

Hepatic expression profiling identifies steatosis-independent and steatosis-driven advanced fibrosis genes

Divya Ramnath,^{1,2} Katharine M. Irvine,^{3,4} Samuel W. Lukowski,^{1,2} Leigh U. Horsfall,^{3,5} Zhixuan Loh,^{1,2} Andrew D. Clouston,³ Preya J. Patel,^{3,5} Kevin J. Fagan,³ Abishek Iyer,^{1,2} Guy Lampe,⁶ Jennifer L. Stow,^{1,2} Kate Schroder,^{1,2} David P. Fairlie,^{1,2} Joseph E. Powell,^{1,2,7} Elizabeth E. Powell,^{3,5} and Matthew J. Sweet^{1,2}

¹Institute for Molecular Bioscience (IMB) and ²IMB Centre for Inflammation and Disease Research, The University of Queensland, Brisbane, Queensland, Australia. ³Centre for Liver Disease Research and ⁴Faculty of Medicine, Mater Research Institute, Translational Research Institute, The University of Queensland, Brisbane, Queensland, Australia. ⁵Department of Gastroenterology and Hepatology, Princess Alexandra Hospital, Brisbane, Queensland, Australia. ⁶Pathology Queensland, Princess Alexandra Hospital, Brisbane, Queensland, Australia. ⁷Garvan Institute of Medical Research, Sydney, New South Wales, Australia.

Chronic liver disease (CLD) is associated with tissue-destructive fibrosis. Considering that common mechanisms drive fibrosis across etiologies, and that steatosis is an important cofactor for pathology, we performed RNA sequencing on liver biopsies of patients with different fibrosis stages, resulting from infection with hepatitis C virus (HCV) (with or without steatosis) or fatty liver disease. In combination with enhanced liver fibrosis score correlation analysis, we reveal a common set of genes associated with advanced fibrosis, as exemplified by those encoding the transcription factor ETS-homologous factor (*EHF*) and the extracellular matrix protein versican (*VCAN*). We identified 17 fibrosis-associated genes as candidate *EHF* targets and demonstrated that *EHF* regulates multiple fibrosis-associated genes, including *VCAN*, in hepatic stellate cells. Serum *VCAN* levels were also elevated in advanced fibrosis patients. Comparing biopsies from patients with HCV with or without steatosis, we identified a steatosis-enriched gene set associated with advanced fibrosis, validating follistatin-like protein 1 (*FSTL1*) as an exemplar of this profile. In patients with advanced fibrosis, serum *FSTL1* levels were elevated in those with steatosis (versus those without). Liver *Fstl1* mRNA levels were also elevated in murine CLD models. We thus reveal a common gene signature for CLD-associated liver fibrosis and potential biomarkers and/or targets for steatosis-associated liver fibrosis.

Introduction

Hepatitis C virus (HCV) infection affects between 63 and 79 million people worldwide (1) and can lead to liver fibrosis, cirrhosis, hepatic decompensation, and hepatocellular carcinoma (HCC). Multiple direct-acting antiviral medications are now available to effectively treat HCV in over 95% of subjects, and, without advanced fibrosis (METAVIR stage 3–4), eradication of virus is regarded as a “cure.” Following a sustained virological response, there is gradual regression of hepatic fibrosis, even in patients with cirrhosis, and this is associated with a beneficial clinical outcome (2, 3). However, despite improvement in fibrosis, the cirrhotic liver does not show resolution of other key features, such as intrahepatic vascular shunts (4), and the risk of HCC persists (5). Therefore, in these subjects, ongoing surveillance for liver-related complications, particularly HCC, is recommended (6). Although the prevalence of cirrhosis in patients with chronic liver disease (CLD) remains unclear, a recent multicenter analysis of an HCV-infected cohort ($n = 9,783$ patients) in the US estimated that 28.5% of patients were cirrhotic and that liver disease severity was significantly underdocumented and underdiagnosed (7).

Fatty liver is a common histological feature in chronic HCV infection, reportedly present in 40%–86% of subjects (8), and is independently associated with fibrosis (9). Both viral and host factors, such as obesity and alcohol consumption, contribute to its development (8). Of concern, steatosis is associated with impaired regression of fibrosis after antiviral treatment (10), and in some studies (11, 12), but not all (13),

Conflict of interest: This work was partially funded by CSL Limited.

Submitted: February 2, 2018

Accepted: June 12, 2018

Published: July 25, 2018

Reference information:

JCI Insight. 2018;3(14):e120274.

<https://doi.org/10.1172/jci.insight.120274>.

insight.120274.

steatosis is a risk factor for development of HCC in patients with chronic HCV, even in subjects who achieve a sustained virological response to interferon therapy (14, 15). Nonalcoholic fatty liver disease (NAFLD) is estimated to affect 25% of the adult population globally (16), and it is likely to be an increasingly common comorbid condition in the HCV-infected population, irrespective of treatment status. A recent study showed that despite HCV eradication and absence of significant alcohol consumption, 10% of patients had persistently elevated liver enzymes and a further 25% had levels above the “healthy” range (<20 U/ml for women and <31 U/ml for men), consistent with ongoing inflammation (17). This study found that BMI significantly correlated with ongoing elevated liver enzymes, and the authors postulated that this may be due to development, or worsening, of preexisting NAFLD (17).

Although a virologic cure is available for most HCV-infected patients, there remains an unmet need for safe, effective therapeutic options to treat ongoing inflammation and fibrosis in some patients, particularly those with comorbid steatosis and metabolic risk factors. With better understanding of NAFLD pathogenesis, several therapies targeting different molecular pathways are being investigated, but current treatment options remain suboptimal (18). We hypothesized that identification of steatosis-driven profibrogenic pathways will provide insight into mechanisms by which comorbid fatty liver contributes to disease progression. To address this, we first performed RNA sequencing of liver biopsies ($n = 69$) to identify gene signatures that are associated with advanced fibrosis in patients with either HCV or fatty liver disease (FLD). We then compared HCV patient gene profiles, in the presence or absence of steatosis, to identify steatosis-enriched fibrosis-associated genes. Key findings from this analysis were subsequently validated in an independent NAFLD cohort. We also assessed, for the first time to our knowledge, the association between hepatic gene expression and the serum enhanced liver fibrosis (ELF) score. The ELF score is calculated from the levels of 3 serum markers, tissue inhibitor of metalloproteinase-1 (TIMP-1), N-terminal propeptide of type-III collagen (PIIINP), and hyaluronic acid (HA), and correlates strongly with fibrosis stage and liver-related clinical outcomes (19, 20). Clinical data included ELF scores, METAVIR fibrosis staging, and a range of other parameters, enabling us to adjust for confounders and perform accurate statistical analyses. Through these approaches, we revealed a gene set associated with advanced fibrosis, including a subset that is particularly elevated in the presence of steatosis. Our approach thus reveals potential new biomarkers and/or therapeutic targets for steatosis-associated liver fibrosis, which is rapidly increasing in prevalence and has no approved treatments.

Results

Patient characteristics at liver biopsy. Of the 69 patients included in the RNA-sequencing analysis, 55 had chronic HCV and 14 had FLD (12 NAFLD and 2 alcohol-related FLD). The HCV cohort included 26 subjects without steatosis and 29 subjects with steatosis (19, grade 1; 6, grade 2; and 4, grade 3 steatosis). The patient cohort was further subdivided into early and advanced fibrosis. Selected demographic and clinical characteristics of the patient cohort are summarized in Table 1.

Identification of genes associated with advanced fibrosis in CLD patients. As the first step in determining gene signatures associated with advanced liver fibrosis, we undertook RNA sequencing of liver biopsies from patients presenting with different etiologies (HCV or FLD) and fibrosis stages ($n = 69$ patients). These patients were categorized into early (stages 0–2) or advanced fibrosis (stages 3–4). To filter the data, a threshold of ≥ 10 counts per million (CPM) in at least 20 samples (29%) was set, with a total of 9,624 genes passing this criterion. Principal component analysis of the corrected counts for the 9,624 genes was performed to identify the components that contribute to maximum variability. As anticipated, the samples clustered based on early or advanced fibrosis (Figure 1A). In Figure 1B, a volcano plot highlights the 168 differentially expressed genes (DEGs) with study-wide significance between early and advanced fibrosis. Of these, 61 showed ≥ 2 -fold difference (Figure 1B, blue), while 107 significant genes were differentially expressed but showed ≤ 2 -fold difference (Figure 1B, orange). The less stringent FDR-corrected analysis revealed 1,595 DEGs between early and advanced fibrosis (Supplemental Data Set 1; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.120274DS1>). We next examined the expression patterns of the 168 DEGs across the patient samples. Figure 1C clearly shows a distinct separation in gene expression patterns in liver biopsies of patients with early versus advanced fibrosis. For a more comprehensive overview of the gene expression signature, we subjected the 1,595 (FDR-corrected) DEGs to KEGG pathway analysis and found that many pathways linked with cancer and/or metabolism, for example, the PI3K/Akt signaling pathway, were enriched in this gene set (Figure 1D). Further pathway analysis by MetaCore also identified predicted cellular processes (e.g., extracellular matrix [ECM] remod-

Table 1. Selected demographic and clinical characteristics of the patient cohort for the initial RNA-sequencing analysis

Variable	HCV without steatosis			HCV + steatosis			FLD			P value
	Early	Advanced	P value	Early	Advanced	P value	Early	Advanced	P value	
No. of patients ^A	16	10	0.92	16	13	0.40	10	4	0.36	NA
Age (yr) ^B	38 (21–55)	49 (39–64)	0.0031	39 (25–57)	51 (43–61)	0.0003	45 (28–62)	51 (47–58)	0.1102	6.7e-07
Sex (male) ^A	9 (56)	7 (70)	0.48	12 (75)	11 (85)	0.52	7 (78)	3 (75)	0.85	0.32
BMI (kg/m ²) ^B	25 (18–29)	27 (17–33)	0.25	27 (21–35)	28 (21–35)	0.47	33 (28–39)	38 (33–45)	0.13	0.21
Alcohol (around biopsy) (g/w) ^B	87 (0–323)	2 (0–10)	0.006	37 (0–160)	69 (0–525)	0.49	77 (0–560)	5 (0–15)	0.22	0.29
ELF score ^B	8.8 (8.06–9.69)	10.3 (8.66–12.64)	0.014	8.7 (7.69–10.26)	10.5 (9.1–12.64)	0.0002	8.6 (8.32–9.02)	9.9 (9.6–10.23)	0.19	6.8e-07
Viral genotype ^A							NA	NA	NA	
1	10 (62)	6 (60)	0.89	6 (38)	7 (54)	0.37				0.63
3	5 (32)	3 (30)	0.94	9 (56)	4 (31)	0.17				0.31
Other	1 (6)	1 (10)	0.72	1 (6)	2 (15)	0.42				0.38
Steatosis grade ^A	NA	NA	NA							
1				12 (75)	7 (54)	0.23	1 (10)	0 (0)	0.51	0.57
2				1 (6)	5 (38)	0.03	3 (30)	1 (25)	0.85	0.13
3				3 (19)	1 (8)	0.39	6 (60)	3 (75)	0.59	0.43
Inflammation grade ^A										
0	1 (6)	0 (0)	0.42	0 (0)	0 (0)	NA	0 (0)	0 (0)	NA	0.41
1	12 (75)	3 (30)	0.023	10 (63)	4 (31)	0.08	7 (70)	0 (0)	0.017	0.0004
2	3 (19)	4 (40)	0.23	6 (37)	8 (62)	0.19	3 (30)	4 (100)	0.017	0.011
3	0 (0)	3 (30)	0.019	0 (0)	1 (7)	0.25	0 (0)	0 (0)	NA	0.01
Fibrosis stage ^A										
0	6 (23)	NA	NA	3 (10)	NA	NA	3 (22)	NA	NA	
1	5 (19)	NA		5 (17)	NA		5 (36)	NA		
2	5 (19)	NA		8 (28)	NA		2 (14)	NA		
3	NA	7 (27)		NA	5 (17)		NA	2 (14)		
4	NA	3 (12)		NA	8 (28)		NA	2 (14)		

^AData presented categorically as *n* (%) and analyzed using χ^2 test. ^BData presented continuously as median \pm range and analyzed using Welch's *t* test. *P* values in bold are statistically significant.

eling) and signaling pathways (e.g., TGF- β , WNTs, and Hedgehog) (Supplemental Data Set 2), which have previously been implicated in CLD progression (21, 22). In addition to the above analyses, we also performed more stringent (≥ 10 CPM in all samples) and relaxed (≥ 10 CPM in at least 1 sample) filtering of the data. Using these criteria, 5,404 and 12,596 genes were captured by the stringent and relaxed filtering approaches, respectively. For the stringently filtered set of 5,404 genes, 71 (study-wide significance) and 791 (FDR-corrected) genes were differentially expressed between early- and advanced-stage fibrosis, while, for the gene list generated through more relaxed filtering (12,596 genes), 314 (study-wide significance), and 2,883 (FDR-corrected) genes were identified. These data are also provided in Supplemental Data Set 1. All subsequent analyses focused on the medium filtered set of 9,624 genes.

