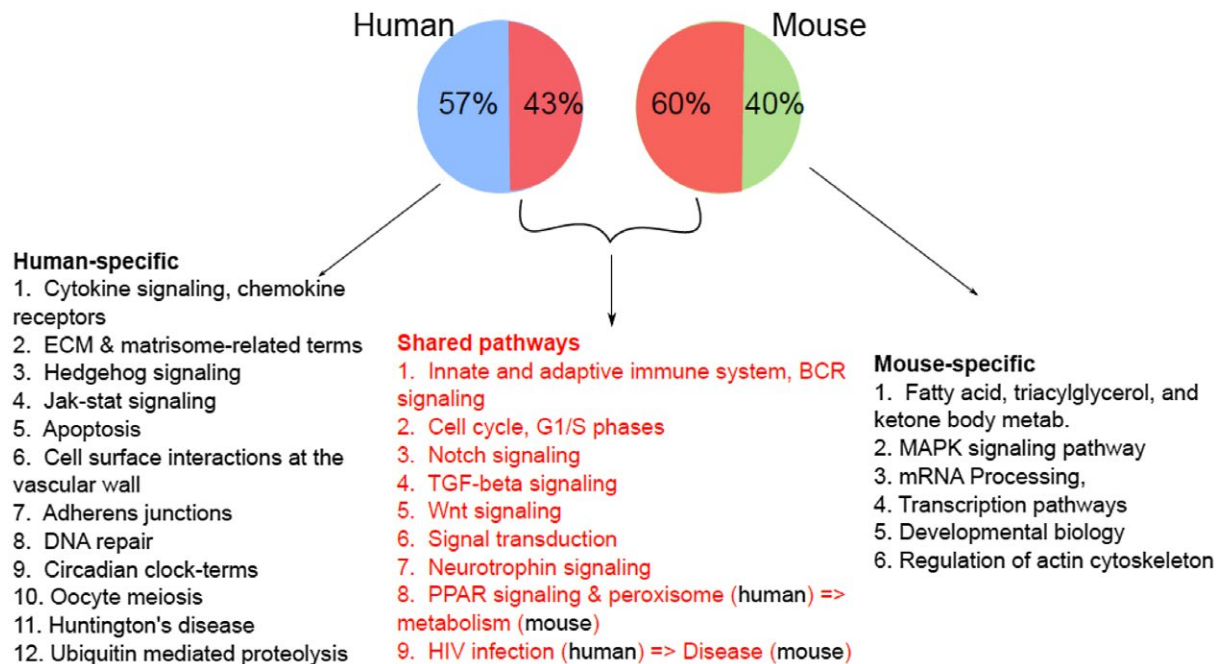


FIG. 7. Comparison between causal pathways associated with human NASH and mouse liver fibrosis. Pathways were identified using the MSEA algorithm.^(17,18) Abbreviations: BCR, B-cell receptor; HIV, human immunodeficiency virus; MAPK, mitogen-activated protein kinase; PPAR, peroxisome proliferator-activated receptor; TGF, transforming growth factor.

of transcripts and metabolites has identified pathways and markers associated with NASH/fibrosis. Finally, we have performed a preliminary comparison of pathways underlying NASH and fibrosis in our mouse model compared to humans, observing very substantial overlap.

Due to the transgenic expression of CETP, these mice exhibit a high-density lipoprotein:LDL lipoprotein profile resembling that of humans, although both apoB48 and apoB100 are present in LDL, whereas human LDL only contains ApoB100. The contribution of apoB48-containing lipoprotein to NASH remains to be studied. Nevertheless, our results indicate that hyperlipidemic transgenic mice expressing human CETP and APOE*3-Leiden provide a useful model for NASH research. First, unlike the commonly used methionine/choline-deficient diet model, which has a drawback of inducing rapid weight loss and decreasing adiposity without insulin resistance, our model showed weight gain and insulin resistance. Second, histological analysis showed that our mice developed the pathological characteristics of human NASH: macrovascular steatosis, hepatocyte ballooning, inflammatory foci, and pericellular patterns of

collagen deposition. Third, at the molecular level, markers for inflammation, apoptosis, and fibrosis were strongly correlated with fibrosis in our mice. Fourth, transcriptomic analysis showed that there was significant overlap between pathways associated with fibrosis in both mice and a cohort of bariatric surgery patients exhibiting NASH. During the course of these studies, others reported a similar finding with APOE*3-Leiden, CETP transgenic mice on a C57BL/6J background.^(22,27)

