A Fibrosis-Independent Hepatic Transcriptomic Signature Identifies Drivers of Disease Progression in Primary Sclerosing Cholangitis

Yevgeniy Gindin, Chuhan Chung, Zhaoshi Jiang, Jing Zhu Zhou, Jun Xu, Andrew N. Billin, Robert P. Myers, Zachary Goodman, Abdolamir Landi, Michael Houghton, Richard M. Green, Cynthia Levy, Kris V. Kowdley, Christopher L. Bowlus, Andrew J. Muir, and Michael Trauner

BACKGROUND AND AIMS: Primary sclerosing cholangitis (PSC) is a heterogeneous cholangiopathy characterized by progressive biliary fibrosis. RNA sequencing of liver tissue from patients with PSC (n = 74) enrolled in a 96-week clinical trial was performed to identify associations between biological pathways that were independent of fibrosis and clinical events.

APPROACH AND RESULTS: The effect of fibrosis was subtracted from gene expression using a computational approach. The fibrosis-adjusted gene expression patterns were associated with time to first PSC-related clinical event (e.g., cholangitis, hepatic decompensation), and differential expression based on risk groups and Ingenuity Pathway Analysis were performed. Baseline demographic data were representative of PSC: median age 48 years, 71% male, 49% with inflammatory bowel disease, and 44% with bridging fibrosis or cirrhosis. The first principle component (PC1) of RNA-seq sequencing data accounted for 18% of variance and correlated with fibrosis stage (ρ = −0.80; P < 0.001). After removing the effect of fibrosis-related genes, the first principle component was not associated with fibrosis (ρ = −0.19; P = 0.11), and a semisupervised clustering approach identified two distinct patient clusters with differential risk of time to first PSC-related event (P < 0.0001). The two groups had similar fibrosis stage, hepatic collagen content, and α-smooth muscle actin expression by morphometry, Enhanced Liver Fibrosis score, and serum liver biochemistry, bile acids, and IL-8 (all P > 0.05). The top pathways identified by Ingenuity Pathway Analysis were eukaryotic translation inhibition factor 2 (eIF2) signaling and regulation of eIF4/p70S6K signaling. Genes involved in the unfolded protein response, activating transcription factor 6 (ATF6) and eIF2, were differentially expressed between the PSC clusters (down-regulated in the high-risk group by log-fold changes of −0.18 [P = 0.02] and −0.16 [P = 0.02], respectively). Clinical events were enriched in the high-risk versus low-risk group (38% [12/32] vs. 2.4% [1/42], P < 0.0001).

CONCLUSIONS: Removing the contribution of fibrosis-related pathways uncovered alterations in the unfolded protein response, which were associated with liver-related complications in PSC. (Hepatology 2021;73:1105-1116).

Primary sclerosing cholangitis (PSC) is a heterogeneous and progressive cholangiopathy characterized by chronic inflammation and...
scarring of the intrahepatic and/or extrahepatic bile ducts. The pathogenesis of PSC remains incompletely understood, and currently no therapy outside of liver transplantation improves clinical outcomes.\(^1,2\) The regional variability of the histologic features of PSC limits the use of liver biopsy in routine clinical practice, although some studies demonstrate a prognostic role for histology-based scores.\(^3\) A recent controlled clinical trial demonstrated that the stage of fibrosis determined histologically or based on noninvasive markers (e.g., Enhanced Liver Fibrosis score [ELF], liver stiffness by transient elastography) is predictive of liver-related complications in patients with PSC.\(^4\) Indeed, the primary endpoint of two ongoing phase 3 trials of therapies for PSC includes non-progression of liver fibrosis determined histologically (NCT03872921; NCT03890120).

