Published in final edited form as:

Gut. 2016 October ; 65(10): 1754–1764. doi:10.1136/gutjnl-2015-309655.

A hepatic stellate cell gene expression signature associated with outcomes in hepatitis C cirrhosis and hepatocellular carcinoma after curative resection

David Y. Zhang¹, Nicolas Goossens^{2,3}, Jinsheng Guo^{1,4}, Ming-chao Tsai¹, Hsin-I Chou¹, Civan Altunkaynak¹, Angelo Sangiovanni⁵, Massimo Ivarone⁵, Massomo Colombo⁵, Masahiro Kobayashi⁶, Hiromitsu Kumada⁶, Augusto Villanueva^{1,2}, Josep M. Llovet^{1,2,7,8}, Yujin Hoshida^{1,2}, and Scott L. Friedman^{1,2}

¹Department of Medicine, Division of Liver Diseases Icahn School of Medicine at Mount Sinai, New York, NY ²Liver Cancer Program, Tisch Cancer Institute Icahn School of Medicine at Mount Sinai, New York, New York ³Division of Gastroenterology and Hepatology Geneva University Hospital, Geneva, Switzerland ⁴Division of Digestive Diseases Zhongshang Hospital and Fudan University, Shanghai, China ⁵M. & A. Migliavacca Center for Liver Disease and 1st Division of Gastroenterology Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy ⁶Toranomon Hospital, Tokyo ⁷Liver Cancer Translational Research Lab, Barcelona Clinic Liver Cancer (BCLC) Group, Liver Unit, Hospital Clínic Barcelona, IDIBAPS, Centro de Investigaciones en Red de Enfermedades Hepáticas y Digestivas (CIBERehd) University of Barcelona, Barcelona, Spain ⁸Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

Abstract

Objective—We used an informatics approach to identify and validate genes whose expression is unique to hepatic stellate cells, and assessed the prognostic capability of their expression in cirrhosis.

Design—We defined a hepatic stellate cell gene signature by comparing stellate, immune, and hepatic transcriptome profiles. We then created a prognostic index using a combination of hepatic stellate cell signature expression and clinical variables, using overall survival as the primary clinical outcome. This signature was derived in a retrospective-prospective cohort of hepatitis C-related early-stage cirrhosis (prognostic index derivation set), and validated in an independent retrospective cohort of post-resection HCC patients (n=82, prognostic index validation set). We then examined association between hepatic stellate cell signature expression and decompensation, hepatocellular carcinoma (HCC) incidence, and progression of Child-Pugh class as additional outcomes in the prognostic index derivation set, and HCC recurrence as an additional outcome in the validation set. We tested whether hepatic stellate cell signature expression is predictive of

Corresponding Author: Scott L. Friedman MD, Division of Liver Diseases, Icahn School of Medicine at Mount Sinai, 1425 Madison Ave, Box 1123, New York, NY 10029, scott.friedman@mssm.edu, Fax 212-849-2574.

COMPETING INTERESTS

No competing interests.

death, decompensation, progression of Child-Pugh class, and HCC incidence in a retrospective-prospective cohort of hepatitis C cirrhotic patients. In the prognostic index derivation cohort (n=216), 66 (31%) died, 71 (34%) developed hepatic decompensation, 66 (31%) experienced progression of Child-Pugh class, and 65 (30%) developed hepatocellular carcinoma. In the prognostic index validation cohort (n=82), 32 (39%) died and 65 (79%) developed tumor recurrence.

Results—The 122-gene hepatic stellate cell signature consists of genes encoding extracellular matrix proteins and developmental factors, and correlates with the extent of fibrosis in human, mouse, and rat datasets. The hepatic stellate cell signature contains several cell surface genes previously established as stellate cell-specific, as well as PCDH7, a novel protocadherin stellate cell surface marker. Importantly, association of clinical prognostic variables with overall survival (c-index: 0.66, 95% CI: 0.59–0.74) was improved by adding the hepatic stellate cell signature (c-index: 0.70, 95%CI: 0.62–0.78); we used these results to define a prognostic index in the derivation set. In the validation set, the same prognostic index was associated with overall survival (c-index: 0.62, 95%CI: 0.51–0.72). For other clinical outcomes examined, the prognostic index was associated with decompensation (c-index: 0.62, 95%CI: 0.55–0.69), HCC (c-index: 0.63, 95%CI: 0.56–0.71), and progression of Child-Pugh class (c-index: 0.70, 95%CI: 0.63–0.78) in the derivation set, and HCC recurrence (c-index: 0.54, 95%CI: 0.46–0.62) in the validation set.

Conclusion—This work highlights the unique transcriptional niche of stellate cells, and identifies potential stellate cell targets for tracking, targeting, and isolation. Hepatic stellate cell signature expression may identify HCV cirrhosis or post resection HCC patients with poor prognosis.

Keywords

cirrhosis; hepatic decompensation; gene set enrichment analysis; hepatocellular carcinoma; transcriptome

INTRODUCTION

Hepatic fibrosis is characterized by progressive deposition of extracellular matrix in patients with chronic liver injury. Among patients with hepatic fibrosis, a significant fraction will progress to cirrhosis, with eventual loss of liver function and an increased risk of hepatocellular carcinoma.[1, 2] Hepatic stellate cells are the primary cellular mediators of hepatic fibrosis through their transdifferentiation, or activation, from a pericytic, vitamin A-storing cell to a contractile, matrix-producing myofibroblast in response to liver injury and inflammation,[3, 4] and specific abrogation of this response has been validated as an anti-fibrotic therapy in many experimental models.[5] In view of their central role in fibrosis and cirrhosis, the identification of specific hepatic stellate cell markers for use as non-invasive diagnostic markers could greatly facilitate preclinical and clinical development of anti-fibrotic therapies in patients with liver disease.