Identification of a gene signature that correlates with ELF score. The ELF score, calculated from serum levels of TIMP-1, PIIINP, and HA, is widely accepted as a noninvasive tool to assess liver fibrosis severity (19, 23). To determine the genes correlating with ELF score, we took the corrected counts for all 9,624 transcripts and performed a linear regression analysis with the corresponding ELF scores. We identified 37 genes that significantly correlated with the ELF score, with an r^2 value of 0.3 or higher (Supplemental Data Set 1). As expected, most of the ELF score-correlated genes (86%) were also identified as DEGs in the early versus advanced fibrosis comparisons above. Genes that showed the strongest correlation ($r^2 \geq 0.39$) are plotted in Figure 2. The genes that most closely correlated with ELF score were cystic fibrosis transmembrane conductance regulator (*CFTR*) and osteopontin (*SPP1*), both of which have previously been implicated in regulating liver fibrosis (24–27), thus validating our approach. Interestingly, the transcriptional regulator ETS-homologous factor (EHF), also known as ESE-3, which itself is a DEG (Figure 1 and Supplemental Data Set 1) that correlates with ELF score (Figure 2), has been functionally linked to other genes in this list,

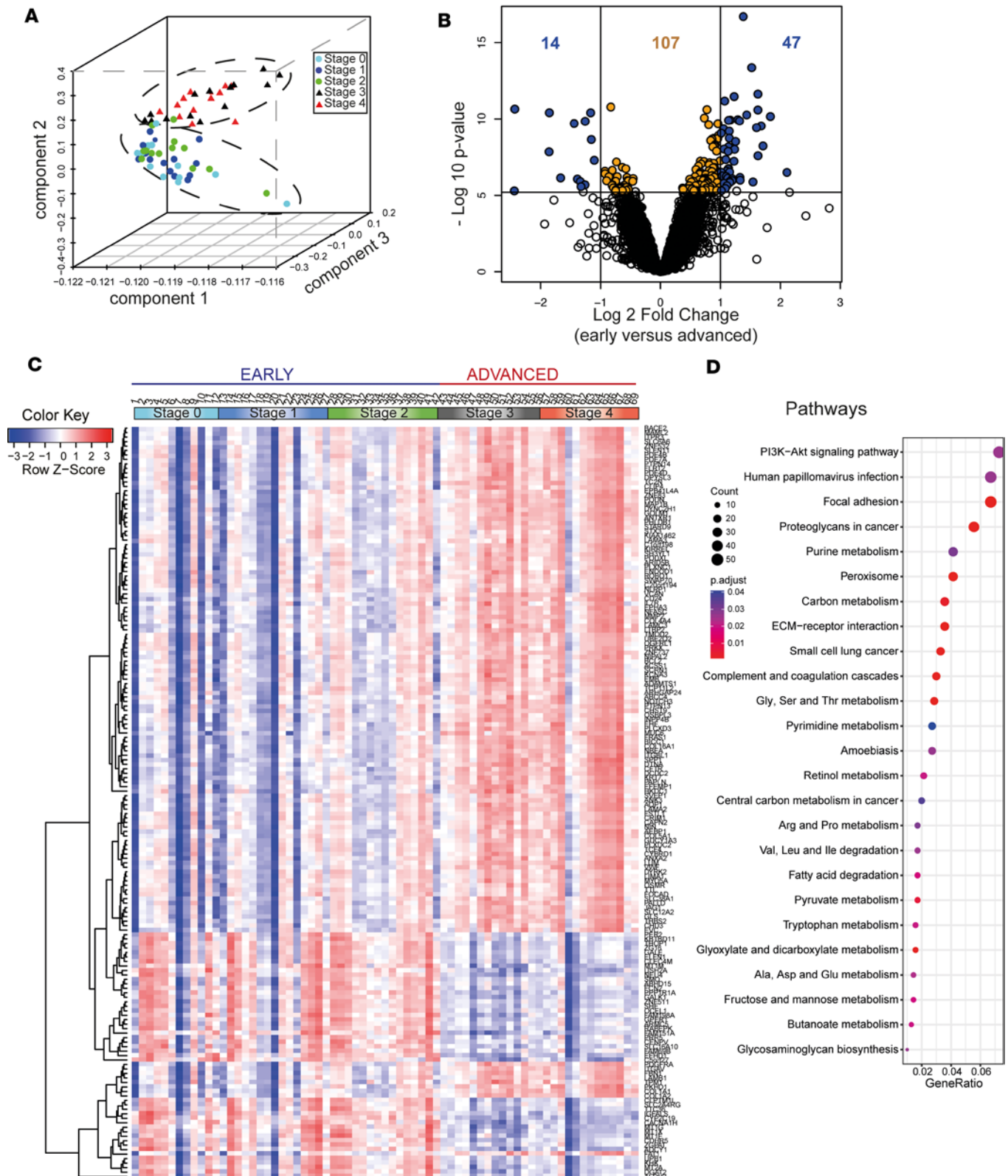


Figure 1. An overview of DEGs in early- versus late-stage liver fibrosis. (A) Principal component analysis of the corrected counts for the filtered genes (9,624 genes), highlighting the components that contribute to maximum variability ($n = 69$ patients; circles, early stages; triangles, advanced stages). (B) Volcano plot illustrating genes, from the filtered data set, that are significantly differentially expressed between early and advanced stages of liver fibrosis, as identified by RNA sequencing ($n = 69$ patients; EdgeR-generated FWER value and log₂ fold change). (C) A heatmap showing the expression patterns of the 168 DEGs after hierarchical clustering using the complete linkage method ($n = 69$ patients). (D) KEGG pathway analysis of the DEGs between early- versus late-stage fibrosis ($n = 69$ patients; 1,595 FDR-corrected list). Circle size correlates with number of genes, and circle color indicates statistical significance.

for example, *CFTR*, *SPPI*, and *CD24* (28–30). To determine if EHF may be more widely linked to core liver fibrosis-associated genes, we compared the list of 37 ELF score–correlating genes to EHF target genes that were previously identified via a ChIP-seq approach in human airway epithelial cells (31). Indeed, 10 of the 37 ELF score–correlated genes that are expressed in fibrotic liver were also characterized as EHF targets in that study (Figure 3, A and B), representing a significant enrichment (hypergeometric test; $P = 0.01$). Additionally, we used an informatics approach (RcisTarget) (32) to identify the overrepresented transcription factor–binding motifs in the ELF score–correlated genes and found that the ETS motif, to which EHF binds, was significantly enriched (enriched in 9 genes, enrichment score 3.16) (Figure 3C and Supplemental Data Set 1). Combined, these two approaches implicated EHF as a candidate regulator of 17 of the 37 ELF score–correlated genes. We next used siRNA-mediated gene silencing of *EHF* in LX-2 hepatic stellate cells (HSCs) (Figure 3D) to determine whether this transcription factor is a likely regulator of TGF- β –inducible expression of ELF score–correlated genes. We examined 10 ELF score–correlated genes and found that 4 of these genes, *VCAN* (encoding versican), *DHRS2* (encoding dehydrogenase/reductase family member 2), *COL1A1* (encoding collagen type I, $\alpha 1$ chain) and *DTNA* (encoding dystrobrevin α), were TGF- β regulated (Figure 3, E–H). *EHF* silencing significantly reduced TGF- β –inducible mRNA expression of *VCAN*, *DHRS2*, and *COL1A1* (Figure 3, E–G), while TGF- β –mediated suppression of *DTNA* mRNA expression was unaffected by *EHF* knockdown (Figure 3H). Silencing of *IRF6*, another epithelial-expressed transcription factor, had no significant effect on TGF- β –inducible *VCAN*, *DHRS2*, or *COL1A1* mRNA levels. Collectively, these data suggest that EHF regulates the expression of multiple liver fibrosis-associated genes.

Genes encoding extracellular proteins that are associated with progressive liver fibrosis. In order to identify gene targets that may represent novel diagnostics and/or be amenable to targeting with biologicals, we subjected the list of 168 DEGs to pathway analysis tools and found that 77 of these genes encode secreted and/or cell-surface proteins (Supplemental Data Set 1). Genes encoding extracellular proteins showing greater than 2.5-fold differences between early and advanced stages are highlighted in Table 2. We also performed this analysis on the FDR-corrected gene list (1,595 genes), identifying 550 genes encoding extracellular proteins (Supplemental Data Set 1). To provide further stringency to our analysis, we compared the 168 DEGs to other published data sets, including an RNA sequencing–based study in a NAFLD cohort (33) and 3 microarray-based studies in NAFLD cohorts (34–36). We observed an overlap of 48 genes from our list with at least 3 other published data sets, with 16 of these genes being present in all 5 data sets (Figure 4). Of these 48 genes, 35 encode extracellular proteins (Figure 4B, highlighted in gray). Our attention was drawn to one of these, *VCAN*, which encodes an ECM protein recently implicated in the activation of HSCs (37) and that has also been linked to inflammation and cancer metastasis (38). *VCAN* mRNA expression strongly correlated with ELF score (r^2 value of 0.41) (Figure 2) and was inducible by the profibrotic cytokine TGF- β in an EHF-dependent manner in the LX-2 HSC line (Figure 3E). Figure 5A shows that *VCAN* mRNA was significantly elevated in biopsies of patients with advanced fibrosis. Four different isoforms of *VCAN* (*V0–V3*) have been described (39) (Figure 5B). To further validate our findings, and determine whether specific *VCAN* isoforms were differentially regulated, we performed qPCR for the different transcripts. The *V0* and *V1* isoforms were readily detectable and significantly elevated in liver biopsies from patients with advanced disease, whereas the *V2* and *V3* isoforms were expressed at very low to undetectable levels and showed no significant differences (Figure 5, C–F). We also examined the serum levels of VCAN in an independent cohort of patients with chronic HCV or NAFLD (Supplemental Table 1). Circulating VCAN was detected in only a subset of these patients (26%), and in this subset, levels were significantly elevated in those with advanced fibrosis (Figure 5G).