The etiology of PSC likely reflects interactions between known genetic risk loci and environmental factors.\(^5,6\) The shared genetic risk loci with other autoimmune diseases emphasizes genetic susceptibility and immune dysfunction. These genetic associations, however, do not explain the widely varying disease manifestations that occur with other autoimmune diseases or within PSC itself, given the rarity of developing PSC in patients with similar risk loci. The absence of clinical benefit in trials using a host of immunosuppressive therapies speaks to the limitation of suppressing inflammation alone in PSC.\(^7\) Thus, patient-specific environmental risk factors and highly variant genetics combine to generate a widely variable clinical phenotype.

Given the elusive nature of PSC pathogenesis, we undertook RNA sequencing (RNA-seq), an unbiased approach to analysis of the hepatic transcriptome, using liver biopsies obtained from patients with PSC enrolled in a randomized controlled trial.\(^4\) Patients with PSC across the spectrum of fibrosis severity were enrolled, and a bioinformatics approach was used to identify drivers of PSC pathogenesis. We hypothesized that removing the influence of fibrosis-related genes could potentially reveal biologic determinants of PSC pathogenesis.

### Materials and Methods

#### STUDY POPULATION AND OUTCOMES

The PSC study population was derived from a phase 2b, placebo-controlled trial of simtuzumab, a...
monoclonal antibody directed against lysyl oxidase-like 2 (NCT01672853). The design and results of this trial have been described.\(^4\) In brief, adult patients with compensated, large-duct PSC were randomized to receive weekly subcutaneous injections of simtuzumab (either 75 mg or 125 mg) or placebo for 96 weeks. Percutaneous liver biopsies were obtained in a protocol-specified manner with a central pathologist (Z.G.) interpreting the adequacy of the biopsy specimen for diagnosis and staging of PSC.\(^4\) The primary objective of the study was to evaluate whether simtuzumab was effective in preventing histologic progression of fibrosis, as measured by changes in hepatic collagen content. The study was approved by the institutional review boards of participating institutions, and all patients provided written informed consent, including for genomic sequencing, before participation in the study.

The hepatic transcriptomic analysis was restricted to baseline liver biopsy samples (see subsequently). Biopsies were reviewed by a central pathologist who staged fibrosis according to the Ishak classification. Patients were categorized as having no to mild fibrosis (Ishak stages 0-1), moderate fibrosis (stages 2-4), or cirrhosis (stages 5-6). Morphometric quantification of hepatic collagen content and alpha-smooth muscle actin (α-SMA) expression were also performed, as previously described.\(^4\) The outcome of interest for the study was time to first PSC-related clinical event, defined as ascending cholangitis, sepsis, hepatic decompensation (i.e., ascites, variceal hemorrhage, hepatic encephalopathy), cholangiocarcinoma, hepatocellular carcinoma, liver transplantation, or death during the 96-week duration of the study. In patients with multiple clinical events, the earliest onset event was reported.

**HEPATIC TRANSCRIPTOMIC ANALYSIS**

RNA-seq (SureSelect protocol) was analyzed on formalin-fixed, paraffin-embedded liver biopsies (n = 74). RNA quality control was assessed by DV200 > 10%, with an initial collection from n = 84, with n = 6 failing to meet the DV200 threshold and another 4 patients lost to clinical follow-up. Gene expression was quantified using Salmon.\(^8\) A semisupervised bioinformatics approach was used to identify in-tissue, fibrosis-independent gene expression signatures associated with time to first PSC-related clinical event. Specifically, the effect of fibrosis was subtracted from gene expression (as implemented in the “removeBatchEffect” function in R package limma)\(^9\) by first fitting a linear model to each gene, relating its expression to fibrosis stage, and then subtracting the portion of gene expression that can be accounted for by fibrosis. This method had the net effect of removing genes with expression that is highly correlated with fibrosis stage. A Cox proportional hazards regression model was fit onto every gene after fibrosis correction to select genes with a significant association with PSC-related clinical events (unadjusted P value ≤ 0.05). Patient clustering, using genes associated with clinical events as input after adjusting for fibrosis, was carried out using a consensus clustering algorithm.\(^10,11\) Consensus clustering ascertains the confidence in the number of clusters (here patient clusters) present in the data, as well as the robustness of these clusters to random sampling. Differential expression based on the identified clusters was performed using limma-voom.\(^12\) P values were corrected for multiple testing where appropriate.