While we are underpowered to detect interstrain differences due to low replicate numbers, association and correlation analyses were done with the entire population of mice consisting of >100 strains and were not dependent on strain differences, allowing sufficient statistical power to detect difference in genetic pathways. Association mapping was done after correction of population structure, and genome-wide significant loci were identified. Association mapping using data from 102 strains of F1 HMDP transgenic mice identified two loci for liver fibrosis, and we prioritized candidate genes in the loci using transcriptomic and protein coding data. At the chromosome 10 locus, we identified two genes (*Phactr2* and *Fuca1*) as

strong candidates. *Phactr2* belongs to member of the largely uncharacterized PHACTR family of protein phosphatase 1-binding and actin-binding proteins involved in inflammation and cell cycle control.^(28,29)

Fuca1 encodes the lysosomal hydrolase α -L-fucosidase implicated in cell growth and signaling.⁽³⁰⁾ At the chromosome 15 locus, *Eif3h*, encoding the eukaryotic translation initiation factor, is a candidate that is involved in the cell cycle.⁽³¹⁾ One limitation of our study is that all F1 mice carry one chromosome from C57BL/6J and one chromosome from the HMDP. Thus, recessive risk alleles not present in the C57BL/6J genome will not be detected and mapped under our mouse model. This is unlikely to be a major issue, however, because the vast majority of common variations exhibit additive inheritance.⁽³²⁾

To help identify markers for NAFLD and to better understand the underlying pathways, we performed metabolomic and transcriptomic analyses. Arginine was strongly negatively correlated with fibrosis, while choline, betaine, and acyl-carnitines were significantly positively correlated. Arginine serves as a precursor not only in protein synthesis and the urea cycle but also in the generation of polyamines, creatine, and nitric oxide. Treatment with arginine appears to block oxidative stress-induced NASH by inhibiting cytochrome P450 2E1 activity, decreasing the TNF- α level and restoring activities of endothelial nitric oxide synthase and antioxidant enzymes as well as glutathione levels.⁽³³⁾ Previous studies have shown that treatment with betaine improved nonalcoholic fatty liver and associated hepatic insulin resistance in mice.⁽³⁴⁾ Short-chain (C2, C3, C4, and C5) acyl derivatives of carnitine were positively correlated with fibrosis, whereas the precursor of carnitine, butyrobetaine, was inversely correlated. Elevated levels of acyl-carnitine are indicative of incomplete oxidation of fatty acids in the mitochondria, consistent with the notion that impaired fatty acid oxidation contributes to lipid accumulation and reactive oxygen species generation in the liver. In addition, elevated short-chain acylcarnitine levels may be contributed by altered branched-chain amino acid metabolism in the skeletal muscle.⁽³⁵⁾

We compared our global mouse transcriptomic data with those from livers from a cohort of bariatric surgery patients selected for a broad range of NAFLD phenotypes, enabling testing for enrichment of canonical pathways and comparison of pathways associated

with NAFLD phenotypes in mice and humans. The results were consistent with the GWAS findings indicating independent pathways contributing to steatosis and to NASH/fibrosis. While the mouse-human comparison must be considered preliminary, given the relatively small number of individuals examined, the pathway overlap was highly significant (Fig. 7). Supporting this conclusion, the hepatic expression of several human GWAS candidate genes, including PNPLA3, was correlated with fibrosis in our mouse model (Table 1). These results are in contrast to a recent study that compared gene expression in human NAFLD livers with nine mouse steatosis models.⁽³⁶⁾ The failure to observe significant overlap in that study may have been due in part to the fact that the mouse samples were obtained from a single time point without histological staging, whereas our study compared a range of NASH/fibrosis phenotypes in both the mouse and human samples.