Mapping the steatosis-associated liver fibrosis gene signature. The above approach identified gene expression changes that occurred in early versus advanced stages of fibrosis, irrespective of the underlying causes of CLD. However, our profiling of HCV patients with or without steatosis provided an opportunity to identify steatosis-associated genes linked to advanced fibrosis. To address this, we reanalyzed the data by comparing patients with and without steatosis across all fibrosis stages. Somewhat surprisingly, this approach did not identify statistically significant DEGs (Supplemental Data Set 3). We presume that the strong fibrosis-associated gene signature overrides any steatosis-specific one. Consequently, we next identified DEGs between early and advanced fibrosis in HCV-infected patients with or without steatosis ($n = 26$ and $n = 29$, respectively). Whereas no genes showed study-wide significant differences in early versus advanced fibrosis in HCV patients without steatosis (Figure 6A), 58 genes had study-wide significance for early versus advanced fibrosis in patients with HCV and steatosis (Figure 6B). 22 genes were excluded from further analyses due to high variability between patient samples

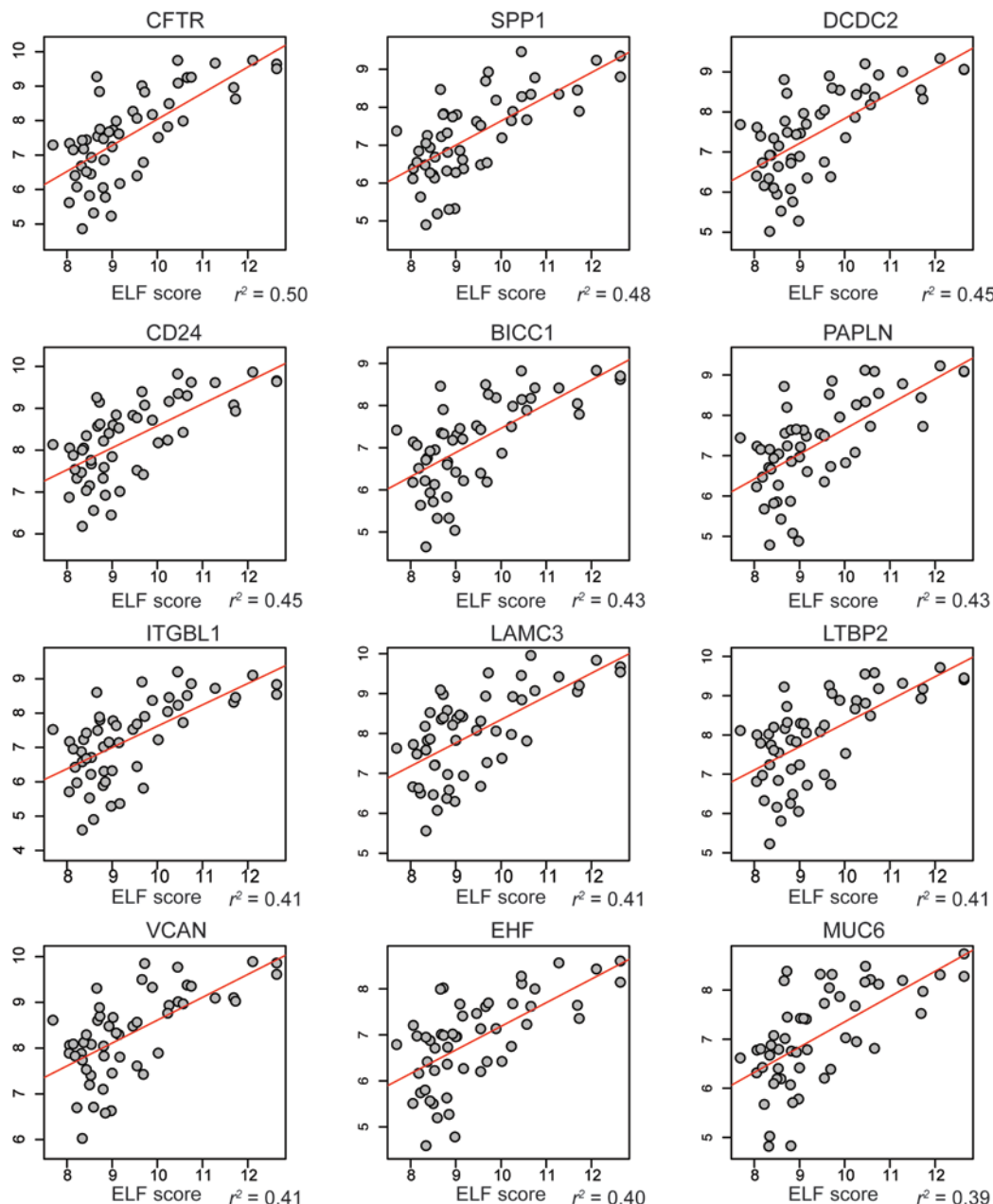


Figure 2. Gene signatures correlating with patient ELF scores. Correlation analysis was performed using log₂-transformed corrected counts of genes and patient ELF scores ($n = 69$ patients). Linear regression of the top 12 genes that most strongly correlate with ELF score (r^2 value of 0.39 or more), showing gene expression as an independent variable and ELF score as a dependent variable.

(coefficient of variance > 1), as this confounded data analysis. We further examined the expression patterns of the remaining 36 genes across the HCV patient subgroups (with or without steatosis) and found that most genes show similar trends in patients with early versus advanced fibrosis; however, the effect was more pronounced in those with steatosis (Figure 6C). To identify the subset of these fibrosis-associated genes that was most highly elevated in patients with steatosis, we calculated the fold change in early versus advanced fibrosis in the HCV versus HCV with steatosis patient subgroups. We identified 12 fibrosis-associated genes that showed greater than 3-fold differences between HCV patients with and without steatosis (Figure 6D), of which 10 encoded extracellular proteins (Table 3). Of the genes falling into this category, follistatin-like protein 1 (*FSTL1*), a member of the SPARC family (40), showed the highest fold difference (5.46-fold).

FSTL1 as a candidate target for steatosis-associated liver fibrosis. *FSTL1* is known to drive pathology in inflammatory disease settings in mouse models, particularly arthritis (41–47). Among its various activities, it has been reported to promote inflammatory cytokine production, including IL-1 β , from macrophages and fibroblasts (41, 47, 48). While *FSTL1* mRNA was significantly elevated in CLD

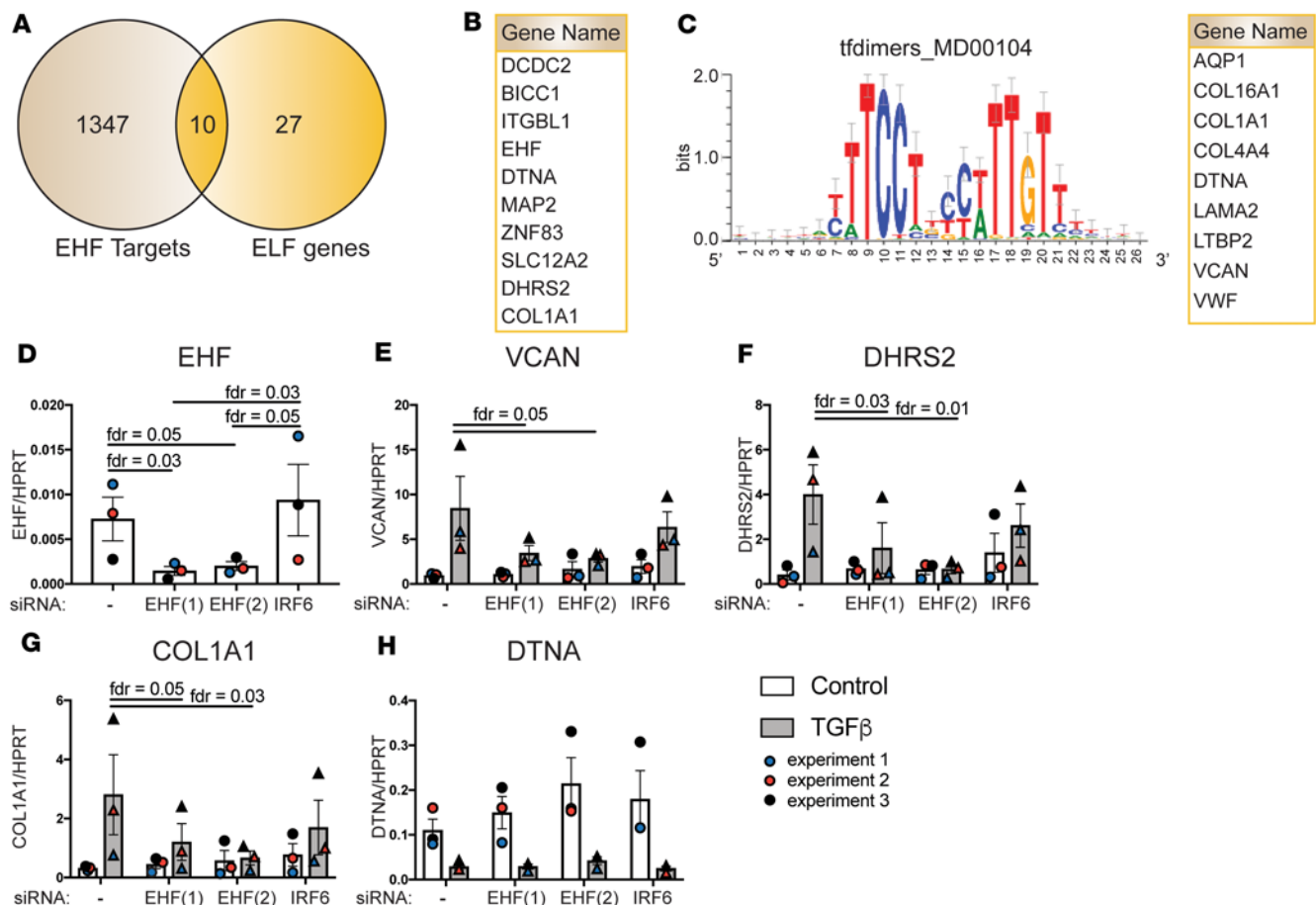


Figure 3. EHF is a candidate regulator of multiple genes associated with advanced liver fibrosis. (A) A Venn diagram showing the overlap between the ELF score-correlated genes and previously identified transcriptional targets of EHF (31). (B) A list of the 10 ELF score-correlating EHF candidate target genes from A. (C) The EHF-binding motif is significantly enriched in 9 ELF score-correlated genes (enrichment score 3.16), as identified by RcisTarget. (D–H) Expression of EHF and IRF6 (control gene) was silenced in LX-2 cells, with two independent siRNAs being used for EHF. After 24 hours, cells were stimulated with 10 ng/ml TGF-β (gray bars) for 24 hours or were left unstimulated (white bars), after which RNA was prepared and qPCR was performed. (D) Basal levels of *EHF* mRNA as well as TGF-β-regulated levels of mRNAs, relative to *HPRT*, for (E) *VCAN*, (F) *DHR2*, (G) *COL1A1*, and (H) *DTNA* were quantified by qPCR. Data represent mean ± SEM from 3 independent experiments. FDR values were calculated using (D) nonparametric ANOVA (Kruskal-Wallis test) or (E–H) 2-way ANOVA followed by Benjamini-Hochberg multiple-testing corrections. FDR values ≤ 0.05 were considered statistically significant.

patients with advanced fibrosis (Figure 7A), it showed only a modest correlation with ELF score (r^2 value of 0.12) (Figure 7B). This lack of correlation is not surprising, given that elevated expression was only apparent in late-stage disease for patients with steatosis. Indeed, *FSTL1* was significantly elevated in advanced versus early fibrosis stages in patients with HCV and steatosis but showed no significant difference in patients with HCV infection alone (Figure 7C). To further validate our findings, we examined serum levels of FSTL1 in an independent cohort of patients with chronic HCV or NAFLD (Supplemental Table 1). Circulating FSTL1 was detected in patients with steatosis, with significantly higher levels being apparent in advanced-stage NAFLD patients (Figure 7D). Given its potential as a steatosis-specific marker of advanced liver fibrosis, we also examined serum levels of FSTL1 in an independent NAFLD patient cohort (Supplemental Table 2) and again found that this protein was significantly elevated in patients with advanced fibrosis (Figure 7E). To determine whether similar effects were apparent during murine CLD, we examined *Fstl1* mRNA levels in two mouse models of progressive liver fibrosis associated with different types of chronic liver injury, the hepatotoxin thioacetamide (TAA) model and the choline-deficient, ethionine-supplemented (CDE) model that develops steatosis. *Fstl1* mRNA levels in the livers of mice subjected to both of these regimes were elevated by comparison with those of control mice (Figure 7, F and G).