**INGENUITY PATHWAY ANALYSIS**

RNA-seq data were analyzed using Ingenuity Pathway Analysis software (Ingenuity Systems, Redwood City, CA) to examine for transcriptional networks associated with biological signaling. Scores were used to rank signaling and pathway networks according to their function.

**BIOMARKERS OF BILE ACID HOMEOSTASIS, LIVER FIBROSIS, AND IL-8**

Levels of total bile acids (BAs) and 15 specific BA species were quantified by liquid chromatography-tandem mass spectrometry (LC/MS-MS; Agilent 1290/Sciex; Metabolon, Durham, NC) in fasting serum samples obtained at baseline (n = 40). In addition, we measured the serum levels of 7α-hydroxy-4-cholesten-3-one (C4), a precursor in BA synthesis, by LC/MS-MS, and serum levels of fibroblast growth factor 19 (FGF19), a hormone that serves as the inhibitory signal for CYP7A1, the rate-limiting enzyme in BA synthesis, by Quantikine enzyme-linked immunosorbent assay (R&D, Minneapolis, MN). ELF score
(Siemens Healthcare, Erlangen, Germany), a noninvasive marker of liver fibrosis, was tested according to manufacturer’s specifications. IL-8 protein was measured by a commercial multiplex assay (Meso Scale Diagnostics, Rockville, MD).

**STATISTICAL ANALYSES**

Comparisons between groups were made using Wilcoxon rank sum tests, and correlations between parameters using Spearman correlations. Time to first PSC-related clinical event was assessed using Kaplan-Meier analysis and Cox proportional hazards regression using the survival R package version 2.41.\(^{(13)}\)

**Results**

**STUDY POPULATION**

In total, 74 patients with compensated, large-duct PSC were included in the current study. Their demographics and baseline characteristics are given in Table 1. The mean age was 46 years, 70% were male, 81% were Caucasian, 50% had ulcerative colitis, and 68% were prescribed ursodeoxycholic acid (UDCA). Eleven patients (15%) had cirrhosis (Ishak stages 5-6) on baseline liver biopsy.

**PRINCIPAL COMPONENT ANALYSIS IDENTIFIES FIBROSIS-RELATED GENES AS PRIMARY DRIVERS OF GENE-EXPRESSION DIFFERENCES ACROSS FIBROSIS STAGES**

The first principal component (PC1) of the RNA-seq data accounted for 18% of variance in hepatic gene expression (Fig. 1A and Supporting Fig. S1). Without adjustment for fibrosis stage, PC1 derived from the gene-expression data was highly correlated with Ishak fibrosis stage at the individual patient level (Spearman \(\rho = -0.80; P < 0.001\)) (Fig. 1B). As expected, the most highly up-regulated pathways according to pathway analysis were those involved in fibrogenesis, specifically genes involved in cell adhesion and cell-matrix interactions \((P = 1.60 \times 10^{-21})\) and stellate cell activation and liver fibrosis \((P = 1.28 \times 10^{-12})\) (Supporting Table S1).