There are a growing number of gene expression datasets available that can facilitate the identification of uniquely or differentially expressed genes across multiple tissues and array formats, especially in liver.[6] In the fields of fibrosis and liver disease, increasingly

sophisticated genomic approaches are being employed; these studies have combined multiple sources of data to identify hepatocellular carcinoma biomarkers,[7] explore regulators of collagen deposition,[8] classify hepatocellular carcinoma into distinct subclasses,[9] and establish prognostic gene expression signatures.[10, 11]

Several genomic approaches have also been employed to elucidate stellate cell biology. Subtraction cloning[12] and microarray[13] have identified differentially expressed transcripts during hepatic stellate cell activation and following their inactivation,[14, 15] for example. While these studies have provided valuable insight into stellate cell expression changes in response to injury and regression, no studies have compared stellate cell gene expression to the gene expression profiles of other cell types in the liver, or between fibrotic tissues in different organs.

To identify novel stellate cell surface or prognostic markers, we leveraged the rapidly expanding availability of cell and tissue expression profiles to identify transcripts specifically expressed in stellate cells. We have generated a hepatic stellate cell gene expression signature that correlates with progressive liver disease in patients and animal models, and have used this signature to identify novel cell surface markers of stellate cells. Importantly, this expression signature correlates with patient prognosis in chronic liver disease, further validating both the hepatic stellate cell signature and the role of hepatic stellate cells in chronic liver disease. We have also used this signature to determine to which tissue expression profiles are similar between hepatic, renal and pulmonary fibrosis. We anticipate that the stellate cell gene signature will be informative for future stellate cell targeting and isolation. Additionally, this signature should facilitate biomarker development, enhancing our ability to assess fibrosis stage and response to therapies.

EXPERIMENTAL PROCEDURES

Identification of the hepatic stellate cell signature

To define the stellate cell gene expression signature, 98 platform-matched transcriptome datasets covering 17 liver cell and tissue types were obtained from previously published work[13] and the National Center for Biotechnology Information Gene Expression Omnibus (GEO) (www.ncbi.nlm.nih.gov/geo) (Table S1, Figure S1). Raw scan data were obtained and converted to normalized data using the Robust Multi-array Average algorithm[16] and quantile normalization, implemented in the GenePattern genomic analysis toolkit (www.broadinstitute.org/genepattern/)[17] ExpressionFileCreator module. Multiple probes corresponding to a gene were collapsed into a median value and subsequently labeled with an official gene symbol provided by the NCBI Entrez Gene database. Genes were included in the hepatic stellate cell signature if their median expression in either quiescent or activated stellate cells was at least twice as high as the highest expression in all other cell and tissue types. Mouse and rat genes were mapped to orthologous human genes based on the NCBI HomoloGene database (release 68), and rat and mouse genes with no known human homologous relationships were excluded from subsequent analysis. Unsupervised hierarchical clustering (Pearson correlation, average-linkage method) was performed on the discovery set using the GenePattern HierarchicalClustering module. Validation datasets, generated on microarrays from a range of species and platforms, were also obtained from the

GEO database. Prenormalized gene expression measurements from each study were used for the analysis summarized in Figure S1.

Cell surface marker discovery

Candidate hepatic stellate cell surface markers were selected from hepatic stellate cell signature genes by integrating two independent databases of protein subcellular localization: Gene Ontology cellular component (geneontology.org/page/cellular-component-ontology-guidelines), and the Mammalian Protein Localization Database (LOCATE, locate.imb.uq.edu.au), [18] with use of terms summarized in Table S4.

Molecular pathway analysis

Association of the hepatic stellate cell signature with gene ontology terms (geneontology.org) was assessed using the hypergeometric test implemented in the DAVID (v6.7) suite of functional annotation tools (david.abcc.ncifcrf.gov). [19, 20] Bonferroni-corrected p-values less than 0.05 were regarded as a statistically significant association. Gene Set Enrichment Analysis (GSEA) was performed as previously published. [21, 22] Gene interaction networks were generated by submitting unranked leading edge gene subsets to the Ingenuity Pathway Analysis core analysis (www.ingenuity.com).

Study population

Association of hepatic stellate cell signature expression with overall survival (primary clinical outcome) was examined in two cohorts with early cirrhosis and minimal hepatic comorbidities: a retrospective-prospective cohort of patients with HCV-related, early-stage (Child-Pugh class A) cirrhosis [10] (prognostic index derivation set), and an independent retrospective cohort of consecutively enrolled patients with curatively resected early-stage HCC followed for a median of 7.8 years [11] (prognostic index validation set). For the prognostic index derivation set, patients were included if they were diagnosed with histologically confirmed liver cirrhosis, but excluded if they had a history of hepatic decompensation, hepatocellular carcinoma, or death. For the prognostic index validation set, patients were included if they were treated with surgery for primary hepatocellular carcinoma between 1990 and 2001, but excluded if no outcomes data was available, or if no formalin-fixed, paraffin-embedded tumor and adjacent tissue were available. Additional details regarding patient enrollment, diagnosis, follow-up and treatment protocols can be found in their respective publications.

Association of the stellate cell signature with development of hepatic decompensation (bleeding varices, ascites, hepatic encephalopathy, hepatorenal syndrome, and infection), progression of Child-Pugh class, and HCC development were evaluated in the discovery set, and association with HCC recurrence was evaluated in the validation set. In the validation set, recurrent HCC tumors were confirmed to be clonally independent from resected primary HCC tumors: poor prognosis in this cohort is attributable to underlying liver cirrhosis, not to dissemination or metastasis of resected primary tumors.