Table 2. List of 11 DEGs (with 2.5-fold difference in expression between early and advanced stages) that encode cell-surface proteins (3 genes), secreted proteins (7 genes), or both (1 gene)

Name	Log ₂ fold change	FWER value	Cell surface/secreted
CFTR	1.83	6.67×10^{-7}	Cell surface
ITGBL1	1.68	2.79×10^{-6}	Both
EFEMP1	1.63	2.45×10^{-4}	Secreted
LTBP2	1.62	2.50×10^{-7}	Secreted
SPP1	1.60	3.60×10^{-6}	Secreted
BICC1	1.52	4.21×10^{-7}	Secreted
PAPLN	1.44	5.27×10^{-7}	Secreted
VCAN	1.38	1.93×10^{-13}	Secreted
LUM	1.32	1.29×10^{-6}	Secreted
USH2A	-1.32	2.60×10^{-2}	Cell surface
CACNA1H	-1.86	1.36×10^{-4}	Cell surface

Genes with increased and decreased expression in advanced fibrosis, compared with early-stage fibrosis, are indicated by positive and negative log₂ fold change. A full list of extracellular proteins is provided in Supplemental Data Set 1.

Discussion

Our combinatorial approach of expression profiling CLD patient liver biopsies in parallel with ELF score analysis has enabled us to robustly identify genes associated with advanced fibrosis. Importantly, our analyses also enabled us to identify genes that were particularly elevated in steatosis-associated advanced fibrosis. Our clinically well-characterized patient cohort allowed us to perform stringent statistical corrections for confounding factors to identify genes that are altered at the population level. Pathway analysis of advanced fibrosis-associated genes revealed those involved in ECM remodeling, cell adhesion, cancer, TGF- β signaling, epithelial-to-mesenchyme transition, and immune response pathways (Supplemental Data Set 2). Many of the liver fibrosis-associated pathway signatures that were identified, including the dysregulation of several metabolic pathways (Figure 1D), are broadly consistent with those reported in other published studies (33, 35, 49–52). Furthermore, some of the individual DEGs have strong associations with liver fibrosis (27, 53–56).

We added another layer of cross-validation to our analysis of fibrosis-associated gene expression by correlating patient ELF scores with gene expression. The ELF score correlates strongly with liver-related clinical outcomes (20), and the 37 genes identified through this approach (Supplemental Data Set 1) will be of particular interest for future functional studies. For example, *CFTR* and *SPP1* showed the strongest correlation with ELF score (Figure 2). Approximately 30% of patients with cystic fibrosis develop liver fibrosis (53), and *Cftr* deficiency in mice predisposes to the development of liver fibrosis upon exposure to a profibrogenic xenobiotic (24). Similarly, *SPP1* was recently identified as a mediator of liver fibrosis via activation of HSC (27). Several other ELF score–correlated genes, including *DCDC2* and *ITGBL1* (57), and *CD24* and *BICC1* (58), have emerging genetic or functional links with liver fibrosis. Furthermore, the identification of the Ets-family transcription factor EHF, which has been linked to the regulation of the ELF score–correlated genes *CFTR*, *SPP1*, and *CD24* (28–30), implies that this transcriptional regulator may be an important driver of fibrogenesis. Indeed, *ITGBL1*, *DCDC2*, *CD24*, and *EHF* (all ELF score–correlated genes) were identified in independent studies as genes closely associated with HBV-associated liver fibrosis (57) and/or cirrhosis (59). Interestingly, 17 of the 37 ELF score–correlated genes, which were all expressed in fibrotic liver, were identified as EHF targets in airway epithelial cells (31) and/or were enriched for the EHF-binding motif, as identified by RcisTarget. Although EHF has been reported to act as a transcriptional repressor (60), individual members of the ETS family can act as both transcriptional activators and repressors in different contexts (61, 62). For example, EHF promoted expression of the death receptor DR-5 in HCT116 colorectal cancer cells (63) but inhibited Ras-inducible expression of target genes in differentiated epithelial cells (60). In our studies, silencing of *EHF* in the LX-2 HSC line impaired TGF- β –inducible mRNA expression of *VCAN*, *DHRS2*, and *COL1A1* (Figure 3, E–G), thereby supporting the view that this transcription factor promotes expression of multiple genes that are associated with advanced fibrosis. Of note, however, *EHF* silencing in LX-2 cells did not affect TGF- β –mediated suppression of *DTNA* (Figure 3H), and the 6 other ELF score–correlated genes that were examined were

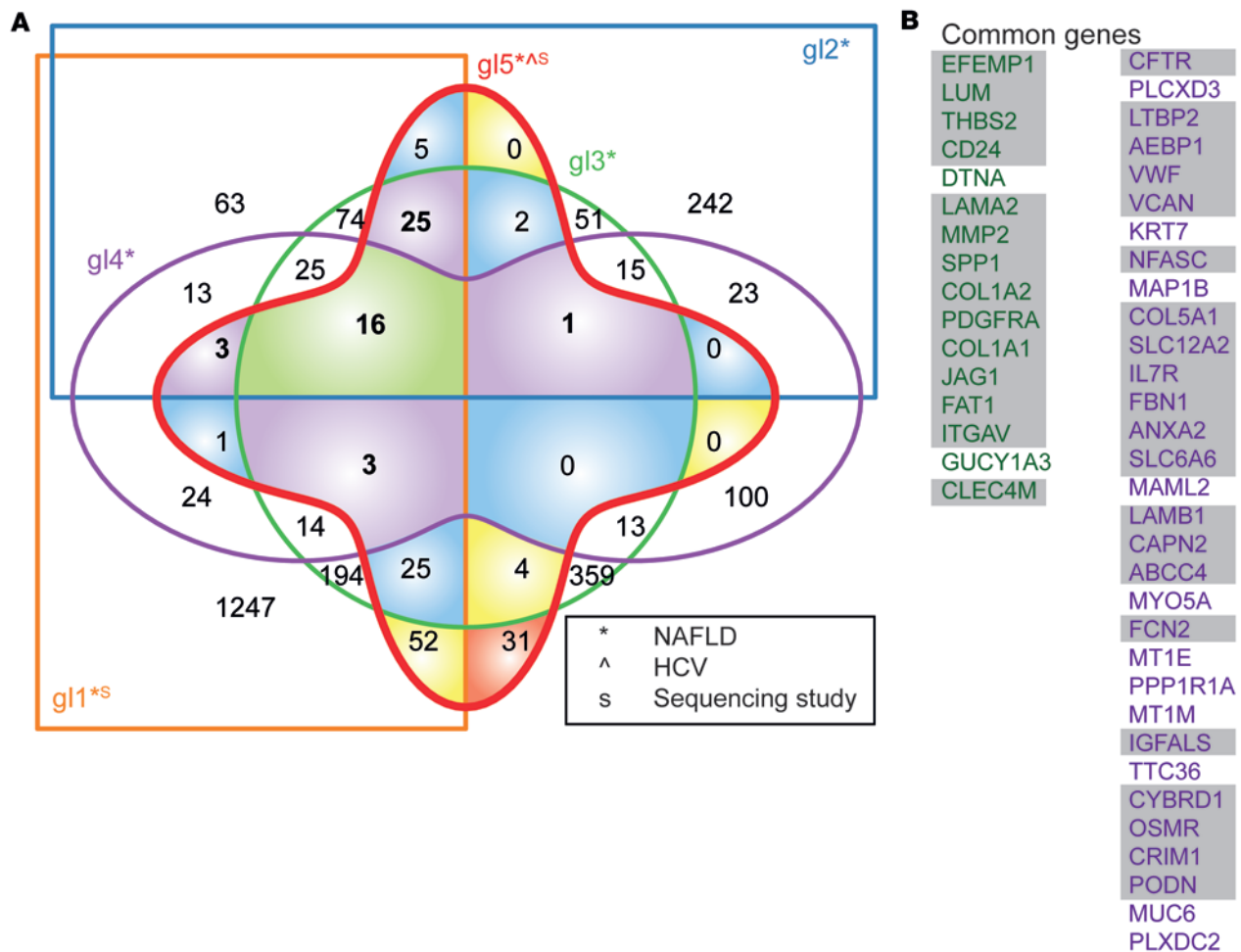


Figure 4. Meta-analysis of existing microarray and RNA-sequencing data sets for liver fibrosis. (A) A Venn diagram comparing the DEGs from this study (gl5, red) to 4 other published gene lists (gl1, gl2, gl3, and gl4). 16 genes (green) were common across all 5 gene lists, while 32 genes (purple) were present in 4 of the 5 gene lists (including this one). 31 genes were unique to this data set (red). (B) The list of 48 DEGs that were common in our data set and at least 3 other published data sets (green, present in all data sets; purple, present in 4 of 5 data sets, including this one; gray, extracellular proteins).

not regulated by TGF- β in these cells (data not shown). It is possible that these genes are controlled by other profibrotic and/or proinflammatory stimuli and/or are expressed in other cell types, given that only TGF- β responses in HSCs were examined here. Whether EHF regulates the expression of these genes in such settings remains to be determined, and it is also conceivable that only a subset of candidate EHF targets identified in our analyses are bona fide targets of this transcription factor. Regardless of this, future functional studies on the role of EHF in the progression and severity of liver fibrosis are clearly warranted. Recent studies have revealed that genetic variation in EHF may be associated with severity of airways disease in cystic fibrosis patients (64–66), a subset of which also develop liver disease (53).