**TABLE 1. Baseline Demographics and Clinical Characteristics of the Study Population**

| Characteristic                        | Total (n = 74) |
|---------------------------------------|---------------|
| **Demographics**                      |               |
| Age, years                            | 46.3 (11.3)   |
| Female                                | 22 (30)       |
| Body mass index, kg/m²                | 26.9 (4.8)    |
| UDCA use                              | 50 (68)       |
| Ulcerative colitis                    | 37 (50)       |
| Ethnicity                             |               |
| European descent                      | 60 (81)       |
| African-American                      | 12 (16)       |
| Asian                                 | 2 (3)         |
| **Liver tests**                       |               |
| Alanine aminotransferase, U/L         | 85.1 (72.7)   |
| Alkaline phosphatase, U/L             | 332.6 (302.4) |
| Gamma-glutamyltransferase, U/L        | 408.6 (457.4) |
| Bilirubin, mg/dL                      | 1.1 (1.3)     |
| Albumin, g/dL                         | 4.0 (0.4)     |
| International normalized ratio        | 1.0 (0.2)     |
| Platelets, \(\times10^3\)/uL         | 260.5 (98.0)  |
| **Other biomarkers**                  |               |
| ELF                                   | 9.47 (1.35)   |
| Hyaluronic acid, ng/mL                | 100.8 (188.2) |
| TIMP-1, ng/mL                         | 331.3 (163.7) |
| PIII-NP, ng/mL                        | 9.6 (5.7)     |
| Hemoglobin, g/dL                      | 13.9 (1.7)    |
| Total bile acids, pg/mL               | 1590.2 (586.2, 3649.2)* |
| C4, ng/mL                             | 8.6 (4.3, 24.6) |
| FGF19, pg/mL                          | 135.8 (60.4, 221.4) |
| IL-8, pg/mL                           | 19.5 (12.0, 41.0) |
| **Liver histology**                   |               |
| Ishak fibrosis stage                  |               |
| F0-F1                                 | 26 (35)       |
| F2-F4                                 | 37 (50)       |
| F5-F6                                 | 11 (15)       |
| Hepatic collagen content, %           | 5.3 (5.1)     |
| \(\alpha\)-SMA expression, %         | 5.3 (7.7)     |

Note: All data are presented as mean (SD) or n (%). Serum bile acids, C4, and FGF19 data are available for \(n = 40\) subjects.

*UDCA and UDCA conjugates were removed from the BA pool.

Abbreviations: PIII-NP, type III procollagen peptide; TIMP-1, tissue inhibitor of metalloproteinase 1.

**REMOVING THE EFFECT OF FIBROSIS IDENTIFIES TWO PATIENT SUBGROUPS WITH DISTINCT RISK OF PSC-RELATED CLINICAL EVENTS**

A bioinformatics approach was used to determine whether patients could be stratified according to the
risk of PSC-related clinical events based on biologic pathways devoid of the influence of fibrosis-related genes and pathways. After removing the effect of fibrosis on gene-expression data, PC1 was no longer associated with the stage of fibrosis ($\rho = -0.19$; $P = 0.11$; Fig. 1C,D). This approach yielded two distinct clusters of patients with PSC according to the risk of PSC-related clinical events, herein referred to as “low-risk” ($n = 42$) and “high-risk” ($n = 32$) groups (Fig. 2A). As shown in Fig. 2B, survival free of clinical events was significantly greater in the low-risk versus high-risk group (log-rank $P < 0.001$). During a median follow-up of 23 months (interquartile range [IQR], 22.52, 23.25), 1 of 42 patients (2.4%) in the low-risk group experienced clinical events ($n = 1$, ascending cholangitis), while 12 of 32 patients (38%) in the high-risk group had events (ascending cholangitis [$n = 7$], cholangiocarcinoma [$n = 2$], hepatic encephalopathy [$n = 1$], jaundice [$n = 1$], and sepsis [$n = 1$]).