Our comprehensive multi-omic data should serve as a rich resource for future studies of NASH and fibrosis. Primary data are available from the authors.

Acknowledgment: We thank Zhiqiang Zhou, Hannah Qi, Judy Wu, Yonghong Meng, and Sarada Charugundla for outstanding technical assistance and Dr. Rita Cantor for advice in statistical analysis. Mice carrying the human ApoE3 Leiden variant were kindly provided by Dr. L. Havekes (Leiden University Medical Center, Leiden, The Netherlands).

REFERENCES

- 1) Musso G, Gambino R, Cassader M, Pagano G. Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. *Ann Med* 2011;43:617-649.
- 2) Kopec KL, Burns D. Nonalcoholic fatty liver disease: a review of the spectrum of disease, diagnosis, and therapy. *Nutr Clin Pract* 2011;26:565-576.
- 3) Loomba R, Schork N, Chen CH, Bettencourt R, Bhatt A, Ang B, et al. Heritability of hepatic fibrosis and steatosis based on a prospective twin study. *Gastroenterology* 2015;149:1784-1793.
- 4) Speliotes EK, Yerges-Armstrong LM, Wu J, Hernaez R, Kim LJ, Palmer CD, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet* 2011;7:e1001324.
- 5) Anstee QM, Seth D, Day CP. Genetic factors that affect risk of alcoholic and nonalcoholic fatty liver disease. *Gastroenterology* 2016;150:1728-1744.
- 6) Hui ST, Parks BW, Org E, Norheim F, Che N, Pan C, et al. The genetic architecture of NAFLD among inbred strains of mice. *Elife* 2015;4:e05607.