By refining the list of DEGs to those encoding extracellular proteins, we describe gene candidates that may be amenable to biological-based therapeutic targeting. A similar approach was recently pursued in a NAFLD patient cohort, where 110 genes were found to be cell-surface proteins (33). *PAPLN* and *JAG1*, for example, were also identified as differentially expressed extracellular proteins in our study. By comparing the advanced fibrosis-associated gene set from this study (168 genes) with those derived from other CLD expression profiling studies (33–36), we were able to refine the list to 35 genes encoding extracellular targets that were common to at least 4 of 5 data sets, including our own (Figure 4B). Several of these (e.g., *SPP1* and *CFTR*) have strong genetic or functional links to liver fibrosis, but we focused on validating expression of *VCAN* (Figure 5), given the limited information that exists on its role in this disease process. Interestingly, we found that the *V0* and *V1* isoforms were elevated in advanced liver fibrosis (Figure 5, C and D), whereas other alternatively spliced transcripts were expressed at low to undetectable levels (Figure 5, E and F). This

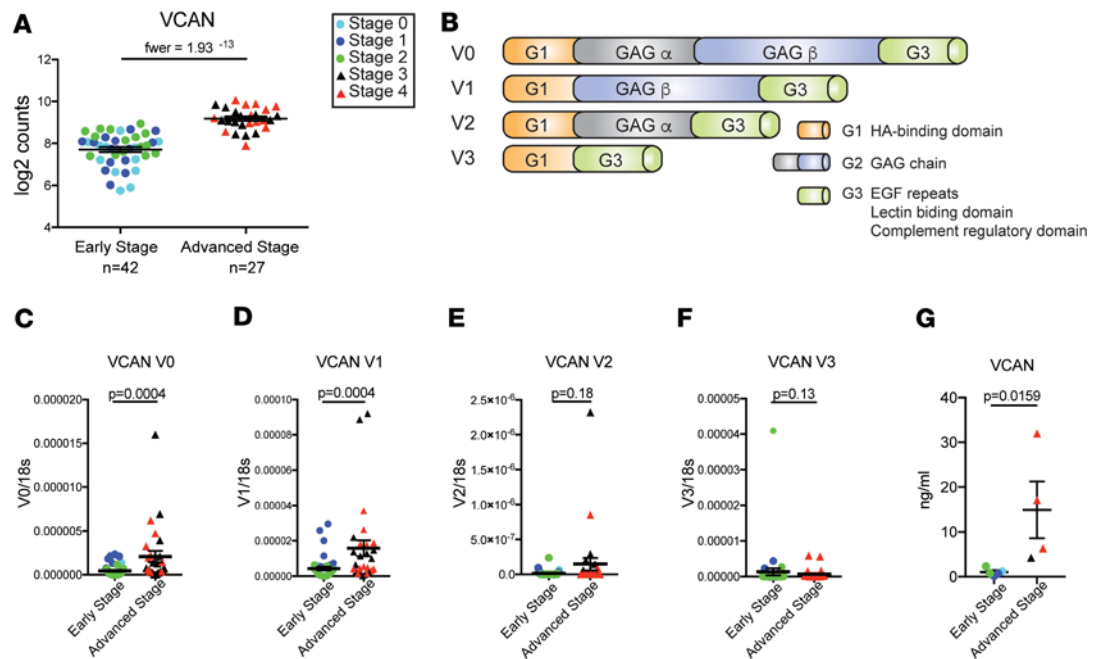


Figure 5. VCAN is upregulated in advanced stages of liver fibrosis. (A) A scatter plot showing the log₂-corrected counts of VCAN for all patients from the RNA-sequencing analysis ($n = 69$ patients; horizontal bars indicate mean \pm SEM; circles, early stages; triangles, advanced stages). (B) A schematic diagram of the different isoforms of VCAN. (C–F) qPCR analysis was performed on the mRNA from all patient biopsies used in the RNA-sequencing study. Levels of the different isoforms of VCAN, (C) *V0*, (D) *V1*, (E) *V2*, and (F) *V3*, were determined, relative to *18S* ribosomal RNA, by qPCR ($n = 69$ patients; horizontal bars indicate mean \pm SEM; circles, early stages; triangles, advanced stages). (G) Circulating levels of VCAN were determined by performing ELISA on human patient sera samples (see Supplemental Table 1 for details of the patient cohort; $n = 35$; detectable levels in 9 patients; horizontal bars indicate mean \pm SEM; circles, early stages; triangles, advanced stages). $P \leq 0.05$, calculated using the Mann-Whitney test.

is in keeping with other studies demonstrating tissue-specific expression of VCAN isoforms (67–70). The *V0* and *V1* isoforms are also reportedly expressed in some tumors (71, 72), whereas the *V2* isoform has been demonstrated to inhibit cell proliferation (73). Thus, it is tempting to speculate that selective expression of the *V0* and *V1* isoforms in advanced fibrosis may have some role in HCC progression. Recently, VCAN was suggested to have a role in hepatic fibrosis, with in vitro studies showing that it promotes HSC activation (37). However, its role in driving inflammation and liver fibrosis during CLD is unknown. Based on existing literature, it might be predicted to have dual roles in liver disease processes, promoting both HSC activation and inflammation. VCAN appears to exert several distinct proinflammatory functions. First, its presence in the ECM promotes leukocyte adhesion and inflammatory responses (74). Second, it can directly trigger Toll-like receptor 2–mediated activation of macrophages, a response that has been linked to metastasis in a breast cancer model (38). VCAN is processed by members of the a disintegrin and metalloproteinase with thrombospondin type-1 motifs (ADAM-TS) family of proteases, including ADAM-TS1, the expression of which was also elevated in patients with advanced fibrosis (Supplemental Data Set 1). Members of this family have previously been functionally linked to liver fibrosis pathology (75, 76). Therefore, it is possible that processing of VCAN generates a profibrotic activity that contributes to disease progression. However, VCAN has a well-recognized role in maintaining the ECM structure by interacting with other ECM components, such as hyaluronan, fibronectin, and type-I collagen (77). A contribution of full-length VCAN to the progression of liver fibrosis is therefore also possible.

Our comparison of patients with HCV with patients with HCV and those with HCV and steatosis enabled us to identify genes that were most highly expressed in patients with both steatosis and advanced fibrosis. This approach revealed 12 such genes (Figure 6D), 10 of which encoded extracellular proteins (Table 3). Since inflammation, associated cell death, and aberrant wound healing are central to the progression of liver fibrosis in NAFLD (78), we pursued FSTL1 as a steatosis-linked gene, because of its strong links to both inflammation and fibrosis (40, 41, 43–45, 79). FSTL1 is a secreted glycopro-

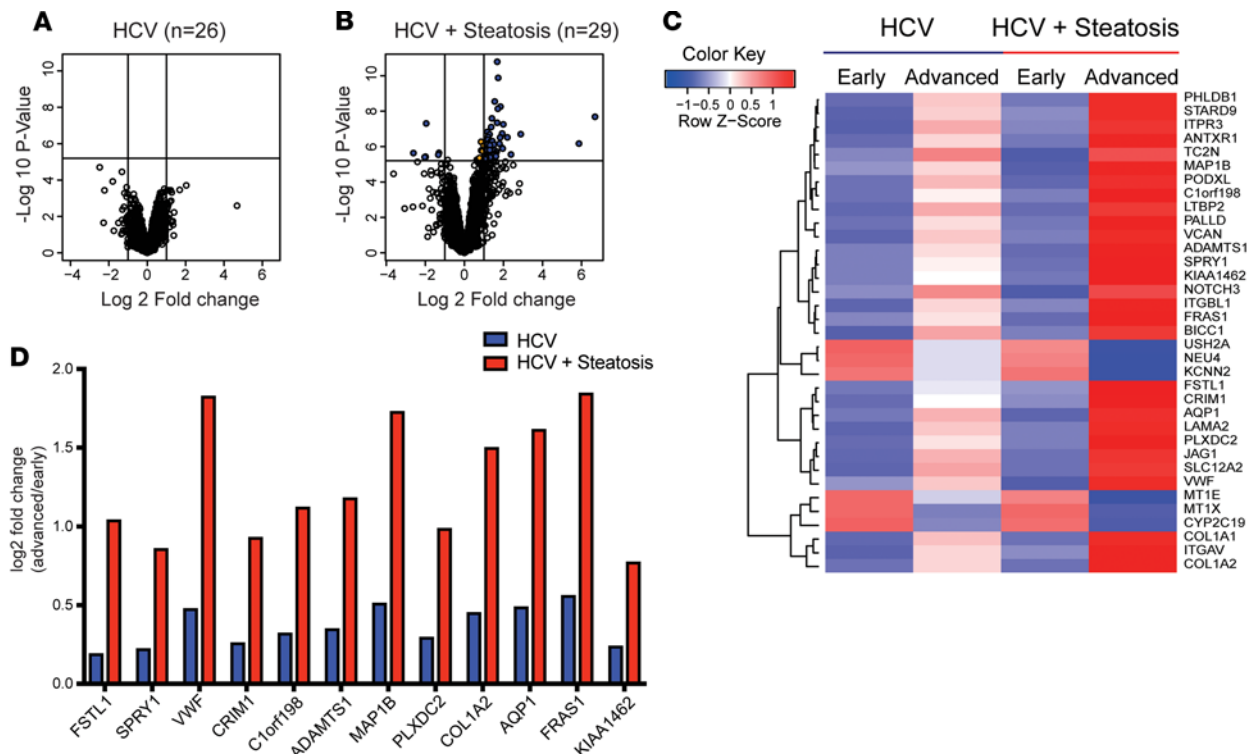


Figure 6. Identification of steatosis-associated late-stage fibrosis genes. A volcano plot illustrating significant differences in genes between early and advanced stages of fibrosis by RNA sequencing in patients with (A) HCV alone ($n = 26$ patients; EdgeR-generated FWER value and \log_2 fold change) and (B) HCV with steatosis ($n = 29$ patients; EdgeR-generated FWER value and \log_2 fold change). (C) A heatmap showing the average expression patterns of the 36 genes, identified as differentially expressed in the HCV and steatosis patient cohort, in the patients with either HCV alone or HCV and steatosis. (D) A bar graph of 12 steatosis-enriched late-stage fibrosis genes, showing fold change in expression between late- versus early-stage fibrosis in HCV patients (blue bars) and HCV with steatosis patients (red bars).

tein belonging to the SPARC family (40). Although it is known to drive pathology in mouse models of inflammatory diseases, particularly arthritis (43, 45, 46), the mechanisms involved remain unclear. However, a consistent theme is that it is linked to proinflammatory cytokine production from multiple cell types. For example, FSTL1 promoted inflammatory cytokine production, including IL-1 β , from macrophages and fibroblasts (44, 48). Overexpression of Fstl1, by in vivo adenoviral transduction in DBA/1 mice, promoted expression of the proinflammatory cytokines TNF- α , IL-1 β , and IL-6 in the liver (47). Although it is currently unclear which receptors are required for FSTL1-mediated inflammatory responses, recent studies have implicated both the NLRP3 inflammasome and IL-17 signaling in downstream responses, both of which are strongly implicated in liver fibrosis and adipogenesis (48, 80–85). Importantly, FSTL1 has also recently been linked to pathology in pulmonary (79) and hepatic fibrosis (86) in mice. Silencing Fstl1 using shRNA reduced *Colla1* mRNA expression and macrophage accumulation, while increasing myofibroblast density and *Timp-1* mRNA expression in carbon tetrachloride injury-mediated liver fibrosis in mice (86). Furthermore, an anti-FSTL1 blocking antibody protected against bleomycin-induced lung fibrosis, with the mechanism linked to a reduction in TGF- β -induced ECM production and myofibroblast differentiation (79). Haplodeletion of *Fstl1* in mice protected against bleomycin-induced epithelial injury and impaired profibrotic signaling responses, restoring the balance between TGF- β and BMP signaling. Thus, the emerging links among FSTL1, inflammation, adipogenesis, and fibrosis, together with our evidence of elevated FSTL1 expression in steatosis-associated liver fibrosis, suggest that this secreted factor may have a role in the progression of NAFLD.