To confirm the ability of the bioinformatics approach to remove the influence of fibrosis-related gene-expression pathways on the patient clustering, the clinical characteristics of the two clusters were compared. As indicated in Table 2, no differences in demographic parameters, UDCA use, liver biochemistry, platelets, ELF score, or hepatic collagen content or $\alpha$-SMA expression by morphometry were observed between the low-risk and high-risk clusters. Importantly, the distribution of fibrosis stages was nearly identical between groups: 14% of the low-risk group and 16% of the high-risk group had cirrhosis ($P = 1.00$). Thus, a fibrosis-independent, hepatic gene-expression signature stratifies patients with PSC into two discrete clusters with distinct prognoses.
Pathway analysis revealed that multiple components of the unfolded protein response (UPR) were significantly down-regulated in the high-risk patient cluster (highlighted in red in Fig. 3). Among these genes were activating transcription factor 6 (ATF6), a transcriptional regulator of endoplasmic reticulum (ER) chaperones that facilitate protein folding, and eukaryotic translation inhibition factor 2α (eIF2α), an enzyme necessary for protein translation. Expression of these genes was down-regulated in the high-risk group compared with the low-risk group by log-fold changes of −0.18 (P = 0.02) and −0.16 (P = 0.02), respectively. The list of the top differentially expressed genes in the high-risk versus low-risk groups revealed other components of the protein translational apparatus and structural adhesion proteins (Table 3).

Experimental models of PSC postulate a role for BAs in the induction of the UPR, an effect that may be mediated in part by activation of the farnesoid X receptor (FXR). \(^{14}\) However, serum levels of multiple BAs, including those that serve as the endogenous ligands for FXR, were similar in the low-risk and high-risk patient clusters (Table 4 and Supporting Table S2). Subtracting the contribution of UDCA resulted in similar levels of BAs between clusters. Pooled assessments of the endogenous FXR agonists (chenodeoxycholic acid, deoxycholic acid, and cholic acid) alone and combined also demonstrated similar levels between the two groups (Supporting Table S2). Moreover, serum levels of C4 and FGF19 were similar, suggesting similarities in FXR signaling between the two clusters. In contrast, differential expression of genes involved in the uptake and excretion of BAs was seen. The gene for OATP-C (SLCO1B1) was significantly reduced (P < 0.001) in the high-risk cluster, and a trend toward significance in the gene for bile salt export pump was also apparent (P = 0.15) (Supporting Fig. S2). Expression of other genes involved in BA uptake, including NTCP (SLC10A1), did not show significant differences. Thus, differences in UPR expression seen in the two patient clusters were not clearly related to BA levels, although differential expression of some transporters was apparent.

Levels of the key neutrophil chemoattractant IL-8 in serum and bile have been associated with worse clinical outcomes in patients with PSC and can be elicited by the UPR.\(^{15-17}\) As such, we compared the hepatic expression and serum levels of IL-8 between the two PSC patient clusters. As shown in Fig. 4, median serum levels of IL-8...
were similar between the low-risk and high-risk clusters (17.5 vs. 21.5 pg/mL; \( P = 0.62 \)), whereas the hepatic expression of the IL-8 gene, CXCL8, was higher in patients at increased risk for clinical events (\( P = 0.076 \)). Whether the decreased expression of genes that comprise an adaptive ER stress response are associated with potentially higher cytokine responses is unknown.

### Discussion

In the current study, we hypothesized that a bioinformatics approach that removed the influence of fibrosis-related genes in the hepatic transcriptome could potentially identify determinants of PSC pathogenesis. An initial effort to remove the influence of fibrosis-related genes identified a prognostic signature of nine genes related to immune responses and cell adhesion proteins, but incompletely removed the influence of fibrosis.\(^{(18)}\) However, the current approach eliminated the influence of fibrosis, as evidenced by similar histologic (e.g., Ishak stage, hepatic collagen content, \( \alpha \)-SMA expression) and nonhistologic (e.g., ELF score, liver biochemistry, platelets) parameters of fibrosis between two patient groups characterized by distinct gene-expression profiles.
The incidence of PSC-related clinical events was markedly different between the two clusters, with a greater number of events, including events of hepatic decompensation, in the high-risk subgroup compared with a single case of ascending cholangitis in the low-risk group. Interrogation of

| Gene     | Log-Fold Change in High vs. Low-Risk Group | Unadjusted P-Value |
|----------|--------------------------------------------|--------------------|
| EIF2A    | -0.16204                                   | 3.18E-04           |
| ATF6     | -0.17518                                   | 3.99E-04           |
| MAPK8    | -0.15154                                   | 0.007365           |
| BCL2     | 0.656579                                   | 0.016592           |
| CALR     | 0.288614                                   | 0.014326           |
| TRAF3    | 0.236773                                   | 0.019753           |
| VCP      | 0.13486                                    | 0.042419           |