Survival analysis

For each patient, relative over-expression of the hepatic stellate cell signature was measured by a gene set enrichment score based on the Kolmogorov-Smirnov (KS) statistic, and patients in each cohort were stratified into hepatic stellate cell signature positive and negative groups based on a 75th percentile cut-off value. Association of the groups with clinical outcomes was evaluated by Kaplan-Meier curve, log-rank test, univariable Cox regression modeling, and multivariable Cox regression modeling. To assess whether adding the stellate signature to baseline clinical data (bilirubin greater than 1mg/dL and platelets count less than 100 000/mm³) improved the discrimination and overall fit of the Cox regression model, we calculated Harrell's c-statistic (a measure of the predictive power of the model ranging from 0.5 to 1) for each model, and compared the two nested models with a likelihood ratio test (assessing whether overall fit of the model is significantly improved by adding the extra variable). MELD and FIB4 scores were approximated for each patient and tested for correlation to outcomes using multivariable Cox regression modeling. For FIB4 score calculation, AST levels were not available, so ALT was substituted for AST. Additionally for MELD score calculation, INR was not available and was assumed to be 1 for all patients. All data analysis was performed using the R statistical language (www.r-project.org) and GenePattern.

Prognostic index

Based on our previously generated multivariate Cox regression model for overall survival in the HCV cohort (discovery cohort), we developed a prognostic index by linearly combining the variables weighted with regression coefficients from the multivariable Cox regression models in the prognostic index derivation set. Tertile values of the index in the derivation set were used as cutoffs to classify patients into high, intermediate, and low risk groups. This prognostic model was tested in the discovery cohort and validated in the HCC cohort (validation cohort) with identical definitions and cutoffs. To evaluate whether the hepatic stellate cell signature improved prediction of clinical outcomes, we compared the c-statistic improvement and 95% confidence interval of a predictive model using bilirubin and platelet count alone, to a model using bilirubin, platelet count and hepatic stellate cell signature expression. A c-statistic above 0.7 would be considered the lower limit for clinical utility.

Cell culture

The Hep3B, Huh7, LX2, Hep1-6, TSEC, and JS1 cells were cultured at 37°C and 5% CO₂ in Dulbecco's modification of Eagle's medium (high glucose, sodium pyruvate-free) (Invitrogen) supplemented with 10% fetal bovine serum and penicillin/streptomycin.

Quantitative reverse transcriptase polymerase chain reaction

RNA was extracted from adherent cell lines using the RNEasy Mini kit (Qiagen). Equimolar concentrations of RNA were converted to cDNA (Clontech Cat. 639549). Copy number was assessed by SYBR green qPCR on the Roche Lightcycler 480 platform. Human PCDH7 expression was measured using the 5'-TTGTGGGAGCAGGAGACAAC-3' forward and 5'-CTCTGCAGTGACCCCTGATG-3' reverse primers, which yielded a 154 base pair amplification product. Mouse PCDH7 expression was measured using the 5'-

TCCACTCCCAGAGGACAACT-3' forward and 5'-GGCTGGCTCTTCTTCCTCTC-3' reverse primers, which yielded a 198 base pair amplification product.

Immunofluorescence staining

Frozen liver tissue sections were prepared from mice treated for 6 weeks with 10% CCl₄ (intraperitoneal injection three times per week). The sections were fixed with ice-cold acetone for 5 minutes, permeabilized with 0.2% Triton X-100 in PBS for 15 minutes, rinsed three times with PBS, and blocked with normal goat serum diluted 1:20 or with 5% BSA/PBS for 30 minutes. Staining was performed with primary mouse McAb to PCDH7 or rabbit PcAb to Desmin (abcam®, Cambridge, MA) at 4° C overnight. Primary antibody detection was performed with Alexa Fluor® 488 -conjugated goat anti-rabbit IgG, e or Alexa Fluor® 647 -conjugated goat anti-mouse. The slides were mounted by ProLong® Gold antifade reagent with DAPI (Invitrogen), and examined using a Zeiss Axiophot microscope (Zeiss Inc.).

RESULTS

Development of the hepatic stellate cell signature

To identify genes highly and specifically expressed in stellate cells, expression profiles representing all liver cell subsets were subtracted from hepatic stellate cell profiles. These datasets were derived from a series of mouse-based transcriptome profiling studies, which was enabled by the extensive availability of isolated liver cell type datasets. Whole liver, hepatocyte, and fetal liver profiles were included, as well as representative epithelial, endothelial, myeloid, and erythroid lineages (Figure 1A and Table S1). In the global transcriptome space, samples clustered together according to cell type, as opposed to study or dataset, supporting the presence of robust transcriptional programs specific to each liver cell type (Figure 1B). A total of 122 genes were identified as highly and uniquely expressed in stellate cells (Figure 1C). As expected, several canonical hepatic stellate cell markers were highly expressed in the stellate cell samples, but excluded from the hepatic stellate cell signature due to concurrent expression in other cell or tissue types (Figure S2). The 122-gene hepatic stellate cell signature was highly enriched with extracellular matrix remodeling and related molecular pathways (Figure S3 and Table S2).

Up-regulation of the hepatic stellate cell signature in human and animal models of fibrosis

To validate the stellate cell gene expression signature for its relevance to cirrhosis-related phenotypes, we tested for hepatic stellate cell signature enrichment in ten different datasets, representing a variety of liver disease etiologies in humans, mice, and rats (Table S3). Hepatic stellate cell signature expression was strikingly induced in both human cirrhotic livers, compared to healthy normal livers (Figure 1D); and fibrotic NAFLD livers compared to non-fibrotic NAFLD livers (Figure 1E). The hepatic stellate cell signature was also significantly associated with the presence of inflammation in livers from NAFLD patients (Figure 1F). In each of the human datasets examined, the diseased population had significantly higher hepatic stellate cell signature enrichment compared to the normal population (Figure 1D–I). We examined gene ontology enrichment within the leading edge of each gene set enrichment analysis, and the 10 most statistically significant networks are

listed in (Figure S3). The same leading edge genes were also subject to Ingenuity pathway analysis, with the highest ranking interaction networks shown in Figures S4–S6. These networks implicate central involvement of many canonical stellate cell factors, including TGF- β , PDGFRA, PDGFRB, IL-1, FAK, collagen, and MMPs, in the regulation of hepatic stellate cell signature genes. Of particular interest to us is the regulation of TAGLN by NOTCH, FAK, TGF- β , and PDGF signaling, which may be a novel final common pathway of pro-fibrotic signaling. Additionally, the regulation of collagen expression by TNXB and ADAMTS2 highlights them as possible targets for antifibrotic therapies.