Since advanced fibrosis and cirrhosis are linked to age, this is a confounding factor that complicates interpretation of most expression profiling studies of CLD patient samples. Thus, our RNA-sequencing findings, and other similar studies (33, 35, 36, 49–51), require independent validation studies for individual genes and/or proteins of interest. In this study, we have verified that levels of circulating VCAN (Figure 5G) and FSTL1 (Figure 7, D and E) were significantly higher in patients with advanced fibrosis compared

Table 3. List of steatosis-enriched late-stage fibrosis genes that encode extracellular proteins

Gene name	Cell surface/secreted
FSTL1	Secreted
VWF	Secreted
CRIM1	Cell surface
SPRY1	Cell surface
MAP1B	Cell surface
AQP1	Cell surface
FRAS1	Cell surface
ADAMTS1	Secreted
PLXDC2	Both
COL1A2	Secreted

with those with early-stage disease in independent, age-matched cohorts. These findings are consistent with the elevated hepatic mRNA levels of *VCAN* and *FSTL1* in patients with advanced fibrosis, as identified in our initial expression profiling analysis (Figure 5A and Figure 7A). Our data are therefore consistent with these two genes being predictive of etiology-independent and steatosis-enriched advanced fibrosis, respectively, rather than merely being associated with aging. The genetic and functional links that exist between many other genes that we identified and

liver fibrosis further support the notion that the gene sets identified in this study are bona fide indicators of CLD. In summary, in this study, we have used a stringent approach to identify hepatic gene expression profiles associated with advanced fibrosis in general as well as genes enriched in the setting of comorbid steatosis. These data sets should provide important resources for the pursuit of novel liver fibrosis targets and/or biomarkers, as exemplified by the validation of *VCAN* and *FSTL1* in this study.

Methods

Clinical information for patient cohorts. This is a retrospective, cross-sectional study in 69 patients with chronic HCV or FLD who had undergone a liver biopsy between 2000 and 2015 at the Princes Alexandra Hospital, Brisbane, Australia. Liver tissue and sera were collected at the time of liver biopsy following an overnight fast for 8–10 hours and stored at -80°C . Diagnosis of liver disease was based on standard biochemical and serological assays and consensus histological assessment of the liver biopsy by 2 blinded, experienced pathologists (ADC and GL) as previously reported (23). All patients with chronic HCV were HCV RNA positive at the time of liver biopsy and/or serum collection. A diagnosis of NAFLD required demonstration of hepatic steatosis in the presence of metabolic risk factors and the exclusion of substantial alcohol consumption (≥ 20 g/d) within the preceding 5 year, steatosis-inducing drugs, or CLD (including a prior history of alcohol-related liver disease). 78% of liver biopsies were ≥ 15 mm in length or showed definite cirrhosis. Fibrosis was assessed using a modified METAVIR score in these patients (87). Advanced fibrosis was defined as fibrosis stage 3 and 4. The METAVIR scoring system was used to assess hepatic inflammatory activity (87). Steatosis was diagnosed when it affected $>5\%$ of the parenchyma, and it was graded as mild (5%–33%, grade 1), moderate (34%–66%, grade 2), or severe ($>66\%$, grade 3). Clinical/demographic data (BMI, age, alcohol consumption, metabolic comorbidities) were obtained from the medical record. An ADVIA Centaur XP system (Siemens Healthcare Diagnostics) was used to quantify serum HA, PIIINP, and TIMP-1 according to manufacturer's instructions. The ELF score was autocalculated by the instrument (23). Subsequently, two independent patient cohorts were used for serum analyses: a cohort of 35 patients with chronic HCV or NAFLD and a cohort of 25 NAFLD/NASH patients. In these cohorts, fibrosis was assessed using a modified METAVIR score (87) in the cohort with chronic HCV or NAFLD and the Brunt score (88) for the NAFLD/NASH cohort.

Animal models. 4- to 5-week-old C57BL/6 female mice were treated with or without 300 mg/l TAA (MilliporeSigma) supplemented in drinking water for up to 12 weeks. For the CDE model, a diet deficient in choline was fed to 4- to 5-week-old C57BL/6 female mice together with drinking water supplemented with 0.15% ethionine (MilliporeSigma) for up to 3 weeks. Control mice were fed normal chow. Mouse livers were harvested at the endpoint of experiments for gene expression analyses.

Cell culture and gene silencing. The human HSC line LX-2 (provided Grant Ramm, Queensland Institute for Medical Research Berghofer, Brisbane, Queensland, Australia) was cultured in DMEM (Life Technologies) with 2% fetal bovine serum, 20 U/ml penicillin, and 20 mg/ml streptomycin. For gene-silencing experiments, EHF siRNAs (premade Stealth siRNA IDs: HSS120049, HSS178139, Life Technologies) and IRF6 siRNA (Stealth sense: CAGGCACCUAUACAGCCCUUCUAUA; antisense: UAUAGAAGGGCUGUAUAG-GUGCCUG; Life Technologies; used as a relevant control gene) were diluted in Opti-MEM reduced serum media (Life Technologies) to a final concentration of 200 pM, after which they were transfected into LX-2

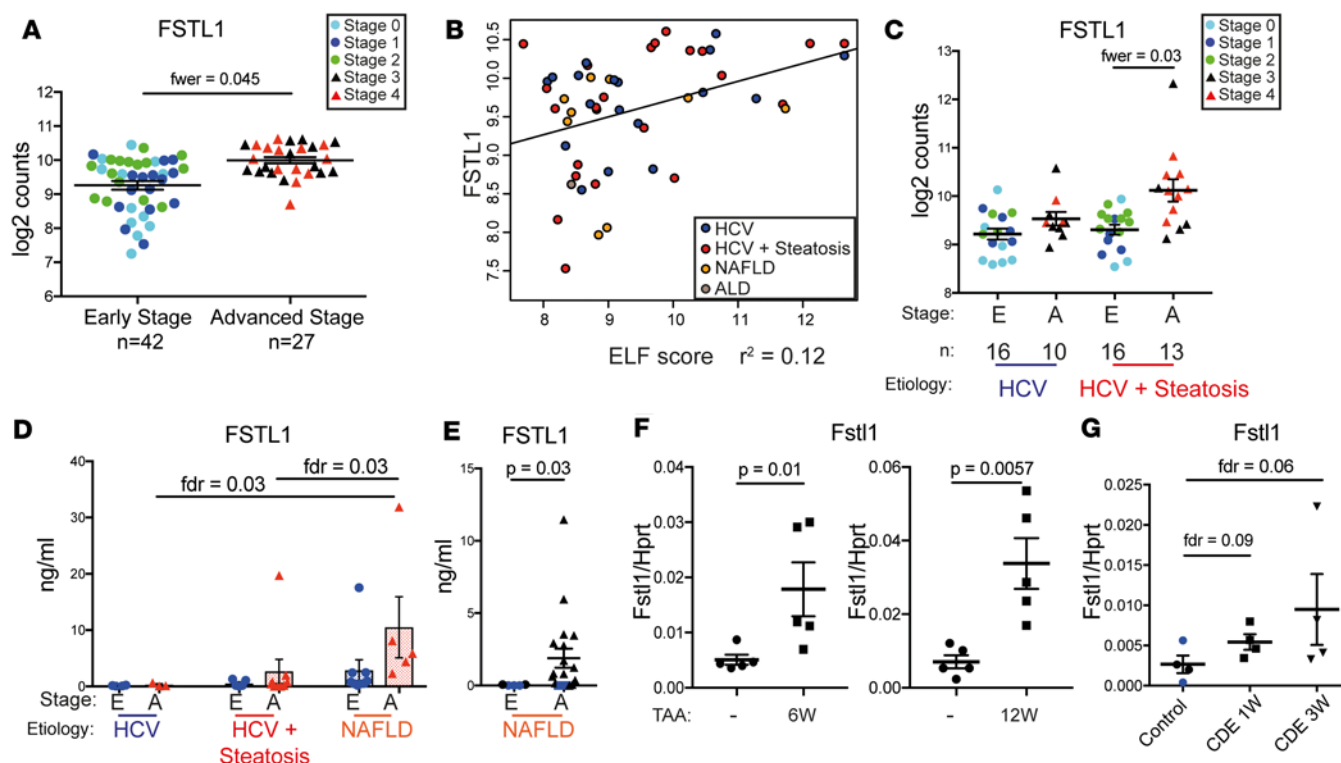


Figure 7. FSTL1 is elevated in patients with steatosis-associated late-stage fibrosis. (A) A scatter plot showing the log₂-corrected counts of *FSTL1* for all patients from the RNA-sequencing analysis ($n = 69$ patients; horizontal bars indicate mean \pm SEM; circles, early stages; triangles, advanced stages). (B) Correlation analysis showing *FSTL1* expression as the independent variable and ELF score as the dependent variable in HCV (blue), HCV with steatosis (red), NAFLD (orange), and alcoholic FLD (brown) patients ($n = 69$ patients). (C) *FSTL1* corrected counts were plotted in early (E) versus advanced (A) stages of fibrosis from patients with HCV alone or HCV and steatosis ($n = 55$ patients; horizontal bars indicate mean \pm SEM; circles, early stages; triangles, advanced stages). (D and E) Circulating levels of *FSTL1* in sera from (D) patients with chronic HCV or NAFLD (see Supplemental Table 1 for details of the cohort; $n = 35$ patients; detectable levels in 29 patients; circles, early stages; triangles, advanced stages) and (E) patients with NAFLD (see Supplemental Table 2 for details of the cohort; $n = 25$ patients; detectable levels in 18 patients; circles, early stages; triangles, advanced stages) were determined by ELISA (data represented as mean \pm SEM). (F and G) mRNA levels of *Fstl1* were measured, relative to the housekeeping gene *Hprt*, in (F) 6- and 12-week-old TAA- or sham-treated mice ($n = 5$ mice per group) and (G) 1- and 3-week-old CDE- or sham-treated mice ($n = 4$ mice per group) (data represented as mean \pm SEM). For A and C, FWER values were calculated using Bonferroni multiple-testing corrections; for D and G, FDR values were calculated using nonparametric ANOVA (Kruskal-Wallis test) followed by Benjamini-Hochberg multiple-testing corrections; and for E and F, P values were calculated using the Mann-Whitney test. P or FDR ≤ 0.05 was considered statistically significant.

cells using Lipofectamine 2000 reagent (2.5 μ l/transfection, Life Technologies) by the reverse transfection method for 24 hours, as described by the manufacturer. Cells were then stimulated with 10 ng/ml TGF- β (R&D Systems) for 24 hours or were left untreated, after which RNA was extracted and qPCR was performed.

RNA purification and qPCR. Total RNA was isolated from human liver biopsy tissue, mouse liver, or LX-2 cells using TRI Reagent (MilliporeSigma), as per the manufacturer's instructions. Up to 1 μ g of the extracted RNA was used to generate cDNA using random hexamer (for human samples) or oligodT (for mouse samples) priming and SuperScript III reverse transcriptase (Invitrogen). The synthesized cDNA was diluted 10-fold in RNase-free water and used for qPCR analysis. Gene expression, relative to the control gene 18S ribosomal RNA (*18S*) (for human biopsy samples) or hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) (for human and mouse samples), was determined by qPCR using the Δ Ct method, SYBR Green PCR Master Mix (Applied Biosystems), and ViiA7 RT-PCR systems (Applied Biosystems). Primers used to amplify specific genes are listed in Supplemental Table 3.