FIG. 3. UPR pathway highlighting gene-expression differences between the low-risk and high-risk PSC clusters. The nodes in the pathway are shaded proportional to the $P$ value of differential gene expression between low-risk and high-risk PSC clusters. Genes with smaller $P$ values are darker. Differentially expressed genes associated with this pathway are provided in the table as an inset. The pathway diagram was generated using Ingenuity Pathway Analysis. Abbreviations: ASK1, mitogen-activated protein kinase kinase kinase 5; ATF4, activating transcription factor 4; ATF6, activating transcription factor 6; BCL2, BCL2 apoptosis regulator; BIP, heat shock protein family A (Hsp70) member 5; c/EBP, CCAAT Enhancer Binding Protein Beta; CALR, calreticulin; CANX, calnexin; CHOP, DNA damage inducible transcript 3; EDEM, ER degradation enhancing alpha-mannosidase like protein 1; EIF2α, eukaryotic translation initiation factor 2A; ERO1-Lβ, endoplasmic reticulum oxidoreductase 1 beta; GADD34, protein phosphatase 1 regulatory subunit 15A; GRP94, heat shock protein 90 beta family member 1; INSIG1, insulin induced gene 1; IRE1, endoplasmic reticulum to nucleus signaling 1; JNK1, mitogen-activated protein kinase 8; MBTPS, membrane bound transcription factor peptidase; M KK7, mitogen-activated protein kinase kinase 7; NRF2, “nuclear factor, erythroid 2 like 2”; P58IPK, DnaJ heat shock protein family (Hsp40) member C3; PDI, protein disulfide isomerase; PERK, eukaryotic translation initiation factor 2 alpha kinase 3; PPARγ, peroxisome proliferator activated receptor gamma; SCAP, SREBF chaperone; SREBP, sterol regulatory element binding transcription factor; TRAF2, TNF receptor associated factor 2; XBP1, X-box binding protein 1.
The UPR is a central cellular homeostatic pathway that becomes active when the protein folding capacity of the endoplasmic reticulum is exceeded. The resultant "ER stress" results in the toxic accumulation of unfolded or misfolded proteins. Three transmembrane protein sensors are involved in this homeostatic mechanism to detect unfolded proteins in the ER: PKR-like ER kinase (PERK), inositol-requiring enzyme 1α (IRE1α), and ATF6. Activation of PERK leads to inactivation of eIF2α, which blocks protein synthesis. ATF6 translocation from the ER to the Golgi apparatus and its subsequent proteolytic cleavage leads to nuclear translocation and transcriptional regulation of other genes that coordinate the UPR, including C/EBP homologous protein (CHOP) and X-box binding protein 1 (XBP1).

Previous experimental work has demonstrated associations between cholestasis and an abnormal UPR. However, whether the UPR serves to protect or worsen cellular injury in PSC remains unclear from these studies. Knockdown of CHOP reduces liver injury in the acute setting of biliary duct ligation, thereby suggesting a harmful role for the UPR in this setting. Bile duct ligation, however, represents a poor model of chronic cholestasis in humans. In contrast, a beneficial role for the UPR is suggested by its induction in response to chronic feeding of toxic BAs such as lithocholic acid, which increases hepatic BA synthesis.