PCDH7 is a novel hepatic stellate cell surface marker

We explored whether genes in the hepatic stellate cell signature encode novel cell surface markers of stellate cells, which would make them particularly appealing as candidate therapeutic targets. To identify this subset of cell surface transcripts within the stellate cell expression signature, we intersected the hepatic stellate cell signature with LOCATE and gene ontology groups that are known to be on the cell surface, on the plasma membrane, at cell junctions, or in the extracellular matrix (Figure 2A, Table S4). This approach yielded several genes previously identified as stellate cell or myofibroblast markers, including DDR2,[23] EDNRA,[24] EDNRB,[25] EREG,[26] and IL1R1.[27] Additionally, this approach highlighted PCDH7, a member of the protocadherin family[28] as a candidate stellate cell surface marker. In the hepatic stellate cell signature discovery data, PCDH7 is highly expressed in both activated and quiescent stellate cells. It decreases slightly in activated stellate cells, but remains substantially higher than in other cell and tissue types examined (Figure 2B). We performed qRT-PCR to compare absolute expression of PCDH7 in immortalized stellate cell lines (LX2, JS1), sinusoidal endothelial cells (TSEC), and hepatocytes (HepG2, Hep3B, Huh7, Hep1–6). PCDH7 expression was enriched in immortalized stellate cells, compared to immortalized hepatocytes and sinusoidal endothelial cells (Figure 2C). We then performed protein characterization by immunofluorescence, observing sinusoidal staining, localizing with desmin, for PCDH7 in fibrotic mouse livers (Figure 2D–E).

Hepatic stellate cell gene signature in fibrosis models

We next assessed the extent of stellate cell enrichment in mouse and rat fibrosis models. For most validation sets, the mean enrichment score varied substantially between diseased and control groups, but statistical significance was not reached due to the low number of samples (Figure S7). However, GSEA demonstrated highly significant hepatic stellate cell signature enrichment in the mouse bile duct ligation, mouse unilateral ureteral obstruction, rat bile duct ligation, and rat diethylnitrosamine models (Table 1). Interestingly, strong association was observed between the hepatic stellate cell signature and the renal unilateral ureteral obstruction dataset, while no association was present between the hepatic stellate cell signature and the lung bleomycin data set; this suggests that liver fibrosis is genomically more closely related to renal fibrosis than to lung fibrosis.

The hepatic stellate cell signature expression is associated with patient prognosis in two human cirrhosis cohorts

To further validate the hepatic stellate cell signature, we explored the association between hepatic stellate cell signature expression and long-term clinical outcomes in cirrhotic cohorts of resected HCC patients[11] and hepatitis C infected cirrhotic patients.[10] Hepatic stellate cell signature expression was calculated for the resected hepatocellular carcinoma (Figure 3A) and hepatitis C (Figure 3B) patient cohorts. Patients were ordered by hepatic stellate cell signature expression (Figure 3C–D), and the top quartile was considered as a separate group in outcomes analysis. In the resected hepatocellular carcinoma cohort, patients with high hepatic stellate cell signature expression had significantly reduced survival and increased tumor recurrence (Figure 3E–F). In the hepatitis C cirrhosis cohort, patients with high hepatic stellate cell signature expression had significantly poorer liver-related outcomes, such as Child-Pugh class progression and hepatic decompensation, as well as decreased survival (Figure 3G–I). There was no significant difference in progression to HCC in the hepatitis C cohort (Figure 3J).

For the hepatitis C cirrhosis cohort, associations between high bilirubin ($>1.0\text{mg/dl}$), low platelet count ($<100,000/\text{mm}^3$), and patient outcome have been previously described.[10] To exclude the possibility that the hepatic stellate cell signature expression is a surrogate for either of these clinical features, we performed multivariable Cox regression to assess the independent contribution of each factor. For overall survival and progression of Child-Pugh class, high hepatic stellate cell signature expression was independently correlated with outcomes (Table 2). We also examined prognostic association between outcomes and albumin, FIB4 score, and MELD score. Although there was an association in univariate analysis of MELD and albumin levels with survival in the index derivation cohort, in multivariable Cox regression only bilirubin, platelet count and hepatic stellate cell signature expression were significantly associated with survival (Table S5). This is most likely because all patients in our prognostic index derivation set have earlier stage disease with limited clinical variable variation – most clinical variables were within normal reference ranges. Using the hepatic stellate cell signature and clinical data together improved prognostic model fit and discrimination in the prognostic index derivation set ($p=0.0015$ for death and $p=0.0043$ for Child-Pugh class progression respectively, see Table S6). Despite being significantly associated in univariable analysis, clinical variables and the hepatic stellate cell signature marker were not independently predictive of decompensation (Table 2). In the prognostic index validation set, the stellate signature was significantly associated with mortality in multivariable analysis, although the hepatic stellate cell signature was not associated with hepatocellular carcinoma development, consistent with the univariable results (Table 2). Since associations between gene expression and patient outcomes have been previously published in these same datasets, we checked for overlap between our hepatic stellate cell signature and the previously published prognostic signature. Only three genes overlapped between the previously published 186-gene prognostic signature and the 122-gene hepatic stellate cell signature, indicating that substantially different sets of genes are being examined (Figure S8).

Prognostic index

Given the prognostic association of the stellate cell signature, we developed a prognostic index combining bilirubin, platelet count and hepatic stellate cell signature expression. In the prognostic index derivation set (n=216), 88 subjects (n=41%) were classified as low-risk, 103 (48%) were classified as intermediate-risk and 25 (12%) were classified as high-risk. The high-risk group was associated with shorter overall survival (HR 8.00, $p<0.0001$), more frequent decompensation (HR 2.9, $p=0.0027$), and faster Child-Pugh class progression (HR 6.5, $p<0.0001$). Association with HCC development was not significant (HR=2.2, $p=0.068$) in this cohort (Table 3, Figure 4A–C). Adding the hepatic stellate cell signature improved the model c-statistic (and 95% confidence interval) from 0.66 (0.59–0.74) to 0.70 (0.62–0.78) for overall survival, 0.61 (0.54–0.68) to 0.62 (0.55–0.69) for decompensation, and 0.68 (0.61–0.75) to 0.70 (0.63–0.78) for progression of Child-Pugh class.