RNA sequencing and analyses. RNA sequencing was performed at the IMB Sequencing Facility at The University of Queensland. RNA quality was assessed using the Agilent 2100 BioAnalyzer, and samples with an RNA integrity number ≥ 6.0 were included for RNA sequencing. Up to 1 μ g RNA was used to synthesize mRNA libraries using the TruSeq Stranded Total RNA (Ribo-Zero GOLD) library preparation kit (Illumina), as per the manufacturer's instructions. Sequencing was performed using NextSeq 75 cycle (1 \times 75 bp) high-output run (Illumina) on a NextSeq500 sequencing system (Illumina). Sequence reads were mapped to

the human hg19 reference genome using STAR aligner v2.3.0e (89), with 12.79–19.5 million reads per sample, and the Sequence Alignment/Map output files were converted into Binary Alignment/Map format using Samtools v0.1.19 (90). Transcript read counts were obtained using the HTSeq python package (91), following which the CPM value was obtained using the edgeR package (92). Genes that passed a threshold of 10 CPM or more in all samples (high stringency), 10 CPM or more in at least 20 patient samples (medium level stringency), or 10 CPM or more in at least 1 patient sample (relaxed stringency) were retained. Most subsequent analyses focused on the medium level stringency gene set. Filtered genes were subjected to normalization by the trimmed means of M value method (Supplemental Figure 1) (93). Normalization was performed after filtering the data, as has been previously recommended for analysis of RNA-sequencing data (94). Correlation analysis was performed on the clinical parameters (sex, ethnicity, etiology, smoking habit, technical factors, alcohol consumption, BMI, age, fibrosis stage, and ELF score) for all patients to determine confounding factors. Factors that correlated more than 25%, such as age, were considered strong confounders and were not corrected for in subsequent analyses. Genes were then subjected to a batch-effect correction based on a generalized linear model and analyzed for differential gene expression by likelihood-ratio test using the edgeR package in R (version 3.3.2)/Bioconductor (95). Differential gene expression analyses were performed by comparing early versus advanced stages of fibrosis across different etiologies and in HCV patients with or without steatosis. Multiple testing was performed using Bonferroni corrections (for family-wise error rate-adjusted [FWER-adjusted] or study-wide significant gene list) or Benjamini-Hochberg corrections (for a less stringent, FDR-adjusted gene list). FWER or FDR values of 0.05 or less were considered significantly different. For ELF score correlation analysis, a linear model was generated using the corrected counts for all genes (y variable) and ELF score (x variable) for each patient. The genes strongly fitting the model were then subjected to multiple testing corrections using the Bonferroni approach, as described above.

Pathway analysis, enrichment analysis, and meta-analysis. Pathway analysis was performed using the Gene-Go MetaCore pathway analysis tool (Thomas Reuters) and KEGG pathway analysis via ClusterProfiler package in R (96). For extracellular protein analysis, GO analysis via ClusterProfiler package in R (96) and MetaCore (GO localization) tools were used. For transcription factor-target enrichment analysis, a hypergeometric test was performed on the ELF score-correlated gene set and 1,000 randomized gene sets using the phyper function in R (97). Transcriptional factor motif enrichment analysis was performed on the ELF score-correlated gene set using the RcisTarget package in R (32). For meta-analysis, lists of DEGs were taken from published data sets (33–35, 98). The Vennerable package (<https://github.com/js229/Vennerable>, commit 46057c9) was used to create the Venn diagram comparing all the data sets.

ELISA. Serum was isolated from patient blood, and VCAN and FSTL1 levels were quantified using a human VCAN ELISA kit (CUSABIO, CSB-E11884h) and human FSTL1 ELISA kit (R&D Systems, RDS-DY1694), respectively, according to the manufacturer's protocol.

Statistics. Statistical analyses of RNA-sequencing data are described above. For other experiments, data were combined from 3 or more independent experiments and statistical analysis was performed using RStudio (version 0.99.442) or GraphPad Prism (version 7) software. Statistical tests used for individual experiments include Bonferroni multiple-testing corrections, nonparametric ANOVA (Kruskal-Wallis test) followed by Benjamini-Hochberg multiple-testing corrections, Mann-Whitney test (2-tailed), 2-way ANOVA followed by Benjamini-Hochberg multiple-testing corrections, χ^2 test, and Welch's t test (2-tailed) and are indicated in individual figure legends. P values less than or equal to 0.05 were considered as statistically significant.

Study approval. Informed written consent was obtained from each patient, and the protocol was approved by the Metro South Health and the University of Queensland Human Research Ethics Committees (ethics reference no. HREC/99/QPAH76 and 2003000092). All animal work was carried out with the approval of the University of Queensland Animal Ethics Committee (animal ethics clearance no. IMB/509/15/UQ and IMB/364/15/CIDR).

Data availability. The raw and processed data from this study have been submitted to ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>) under accession E-MTAB-6863.

Author contributions

DR designed and performed experiments; acquired, analyzed, and interpreted the data; and drafted the manuscript. KMI contributed to study design and data interpretation and critically revised the manuscript. SWL contributed to data analyses and interpretation and critically revised the manuscript. LUH collected clinical samples and information from patients. ZL performed the animal model studies. ADC provided clinical

cal expertise and histological assessment of patient samples and critically revised the manuscript. PJP and KJF collected clinical samples and provided assistance with clinical information. AI assisted with the animal model studies. GL provided histological assessment of patient samples. JLS and KS contributed to study design, supervision, and data interpretation. DPF contributed to study design, supervision of the animal model experiments, and data interpretation. JEP contributed to study design and data interpretation, provided expertise in statistical analyses, and critically revised the manuscript. EEP designed and supervised the study, collected clinical samples and data, and drafted and critically revised the manuscript. MJS designed and supervised the study, interpreted the data, and drafted and critically revised the manuscript.

Acknowledgments

This work was supported by the IMB Centre for Inflammation and Disease Research at The University of Queensland as well as CSL Limited. The latter had no role in study design or data analyses. MJS is supported by a National Health and Medical Research Council (NHMRC) Senior Research Fellowship (APP1107914), DPF is supported by an NHMRC Senior Principal Research Fellowship (APP1117017), and KS is supported by an NHMRC Career Development Fellowship (APP1141131). We thank Siemens Healthineers for providing the ELF test.

Address correspondence to: Matthew Sweet, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Qld 4072, Australia. Phone: 61.7.3346.2082; Email: m.sweet@imb.uq.edu.au. Or to: Elizabeth Powell, Centre for Liver Disease Research, The University of Queensland, Translational Research Institute, Brisbane, Qld 4102, Australia. Phone: 61.7.3443.8015; Email: e.powell@uq.edu.au.

1. Polaris Observatory HCV Collaborators. Global prevalence and genotype distribution of hepatitis C virus infection in 2015: a modelling study. *Lancet Gastroenterol Hepatol*. 2017;2(3):161–176.
2. van der Meer AJ, Berenguer M. Reversion of disease manifestations after HCV eradication. *J Hepatol*. 2016;65(1 Suppl):S95–S108.
3. Nahon P, et al. Eradication of hepatitis C virus infection in patients with cirrhosis reduces risk of liver and non-liver complications. *Gastroenterology*. 2017;152(1):142–156.e2.
4. Desmet VJ, Roskams T. Reversal of cirrhosis: evidence-based medicine? *Gastroenterology*. 2003;125(2):629–30; author reply 630.
5. Kwok RM, Tran TT. Management of cirrhotic patients after successful HCV eradication. *Curr Treat Options Gastroenterol*. 2017;15(2):305–315.
6. Jacobson IM, Lim JK, Fried MW. American Gastroenterological Association Institute Clinical Practice Update-Expert Review: Care of patients who have achieved a sustained virologic response after antiviral therapy for chronic hepatitis C infection. *Gastroenterology*. 2017;152(6):1578–1587.
7. Gordon SC, et al. Prevalence of cirrhosis in hepatitis C patients in the Chronic Hepatitis Cohort Study (CHeCS): a retrospective and prospective observational study. *Am J Gastroenterol*. 2015;110(8):1169–1178.
8. Adinolfi LE, et al. NAFLD and NASH in HCV Infection: prevalence and significance in hepatic and extrahepatic manifestations. *Int J Mol Sci*. 2016;17(6):803.
9. Hourigan LF, et al. Fibrosis in chronic hepatitis C correlates significantly with body mass index and steatosis. *Hepatology*. 1999;29(4):1215–1219.
10. Hung CH, et al. Impact of steatosis on long-term histological outcome in chronic hepatitis C after antiviral therapy. *Antivir Ther (Lond)*. 2006;11(4):483–489.
11. Pekow JR, Bhan AK, Zheng H, Chung RT. Hepatic steatosis is associated with increased frequency of hepatocellular carcinoma in patients with hepatitis C-related cirrhosis. *Cancer*. 2007;109(12):2490–2496.
12. Ohata K, et al. Hepatic steatosis is a risk factor for hepatocellular carcinoma in patients with chronic hepatitis C virus infection. *Cancer*. 2003;97(12):3036–3043.
13. Kumar D, Farrell GC, Kench J, George J. Hepatic steatosis and the risk of hepatocellular carcinoma in chronic hepatitis C. *J Gastroenterol Hepatol*. 2005;20(9):1395–1400.
14. Nirei K, et al. Steatosis influences the clinical profiles and long-term outcomes of interferon-treated chronic hepatitis C and liver cirrhosis patients. *Int J Med Sci*. 2017;14(1):45–52.
15. Tanaka A, et al. Hepatic steatosis as a possible risk factor for the development of hepatocellular carcinoma after eradication of hepatitis C virus with antiviral therapy in patients with chronic hepatitis C. *World J Gastroenterol*. 2007;13(39):5180–5187.
16. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology*. 2016;64(1):73–84.
17. Welsch C, et al. Ongoing liver inflammation in patients with chronic hepatitis C and sustained virological response. *PLoS ONE*. 2017;12(2):e0171755.
18. Issa D, Wattacheril J, Sanyal AJ. Treatment options for nonalcoholic steatohepatitis - a safety evaluation. *Expert Opin Drug Saf*. 2017;16(8):903–913.
19. Rosenberg WM, et al. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology*. 2004;127(6):1704–1713.
20. Irvine KM, et al. The Enhanced liver fibrosis score is associated with clinical outcomes and disease progression in patients with chronic liver disease. *Liver Int*. 2016;36(3):370–377.
21. Jiao J, Friedman SL, Aloman C. Hepatic fibrosis. *Curr Opin Gastroenterol*. 2009;25(3):223–229.