- 7) van den Maagdenberg AM, Hofker MH, Krimpenfort PJ, de Bruijn I, van Vlijmen B, van der Boom H, et al. Transgenic mice carrying the apolipoprotein E3-Leiden gene exhibit hyperlipoproteinemia. *J Biol Chem* 1993;268:10540-10545.
- 8) Bennett BJ, Davis RC, Civelek M, Orozco L, Wu J, Qi H, et al. Genetic architecture of atherosclerosis in mice: a systems genetics analysis of common inbred strains. *PLoS Genet* 2015;11:e1005711.
- 9) Wang Z, Levison BS, Hazen JE, Donahue L, Li XM, Hazen SL. Measurement of trimethylamine-N-oxide by stable isotope dilution liquid chromatography tandem mass spectrometry. *Anal Biochem* 2014;455:35-40.
- 10) **Bennett BJ, Farber CR, Orozco L, Kang HM**, Ghazalpour A, Siemers N, et al. A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Res* 2010;20:281-290.
- 11) **Lippert C, Listgarten J**, Liu Y, Kadie CM, Davidson RI, **Heckerman D**. FaST linear mixed models for genome-wide association studies. *Nat Methods* 2011;8:833-835.
- 12) Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
- 13) **da Huang W, Sherman BT**, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4:44-57.
- 14) Rau CD, Parks B, Wang Y, Eskin E, Simecek P, Churchill GA, et al. High-density genotypes of inbred mouse strains: improved power and precision of association mapping. *G3: Genes - Genomes - Genetics* 2015;5:2021-2026.
- 15) Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 2015;31:2745-2747.
- 16) Leon-Mimila P, Vega-Badillo J, Gutierrez-Vidal R, Villamil-Ramirez H, Villareal-Molina T, Larrieta-Carrasco E, et al. A genetic risk score is associated with hepatic triglyceride content and non-alcoholic steatohepatitis in Mexicans with morbid obesity. *Exp Mol Pathol* 2015;98:178-183.
- 17) Shu L, Zhao Y, Kurt Z, Byars SG, Tukiainen T, Kettunen J, et al. Mergeomics: multidimensional data integration to identify pathogenic perturbations to biological systems. *BMC Genom* 2016;17:874.
- 18) Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;102:15545-15550.
- 19) Browning BL, Browning SR. Genotype imputation with millions of reference samples. *Am J Hum Genet* 2016;98:116-126.
- 20) Zhong H, Beaulaurier J, Lum PY, Molony C, Yang X, Macneil DJ, et al. Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS Genet* 2010;6:e1000932.
- 21) Consortium CAD, **Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL**, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet* 2013;45:25-33.
- 22) Mulder P, Liang W, Wielinga PY, Verschuren L, Toet K, Havekes LM, et al. Macrovesicular steatosis is associated with development of lobular inflammation and fibrosis in diet-induced non-alcoholic steatohepatitis (NASH). *Inflamm Cell Signal* 2015;2:e804.
- 23) **Westerterp M, van der Hoogt CC**, de Haan W, Offerman EH, Dallinga-Thie GM, Jukema JW, et al. Cholesteryl ester transfer protein decreases high-density lipoprotein and severely aggravates atherosclerosis in APOE*3-Leiden mice. *Arterioscler Thromb Vasc Biol* 2006;26:2552-2559.
- 24) **Tomita K, Teratani T**, Suzuki T, Shimizu M, Sato H, Narimatsu K, et al. Free cholesterol accumulation in hepatic stellate cells: mechanism of liver fibrosis aggravation in nonalcoholic steatohepatitis in mice. *Hepatology* 2014;59:154-169.
- 25) Kalhan SC, Guo L, Edmison J, Dasarathy S, McCullough AJ, Hanson RW, et al. Plasma metabolomic profile in nonalcoholic fatty liver disease. *Metabolism* 2011;60:404-413.
- 26) Farber CR, Bennett BJ, Orozco L, Zou W, Lira A, Kostem E, et al. Mouse genome-wide association and systems genetics identify *Asxl2* as a regulator of bone mineral density and osteoclastogenesis. *PLoS Genet* 2011;7:e1002038.
- 27) Liang W, Lindeman JH, Menke AL, Koonen DP, Morrison M, Havekes LM, et al. Metabolically induced liver inflammation leads to NASH and differs from LPS- or IL-1beta-induced chronic inflammation. *Lab Invest* 2014;94:491-502.
- 28) Cao F, Liu M, Zhang QZ, Hao R. PHACTR4 regulates proliferation, migration and invasion of human hepatocellular carcinoma by inhibiting IL-6/Stat3 pathway. *Eur Rev Med Pharmacol Sci* 2016;20:3392-3399.
- 29) **Jarray R, Pavoni S**, Borriello L, Allain B, Lopez N, Bianco S, et al. Disruption of phactr-1 pathway triggers pro-inflammatory and pro-atherogenic factors: new insights in atherosclerosis development. *Biochimie* 2015;118:151-161.
- 30) Van Hoof F, Hers HG. Mucopolysaccharidosis by absence of alpha-fucosidase. *Lancet* 1968;1:1198.
- 31) **Zhu Q, Qiao GL, Zeng XC**, Li Y, Yan JJ, Duan R, et al. Elevated expression of eukaryotic translation initiation factor 3H is associated with proliferation, invasion and tumorigenicity in human hepatocellular carcinoma. *Oncotarget* 2016;7:49888-49901.
- 32) **Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH**, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;518:197-206.
- 33) Abu-Serie MM, El-Gamal BA, El-Kersh MA, El-Saadani MA. Investigation into the antioxidant role of arginine in the treatment and the protection for intralipid-induced non-alcoholic steatohepatitis. *Lipids Health Dis* 2015;14:128.
- 34) Kathirvel E, Morgan K, Nandgiri G, Sandoval BC, Caudill MA, Bottiglieri T, et al. Betaine improves nonalcoholic fatty liver and associated hepatic insulin resistance: a potential mechanism for hepatoprotection by betaine. *Am J Physiol Gastrointest Liver Physiol* 2010;299:G1068-G1077.
- 35) Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab* 2009;9:311-326.
- 36) Teufel A, Itzel T, Erhart W, Brosch M, Wang XY, Kim YO, et al. Comparison of gene expression patterns between mouse models of nonalcoholic fatty liver disease and liver tissues from patients. *Gastroenterology* 2016;151:513-525.

Author names in bold designate shared co-first authorship.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.30113/supinfo.