In the prognostic index validation cohort (n=82), 42 subjects (51%) were classified as low-risk, 30 (37%) were classified as intermediate-risk and 10 (12%) were classified as high-risk using cut-off values predefined in the derivation set. The high-risk group was associated with shorter overall survival (HR=3.9, $p=0.0077$), but not associated with HCC recurrence (HR=1.7, $p=0.20$) (Table 3, Figure 4D). Addition of hepatic stellate cell signature expression improved the model c-statistic (and 95% confidence interval) from 0.58 (0.48–0.68) to 0.62 (0.51–0.72) for death, and 0.51 (0.44–0.59) to 0.54 (0.46–0.62) for HCC recurrence.

DISCUSSION

We have used a bioinformatics approach to identify a group of genes that are specifically expressed by hepatic stellate cells. Furthermore, we show that these genes correlate with stellate cell abundance and patient outcomes, and contain transcripts that encode a novel cell surface marker of stellate cells.

The selection strategy for hepatic stellate cell signature genes in the discovery set yielded several different types of genes. Although all genes in the hepatic stellate cell signature were highly enriched in stellate cells, some were substantially higher in quiescent stellate cells, others were substantially higher in activated stellate cells, and some were equally expressed between quiescent and activated stellate cells. We initially explored using separate subsets of the hepatic stellate cell signature to capture this difference. We expected higher correlation between gene expression and fibrotic phenotype for the set of activated stellate cell genes. Instead, we found that no subgroup had a substantially stronger correlation with phenotype (Table S7). In addition, almost every gene in the signature was enriched in several disease phenotypes (Table S8). These results suggest that the hepatic stellate cell signature enrichment we observe in diseased phenotypes is a result of increased stellate cell number, rather than changes in stellate cell activation. This is important because it implies that stellate cell proliferation is a dominant driving factor for hepatic fibrosis. Hepatic stellate cell proliferation during stellate cell activation was described when the cell type was defined, [29, 30, 31] but stellate cell activation and proliferation have seldom been mechanistically uncoupled. Instead, stellate cell number is generally normalized to measure changes in RNA and protein expression. In addition to therapeutic approaches that limit stellate cell

activation, our data indicate that strategies limiting hepatic stellate cell proliferation merit strong consideration as potential antifibrotic therapies.

Since stellate cells comprise a minority of the cell volume in the liver, stellate cell gene expression changes are difficult to detect in whole liver expression profiles. By examining genes that are uniquely expressed in stellate cells – the genes in the hepatic stellate cell signature – tracking of stellate cell genes in whole liver samples becomes tractable. This allows us to rapidly assess stellate cell abundance in any mouse, rat, or human liver sample, provided that a microarray or RNA-seq procedure has been performed. This approach will be especially useful for characterizing novel genetic or etiological models of liver injury.

Although the hepatic stellate cell signature is independently predictive of patient outcomes, translating it into a clinical test would require further development. The prognostic index c-statistic for overall survival, the primary outcome, was 0.70 – this is the threshold generally regarded as clinically useful. The dependence on liver biopsy and cost of measuring the expression of 122 genes are obstacles to adoption, but do not preclude their future clinical implementation. Nonetheless, we primarily view this analysis as a validation of the hepatic stellate cell signature. Additionally, we hope that study of individual genes in the signature will help us uncover fundamental etiological drivers of chronic liver disease.

For easier clinical implementation, we explored whether subsets of the hepatic stellate cell signature were equally predictive of outcomes. We discovered that smaller gene subsets (31 to 55 genes, instead of 122) were similarly associated with overall survival (Table S9). The reduced 55 gene signature is a promising candidate clinical biomarker for evaluation in future studies.

The novel cell surface markers uncovered in our analysis have several potential applications. From an experimental perspective, isolation of pure stellate cell populations has been an ongoing challenge for the field. Use of these markers for MACS or flow cytometry would generate a purer population of isolated stellate cells. Additionally, the promoter of each gene in the hepatic stellate cell signature is potentially useful to drive stellate-targeted deletion of genes in animal models. From a clinical perspective, novel stellate cell surface markers could be used to target pharmacologic agents to stellate cells. Additionally, secreted members of the hepatic stellate cell signature may represent surrogate serum markers of fibrosis.

Our method for identifying stellate cell surface markers proved to be extremely robust. Out of the six cell surface markers computationally identified, five of them were previously described, and we identified an additional novel marker, PCDH7, which we have validated as a specific cell surface marker of stellate cells. PCDH7 is a transmembrane protocadherin with seven cadherin repeats on its extracellular domain.[28] Although there is a paucity of previously published work on PCDH7, protocadherins typically function in signal transduction and cell-cell recognition. These characteristics make PCDH7 an interesting molecule not only for its targetable, cell-surface location, but also as a possible regulator of hepatic stellate cell fibrogenesis.

Another major goal of our study was to seek a correlation between the hepatic stellate cell signature and fibrotic responses in other tissues. Specifically, we compared the signature to tissues from murine renal fibrosis (unilateral urethral obstruction), and murine lung fibrosis (bleomycin). The renal fibrosis model correlated very strongly with the hepatic stellate cell signature, while the lung fibrosis model did not (Table 1). This correlation suggests a stronger similarity between hepatic and renal fibrosis than to pulmonary fibrosis. This is consistent with our understanding of fibrosis pathogenesis among the three organs, as the biology of fibrosis in liver and kidney are more akin to each other through the presence of kidney and liver pericytes (ie. stellate cells), whereas the cellular source(s) of fibrogenic cells in lung is less certain. This is also consistent with the divergent clinical course of fibrotic disease in lung, which is typically more aggressive than those of liver and kidney fibrosis.