22. Miao CG, et al. Wnt signaling in liver fibrosis: progress, challenges and potential directions. *Biochimie*. 2013;95(12):2326–2335.
23. Fagan KJ, et al. ELF score ≥ 9.8 indicates advanced hepatic fibrosis and is influenced by age, steatosis and histological activity. *Liver Int*. 2015;35(6):1673–1681.
24. Martin CR, et al. CFTR dysfunction predisposes to fibrotic liver disease in a murine model. *Am J Physiol Gastrointest Liver Physiol*. 2012;303(4):G474–G481.
25. Coombes JD, et al. Osteopontin neutralisation abrogates the liver progenitor cell response and fibrogenesis in mice. *Gut*. 2015;64(7):1120–1131.
26. Wang X, et al. Osteopontin induces ductular reaction contributing to liver fibrosis. *Gut*. 2014;63(11):1805–1818.
27. Seth D, Duly A, Kuo PC, McCaughan GW, Haber PS. Osteopontin is an important mediator of alcoholic liver disease via hepatic stellate cell activation. *World J Gastroenterol*. 2014;20(36):13088–13104.
28. Galang CK, Muller WJ, Foos G, Oshima RG, Hauser CA. Changes in the expression of many Ets family transcription factors and of potential target genes in normal mammary tissue and tumors. *J Biol Chem*. 2004;279(12):11281–11292.
29. Gonzalez HE, et al. Identification of 9 genes differentially expressed in head and neck squamous cell carcinoma. *Arch Otolaryngol Head Neck Surg*. 2003;129(7):754–759.
30. Mutolo MJ, Leir SH, Fossum SL, Browne JA, Harris A. A transcription factor network represses CFTR gene expression in airway epithelial cells. *Biochem J*. 2018;475(7):1323–1334.
31. Fossum SL, et al. Ets homologous factor regulates pathways controlling response to injury in airway epithelial cells. *Nucleic Acids Res*. 2014;42(22):13588–13598.
32. Aibar S, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. 2017;14(11):1083–1086.
33. 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*. 2017;47(13):1445–1458.
34. Arendt BM, 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*. 2015;61(5):1565–1578.
35. Lou Y, et al. Characterization of transcriptional modules related to fibrosing-NAFLD progression. *Sci Rep*. 2017;7(1):4748.
36. Lefebvre P, et al. Interspecies NASH disease activity whole-genome profiling identifies a fibrogenic role of PPAR α -regulated dermatopontin. *JCI Insight*. 2017;2(13):92264.
37. Bukong TN, Maurice SB, Chahal B, Schaeffer DF, Winwood PJ. Versican: a novel modulator of hepatic fibrosis. *Lab Invest*. 2016;96(3):361–374.
38. Kim S, et al. Carcinoma-produced factors activate myeloid cells through TLR2 to stimulate metastasis. *Nature*. 2009;457(7225):102–106.
39. Wight TN, Kinsella MG, Evanko SP, Potter-Perigo S, Merrilees MJ. Versican and the regulation of cell phenotype in disease. *Biochim Biophys Acta*. 2014;1840(8):2441–2451.
40. Chaly Y, Hostager B, Smith S, Hirsch R. Follistatin-like protein 1 and its role in inflammation and inflammatory diseases. *Immunol Res*. 2014;59(1-3):266–272.
41. Liu Y, et al. Follistatin-like protein 1 promotes inflammatory reactions in nucleus pulposus cells by interacting with the MAPK and NF κ B signaling pathways. *Oncotarget*. 2017;8(26):43023–43034.
42. Kim HJ, Kang WY, Seong SJ, Kim SY, Lim MS, Yoon YR. Follistatin-like 1 promotes osteoclast formation via RANKL-mediated NF- κ B activation and M-CSF-induced precursor proliferation. *Cell Signal*. 2016;28(9):1137–1144.
43. Ni S, et al. The involvement of follistatin-like protein 1 in osteoarthritis by elevating NF- κ B-mediated inflammatory cytokines and enhancing fibroblast like synovocyte proliferation. *Arthritis Res Ther*. 2015;17:91.
44. Fan N, et al. Follistatin-like 1: a potential mediator of inflammation in obesity. *Mediators Inflamm*. 2013;2013:752519.
45. Li D, et al. Follistatin-like protein 1 is elevated in systemic autoimmune diseases and correlated with disease activity in patients with rheumatoid arthritis. *Arthritis Res Ther*. 2011;13(1):R17.
46. Clutter SD, Wilson DC, Marinov AD, Hirsch R. Follistatin-like protein 1 promotes arthritis by up-regulating IFN- γ . *J Immunol*. 2009;182(1):234–239.
47. Miyamae T, et al. Follistatin-like protein-1 is a novel proinflammatory molecule. *J Immunol*. 2006;177(7):4758–4762.
48. Chaly Y, et al. Follistatin-like protein 1 enhances NLRP3 inflammasome-mediated IL-1 β secretion from monocytes and macrophages. *Eur J Immunol*. 2014;44(5):1467–1479.
49. Smalling RL, et al. Genome-wide transcriptome analysis identifies novel gene signatures implicated in human chronic liver disease. *Am J Physiol Gastrointest Liver Physiol*. 2013;305(5):G364–G374.
50. Leti F, et al. High-throughput sequencing reveals altered expression of hepatic microRNAs in nonalcoholic fatty liver disease-related fibrosis. *Transl Res*. 2015;166(3):304–314.
51. Robinson MW, et al. Viral genotype correlates with distinct liver gene transcription signatures in chronic hepatitis C virus infection. *Liver Int*. 2015;35(10):2256–2264.
52. Yang W, et al. Integrative transcriptomic analysis of NAFLD animal model reveals dysregulated genes and pathways in metabolism. *Gene*. 2016;595(1):99–108.
53. Leeuwen L, Fitzgerald DA, Gaskin KJ. Liver disease in cystic fibrosis. *Paediatr Respir Rev*. 2014;15(1):69–74.
54. Bataller R, Brenner DA. Liver fibrosis. *J Clin Invest*. 2005;115(2):209–218.
55. Williams MJ, Clouston AD, Forbes SJ. Links between hepatic fibrosis, ductular reaction, and progenitor cell expansion. *Gastroenterology*. 2014;146(2):349–356.
56. Gadd VL, et al. Portal, but not lobular, macrophages express matrix metalloproteinase-9: association with the ductular reaction and fibrosis in chronic hepatitis C. *Liver Int*. 2013;33(4):569–579.
57. Wang M, et al. Characterization of gene expression profiles in HBV-related liver fibrosis patients and identification of ITGBL1 as a key regulator of fibrogenesis. *Sci Rep*. 2017;7:43446.
58. Moylan CA, et al. Hepatic gene expression profiles differentiate presymptomatic patients with mild versus severe nonalcoholic fatty liver disease. *Hepatology*. 2014;59(2):471–482.
59. Xu MY, Qu Y, Li Z, Li F, Xiao CY, Lu LG. A 6 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis B. *Front Biosci (Landmark Ed)*. 2016;21:479–486.

60. Tugores A, et al. The epithelium-specific ETS protein EHF/ESE-3 is a context-dependent transcriptional repressor downstream of MAPK signaling cascades. *J Biol Chem*. 2001;276(23):20397–20406.
61. Sharrocks AD. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol*. 2001;2(11):827–837.
62. Kar A, Gutierrez-Hartmann A. Molecular mechanisms of ETS transcription factor-mediated tumorigenesis. *Crit Rev Biochem Mol Biol*. 2013;48(6):522–543.
63. Lim JH, Cho JY, Park YB, Park JW, Kwon TK. ESE-3 transcription factor is involved in the expression of death receptor (DR)-5 through putative Ets sites. *Biochem Biophys Res Commun*. 2006;350(3):736–741.
64. Stolzenburg LR, et al. Regulatory dynamics of 11p13 suggest a role for EHF in modifying CF lung disease severity. *Nucleic Acids Res*. 2017;45(15):8773–8784.
65. Corvol H, et al. Genome-wide association meta-analysis identifies five modifier loci of lung disease severity in cystic fibrosis. *Nat Commun*. 2015;6:8382.
66. Wright FA, et al. Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2. *Nat Genet*. 2011;43(6):539–546.
67. Cattaruzza S, et al. Distribution of PG-M/versican variants in human tissues and de novo expression of isoform V3 upon endothelial cell activation, migration, and neoangiogenesis in vitro. *J Biol Chem*. 2002;277(49):47626–47635.
68. Zimmermann DR, Dours-Zimmermann MT, Schubert M, Bruckner-Tuderman L. Versican is expressed in the proliferating zone in the epidermis and in association with the elastic network of the dermis. *J Cell Biol*. 1994;124(5):817–825.
69. Schmalfeldt M, Dours-Zimmermann MT, Winterhalter KH, Zimmermann DR. Versican V2 is a major extracellular matrix component of the mature bovine brain. *J Biol Chem*. 1998;273(25):15758–15764.
70. Kang I, Yoon DW, Braun KR, Wight TN. Expression of versican V3 by arterial smooth muscle cells alters tumor growth factor β (TGF β)-, epidermal growth factor (EGF)-, and nuclear factor κ B (NF κ B)-dependent signaling pathways, creating a microenvironment that resists monocyte adhesion. *J Biol Chem*. 2014;289(22):15393–15404.
71. Touab M, Villena J, Barranco C, Arumí-Uría M, Bassols A. Versican is differentially expressed in human melanoma and may play a role in tumor development. *Am J Pathol*. 2002;160(2):549–557.
72. Arslan F, Bosserhoff AK, Nickl-Jockschat T, Doerfler A, Bogdahn U, Hau P. The role of versican isoforms V0/V1 in glioma migration mediated by transforming growth factor-beta2. *Br J Cancer*. 2007;96(10):1560–1568.
73. Sheng W, et al. The roles of versican V1 and V2 isoforms in cell proliferation and apoptosis. *Mol Biol Cell*. 2005;16(3):1330–1340.
74. Wight TN, Kang I, Merrilees MJ. Versican and the control of inflammation. *Matrix Biol*. 2014;35:152–161.
75. Kesteloot F, et al. ADAM metalloproteinase with thrombospondin type 1 motif 2 inactivation reduces the extent and stability of carbon tetrachloride-induced hepatic fibrosis in mice. *Hepatology*. 2007;46(5):1620–1631.
76. Bauters D, et al. ADAMTS5 deficiency protects against non-alcoholic steatohepatitis in obesity. *Liver Int*. 2016;36(12):1848–1859.
77. Wu YJ, La Pierre DP, Wu J, Yee AJ, Yang BB. The interaction of versican with its binding partners. *Cell Res*. 2005;15(7):483–494.
78. Ganz M, Szabo G. Immune and inflammatory pathways in NASH. *Hepatol Int*. 2013;7 Suppl 2:771–781.
79. Dong Y, et al. Blocking follistatin-like 1 attenuates bleomycin-induced pulmonary fibrosis in mice. *J Exp Med*. 2015;212(2):235–252.
80. Campfield BT, et al. Follistatin-like protein 1 modulates IL-17 signaling via IL-17RC regulation in stromal cells. *Immunol Cell Biol*. 2017;95(8):656–665.
81. Wree A, et al. NLRP3 inflammasome activation results in hepatocyte pyroptosis, liver inflammation, and fibrosis in mice. *Hepatology*. 2014;59(3):898–910.
82. Negash AA, et al. IL-1 β production through the NLRP3 inflammasome by hepatic macrophages links hepatitis C virus infection with liver inflammation and disease. *PLoS Pathog*. 2013;9(4):e1003330.
83. Giles DA, Moreno-Fernandez ME, Divanovic S. IL-17 axis driven inflammation in non-alcoholic fatty liver disease progression. *Curr Drug Targets*. 2015;16(12):1315–1323.
84. Du WJ, et al. Expression of interleukin-17 associated with disease progression and liver fibrosis with hepatitis B virus infection: IL-17 in HBV infection. *Diagn Pathol*. 2013;8:40.
85. Meng F, et al. Interleukin-17 signaling in inflammatory, Kupffer cells, and hepatic stellate cells exacerbates liver fibrosis in mice. *Gastroenterology*. 2012;143(3):765–776.e3.
86. Vollmann EH, et al. Identification of novel fibrosis modifiers by in vivo siRNA silencing. *Mol Ther Nucleic Acids*. 2017;7:314–323.
87. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology*. 1996;24(2):289–293.
88. Kleiner DE, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology*. 2005;41(6):1313–1321.
89. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
90. Li H, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–2079.
91. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–169.
92. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15(2):R29.
93. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25.
94. Conesa A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol*. 2016;17:13.
95. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–140.
96. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–287.
97. Kachitvichyanukul V, Schmeiser B. Computer generation of hypergeometric random variates. *J Stat Comput Simul*. 1985;22(2):127–145.
98. Ryaboshapkina M, Hammar M. Human hepatic gene expression signature of non-alcoholic fatty liver disease progression, a meta-analysis. *Sci Rep*. 2017;7(1):12361.