In summary, our work establishes a novel and clinically relevant platform for unbiased discovery of novel stellate cell genes. These genes correlate with clinical outcomes in human liver disease, and may contribute to fibrosis in vivo. This creates a unique opportunity to systematically uncover novel fibrosis biology, identify candidate drug targets, and define novel candidate biomarkers that correlate with disease outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

FUNDING

To SLF: NIH-DK056621, NIH-AA020709; To YH: NIH-DK099558, Irma T Hirschl Trust, Dr. Harold and Golden Lamport Research Award, To Icahn School of Medicine Medical Scientist Training Program: NIH-GM007280

REFERENCE LIST

1. Wynn TA. Cellular and molecular mechanisms of fibrosis. *J Pathol.* 2008; 214:199–210. [PubMed: 18161745]
2. Friedman SL. Mechanisms of hepatic fibrogenesis. *Gastroenterology.* 2008; 134:1655–69. [PubMed: 18471545]
3. Hinz B, Phan SH, Thannickal VJ, Galli A, Bochaton-Piallat ML, Gabbiani G. The myofibroblast: one function, multiple origins. *Am J Pathol.* 2007; 170:1807–16. [PubMed: 17525249]
4. Friedman SL. Hepatic stellate cells: protean, multifunctional, and enigmatic cells of the liver. *Physiol Rev.* 2008; 88:125–72. [PubMed: 18195085]
5. Puche JE, Lee YA, Jiao J, Aloman C, Fiel MI, Munoz U, et al. A novel murine model to deplete hepatic stellate cells uncovers their role in amplifying liver damage in mice. *Hepatology.* 2013; 57:339–50. [PubMed: 22961591]
6. Tseng GC, Ghosh D, Feingold E. Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Res.* 2012; 40:3785–99. [PubMed: 22262733]
7. Kim H, Kim K, Yu SJ, Jang ES, Yu J, Cho G, et al. Development of Biomarkers for Screening Hepatocellular Carcinoma Using Global Data Mining and Multiple Reaction Monitoring. *PLoS One.* 2013; 8
8. Blaauboer ME, Emson CL, Verschuren L, van Erk M, Turner SM, Everts V, et al. Novel combination of collagen dynamics analysis and transcriptional profiling reveals fibrosis-relevant genes and pathways. *Matrix Biology.* 2013; 32:424–31. [PubMed: 23648810]

9. Hoshida Y, Nijman SM, Kobayashi M, Chan JA, Brunet JP, Chiang DY, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res.* 2009; 69:7385–92. [PubMed: 19723656]
10. Hoshida Y, Villanueva A, Sangiovanni A, Sole M, Hur C, Andersson KL, et al. Prognostic gene expression signature for patients with hepatitis C-related early-stage cirrhosis. *Gastroenterology.* 2013; 144:1024–30. [PubMed: 23333348]
11. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med.* 2008; 359:1995–2004. [PubMed: 18923165]
12. Lalazar A, Wong L, Yamasaki G, Friedman SL. Early genes induced in hepatic stellate cells during wound healing. *Gene.* 1997; 195:235–43. [PubMed: 9305769]
13. De Minicis S, Seki E, Uchinami H, Kluwe J, Zhang Y, Brenner DA, et al. Gene expression profiles during hepatic stellate cell activation in culture and in vivo. *Gastroenterology.* 2007; 132:1937–46. [PubMed: 17484886]
14. Troeger JS, Mederacke I, Gwak GY, Dapito DH, Mu X, Hsu CC, et al. Deactivation of hepatic stellate cells during liver fibrosis resolution in mice. *Gastroenterology.* 2012; 143:1073–83, e22. [PubMed: 22750464]
15. Kisseleva T, Brenner DA. Inactivation of myofibroblasts during regression of liver fibrosis. *Cell Cycle.* 2013; 12:381–2. [PubMed: 23324345]
16. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics.* 2003; 4:249–64. [PubMed: 12925520]
17. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. *Nature Genetics.* 2006; 38:500–1. [PubMed: 16642009]
18. Sprenger J, Lynn Fink J, Karunaratne S, Hanson K, Hamilton NA, Teasdale RD. LOCATE: a mammalian protein subcellular localization database. *Nucleic Acids Res.* 2008; 36:D230–3. [PubMed: 17986452]
19. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols.* 2009; 4:44–57. [PubMed: 19131956]
20. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research.* 2009; 37:1–13. [PubMed: 19033363]
21. 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. *P Natl Acad Sci USA.* 2005; 102:15545–50.
22. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1 alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics.* 2003; 34:267–73. [PubMed: 12808457]
23. Ankoma-Sey V, Matli M, Chang KB, Lalazar A, Donner DB, Wong L, et al. Coordinated induction of VEGF receptors in mesenchymal cell types during rat hepatic wound healing. *Oncogene.* 1998; 17:115–21. [PubMed: 9671320]
24. Yokomori H, Oda M, Ogi M, Kamegaya Y, Tsukada N, Nakamura M, et al. Enhanced expression of endothelin receptor subtypes in cirrhotic rat liver. *Liver.* 2001; 21:114–22. [PubMed: 11318980]
25. Yokomori H, Oda M, Yasogawa Y, Nishi Y, Ogi M, Takahashi M, et al. Enhanced expression of endothelin B receptor at protein and gene levels in human cirrhotic liver. *Am J Pathol.* 2001; 159:1353–62. [PubMed: 11583963]
26. Inatomi O, Andoh A, Yagi Y, Bamba S, Tsujikawa T, Fujiyama Y. Regulation of amphiregulin and epi-regulin expression in human colonic subepithelial myofibroblasts. *Int J Mol Med.* 2006; 18:497–503. [PubMed: 16865236]
27. Gieling RG, Wallace K, Han YP. Interleukin-1 participates in the progression from liver injury to fibrosis. *Am J Physiol Gastrointest Liver Physiol.* 2009; 296:G1324–31. [PubMed: 19342509]
28. Yoshida K, Yoshitomo-Nakagawa K, Seki N, Sasaki M, Sugano S. Cloning, expression analysis, and chromosomal localization of BH-protocadherin (PCDH7), a novel member of the cadherin superfamily. *Genomics.* 1998; 49:458–61. [PubMed: 9615233]

29. Friedman SL, Arthur MJ. Activation of cultured rat hepatic lipocytes by Kupffer cell conditioned medium. Direct enhancement of matrix synthesis and stimulation of cell proliferation via induction of platelet-derived growth factor receptors. *J Clin Invest.* 1989; 84:1780–5. [PubMed: 2556445]
30. Gressner AM. Cytokines and cellular crosstalk involved in the activation of fat-storing cells. *J Hepatol.* 1995; 22:28–36. [PubMed: 7665846]
31. Kinnman N, Gorla O, Wendum D, Gendron MC, Rey C, Poupon R, et al. Hepatic stellate cell proliferation is an early platelet-derived growth factor-mediated cellular event in rat cholestatic liver injury. *Lab Invest.* 2001; 81:1709–16. [PubMed: 11742041]
32. Wurmbach E, Chen YB, Khitrov G, Zhang W, Roayaie S, Schwartz M, et al. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. *Hepatology.* 2007; 45:938–47. [PubMed: 17393520]
33. Mellentin JD, Murre C, Donlon TA, McCaw PS, Smith SD, Carroll AJ, et al. The gene for enhancer binding proteins E12/E47 lies at the t(1;19) breakpoint in acute leukemias. *Science.* 1989; 246:379–82. [PubMed: 2799390]

SUMMARY

What is already known about the subject

- In the setting of chronic liver injury, activated hepatic stellate cells are central mediators of liver fibrosis and cirrhosis.
- Patients with cirrhosis have variable progression to liver failure and hepatocellular carcinoma.
- Biomarkers derived from stellate cells and/or fibrotic tissue are urgently needed to facilitate clinical trials for anti-fibrotic agents.

What are the new findings

- We systematically define highly and uniquely expressed stellate cell genes, and show that these genes are enriched in human and animal models of liver disease.
- There are strong similarities between the hepatic stellate cell signature and renal fibrosis expression changes, but the hepatic stellate cell signature is distinct from pulmonary fibrosis expression changes.
- Our approach has identified PCDH7 as a novel extracellular marker of stellate cells. PCDH7 is a candidate marker for stellate cell tracking, targeting, and isolation.
- Increased expression of hepatic stellate cell signature genes in patients is predictive of clinical outcomes, including decompensation, progression of Child-Pugh class, and overall survival. Combining these predictions with clinical factors leads to more accurate predictions than using clinical factors alone.

How might it impact on clinical practice in the foreseeable future?

The hepatic stellate cell signature, alone or in combination with routine clinical measurements, may provide a tool for improved prediction of cirrhosis progression, guiding anti-cirrhosis monitoring or intervention. Additionally, the hepatic stellate cell signature may serve as a surrogate marker of long-term outcome for anti-fibrotic agents in clinical trials.

Figure 1. Identification of genes uniquely enriched in stellate cells

(A) Summary of study design. (B) Expression of all genes examined, across all tissue and cell types examined. The samples are ordered by similarity in hierarchical clustering, with the dendrogram above the heat map illustrating degree of similarity. Highly expressed genes are present for each cell and tissue type, but there is considerable expression overlap between different cell and tissue types. (C) Expression of genes in the hepatic stellate cell signature. Only genes that were uniquely and highly expressed in stellate cells were included in the signature. Enrichment of the stellate signature was confirmed in a hepatitis C cohort,

[32] comparing normal and cirrhotic patients (D), a NAFLD cohort,[33] comparing fibrotic and non-fibrotic patients (E), and the same NAFLD cohort, comparing patients with inflammation and no inflammation (F). In each cohort, stellate signature expression strongly correlates with the diseased phenotype. In (D–F), hepatic stellate cell signature expression was assessed for each patient sample, and assigned a numeric value using the mean Kolmogorov-Smirnov statistic. Mean stellate cell enrichment scores are shown for each patient group, and plotted $\pm$ standard error of the mean. Welch's t-test was used to calculate the reported p-values. In (G–I) we show gene set enrichment analysis of the stellate signature in each patient dataset. Array genes were ordered from highest in healthy liver (left side) to highest in diseased liver (right side). Vertical black bars indicate the location of each gene in the stellate signature. The normalized enrichment score (NES) and statistical significance are included

Figure 2. PCDH7 is a candidate hepatic stellate cell surface marker

(A) Overlap between stellate signature genes, cell surface gene ontologies, and LOCATE membrane protein predictions highlight potential cell surface stellate cell markers. (B–C) Of the 6 putative cell surface markers, we confirmed expression of PCDH7 in our original array data (B) and by qRT-PCR in human and mouse cell lines (C). For the array data, PCDH7 expression is shown $\pm$ SEM for three or more samples. For the qRT-PCR results, absolute expression is shown, with the black bars indicating $\pm$ SEM between three replicates. P-values are from unpaired, two-tailed t-tests for difference of means. (C–D) Staining of

PCDH7, a putative stellate cell surface marker, in mice treated with carbon tetrachloride for 6 weeks. Good overlap is observed with desmin, especially in the fibrotic septa.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. Correlation between stellate cell signature expression and clinical outcomes
 (A–B) Heat maps of patient stellate signature expression in the hepatitis C (A) and resected HCC (B) cirrhotic cohorts, sorted by enrichment score (datasets from [10] and [11] respectively). Each column is one patient, and each row is one stellate cell signature gene. The vertical dashed line separates the two patient groups analyzed in subsequent panels. (C–D) Stellate signature enrichment scores for each patient. The 25% of patients with the highest hepatic stellate cell signature enrichment are separated by the dotted line, and analyzed as a separate group for adverse outcomes in (E–J). The red survival curves represent the 25% of patients with the highest stellate cell signature enrichment, and the blue survival curves represent that 75% of patients with lower stellate cell signature enrichment. For the hepatitis C cohort, survival (E), progression of Child-Pugh class (F), decompensation (G), and HCC incidence (H) are shown. For the resected HCC cohort, survival (I) and HCC recurrence (J) are shown. Patients with high hepatic stellate cell signature expression had significantly worse survival, increased Child-Pugh class progression, and increased decompensation.

Figure 4. A prognostic index comprised of hepatic stellate cell signature expression, bilirubin, and platelet count, is associated with overall survival in prognostic index derivation and validation sets

Multivariable Cox regression was used to assign relative weights to serum bilirubin (0.8082), platelet count (0.9896), and hepatic stellate cell signature expression (0.8516). Patients received a higher prognostic index for increased serum bilirubin, decreased platelet count, or increased hepatic stellate cell signature expression. Patients were assigned to one of three risk groups based on their prognostic index, and risk group cutoffs were set such that one third of the derivation cohort was in each risk group. (A–C) Clinical outcomes for the prognostic index derivation set (n=216). There is robust separation between each of the risk groups for survival, hepatic decompensation, and progression of Child-Pugh class. (D) Survival in the prognostic index validation set (n=82), with a substantial survival difference in the high risk group, compared to the medium and low risk groups. Hazard ratios, 95% confidence intervals, and p-values are given for the difference between the high and low risk groups. We provide additional, detailed statistics in Table 3.

Table 1

Association between hepatic stellate cell signature expression and clinical or experimental phenotypes (gene set enrichment analysis)

Organism	Dataset	Groups Compared	NES*	FDR**	p-value***
Human	Non-alcoholic fatty liver disease	Inflammation present vs. absent	1.49	0.071	0.009
		Severe versus moderate fibrosis	1.97	0.002	<0.001
	Cirrhosis	Cirrhosis vs. normal	2.08	<0.001	<0.001
Mouse	Carbon tetrachloride (CCl ₄) injection	CCl ₄ vs. vehicle	1.39	0.141	0.067
	Bile duct ligation (BDL)	28 days post BDL vs. sham surgery	2.01	<0.001	<0.001
	Primary sclerosing cholangitis	Cholic acid vs. ursodeoxycholic acid diet, Abcb4 mouse	****	****	****
	Unilateral urethral obstruction (UUO, kidney)	9 days UUO vs. sham surgery	2.24	<0.001	<0.001
	Bleomycin (lung)	5 weeks bleomycin vs. vehicle	****	****	****
Rat	Carbon tetrachloride injection	3 days CCl ₄ vs. vehicle	0.33	0.305	0.135
	Bile duct ligation	2 weeks post BDL vs. sham surgery	2.39	<0.001	<0.001
	Diethylnitrosamine (DEN)	18 weeks DEN vs. vehicle	1.84	<0.001	<0.001

Gray boxes indicate non-significant p-values (p>0.05)

* NES = normalized enrichment score

** FDR = false discovery rate

*** p-value = gene set enrichment analysis nominal p-value

**** no enrichment

Table 2

Association of hepatic stellate cell signature with clinical outcomes in the prognostic index derivation set (n=216, multivariable Cox regression)

	Outcome			
	Child progression *	Decompensation **	Death	HCC
Bilirubin (>1.0 mg/dL)				
hazard ratio	2.54	1.62	2.24	2.03
HR, 95% CI	[1.46 – 4.44]	[0.99 – 2.67]	[1.3 – 3.87]	[1.18 – 3.48]
p-value	0.001	0.06	0.004	0.01
Platelet count (<100,000 / mm ³)				
hazard ratio	2.26	1.47	2.69	1.56
HR, 95% CI	[1.32 – 3.89]	[0.90 – 2.4]	[1.52 – 4.75]	[0.91 – 2.69]
p-value	0.003	0.12	0.0006	0.11
Stellate signature (>75th percentile)				
hazard ratio	2.16	1.60	2.34	1.10
HR, 95% CI	[1.30 – 3.60]	[0.97 – 2.64]	[1.41 – 3.88]	[0.62 – 1.94]
p-value	0.003	0.07	<0.001	0.75

* Progression of Child-Pugh class

** Hepatic decompensation: a composite score englobing gastrointestinal bleeding, ascites and/or hepatic encephalopathy

Table 3

Association of the prognostic index with clinical outcomes

Cohort	Outcome	Risk group	HR (95% confidence interval)	p-value
Derivation cohort: Hepatitis C cirrhosis (n=216)	Overall survival (n=66, 31%)	Intermediate	3.40 (1.80–6.40)	<0.001
		High	8.00 (3.70–17.00)	<0.001
	Child progression [*] (n=66, 31%)	Intermediate	2.40 (1.30–4.40)	0.004
		High	6.5 (3.2–13)	<0.001
	Decompensation ^{**} (n=71, 34%)	Intermediate	1.6 (0.94–2.70)	0.09
		High	2.9 (1.4–5.7)	0.003
	Development of HCC (n=65, 30%)	Intermediate	1.7 (0.98–3.0)	0.06
		High	2.2 (0.94–4.9)	0.07
Validation cohort: Surgically treated HCC (n=82)	Overall survival (n=32, 39%)	Intermediate	1.2 (0.55–2.6)	0.64
		High	3.9 (1.4–11)	0.008
	HCC recurrence (n=65, 79%)	Intermediate	1.2 (0.72–2.1)	0.45
		High	1.7 (0.76–3.7)	0.20

Low risk group as reference

^{*}
HR: hazard ratio^{*}
progression of Child-Pugh class^{**}
hepatic decompensation: a composite score englobing gastrointestinal bleeding, ascites and/or hepatic encephalopathy