| Gene*       | Log-Fold Change in High With High Risk of Clinical Events | Unadjusted P Value | Adjusted P Value |
|-------------|----------------------------------------------------------|-------------------|------------------|
| EIF3E       | -0.30195                                                 | 6.08E−09          | 0.000112         |
| ARHGAP5     | -0.22756                                                 | 2.44E−08          | 0.000184         |
| AGRN        | 0.495281                                                 | 3.00E−08          | 0.000184         |
| CEBPZOS     | -0.43928                                                 | 1.14E−07          | 0.000487         |
| CEP170B     | 0.310916                                                 | 1.47E−07          | 0.000487         |
| SBNO2       | 0.356272                                                 | 1.59E−07          | 0.000487         |
| EIF4A2      | -0.20602                                                 | 2.83E−07          | 0.000666         |
| NOTCH1      | 0.296809                                                 | 2.90E−07          | 0.000666         |
| GRK2        | 0.25553                                                  | 4.96E−07          | 0.001013         |
| HCFC1       | 0.236412                                                 | 5.81E−07          | 0.001068         |
| CEBPZ       | -0.19753                                                 | 7.88E−07          | 0.001315         |
| PIEZO1      | 0.371143                                                 | 9.00E−07          | 0.001378         |
| MCRS1       | 0.298438                                                 | 1.07E−06          | 0.001512         |
| ACAP3       | 0.31977                                                  | 1.29E−06          | 0.001621         |
| GOLGA3      | 0.298354                                                 | 1.41E−06          | 0.001621         |
| TECPR1      | 0.307899                                                 | 1.49E−06          | 0.001621         |
| FANCL       | -0.27802                                                 | 1.50E−06          | 0.001621         |
| AUH         | -0.31464                                                 | 1.60E−06          | 0.001631         |
| FLII        | 0.270343                                                 | 1.70E−06          | 0.001639         |
| MICALL1     | 0.338758                                                 | 1.95E−06          | 0.001704         |

Abbreviations: ACAP3, ArfGAP with coiled-coil, ankyrin repeat and PH domains 3; AGRN, agrin; ARHGAP5, Rho GTPase activating protein 5; AUH, AU RNA binding methylglutaconyl-CoA hydratase; CEBPZ, CCAAT enhancer binding protein zeta; CEBPZOS, CEBPZ opposite strand; CEP170B, centrosomal protein 170B; EIF3E, eukaryotic translation initiation factor 3 subunit E; EIF4A2, eukaryotic translation initiation factor 4A2; FANCL, FA complementation group L; FLII, FLII actin remodeling protein; GOLGA3, golgin A3; GRK2, G protein-coupled receptor kinase 2; HCFC1, host cell factor C1; MCRS1, microspherule protein 1; MICALL1, MICAL like 1; NOTCH1, notch receptor 1; PIEZO1, piezo type mechanosensitive ion channel component 1; SBNO2, strawberry notch homolog 2; TECPR1, tectonin beta-propeller repeat containing 1.

*Genes are sorted according to increasing order of P values.
the presence of advanced fibrosis in all groups.\(^{(27-29)}\)

Related to these findings, multiple components of the UPR have been shown to decrease with advancing age and may contribute to multiple aging-related diseases.\(^{(30)}\) Whether the age acceleration identified in PSC contributes to the functional declines in the UPR with aging remains unknown. Determining this relationship will require future investigation of

| Parameter                        | Low-Risk Subgroup (n = 24)       | High-Risk Subgroup (n = 16)       | PValue |
|----------------------------------|----------------------------------|----------------------------------|--------|
| BA, pg/mL                        | 2,470.8 (656.47, 7,245.2)        | 2,017 (794.98, 7,500.33)         | 0.70   |
| Total primary BA                 | 1,635.1 (633.16, 6,235.3)        | 1,713.7 (581.15, 6,726.35)       | 0.81   |
| Total secondary BA*              | 194.1 (34.79, 724.8)            | 202.8 (22.4, 457.4)             | 0.35   |
| Total conjugated BA*             | 2,163.2 (642.8, 7,064.8)        | 1,871.4 (604.3, 7,368.9)        | 0.66   |
| Total unconjugated BA*           | 116.0 (32, 237.0)               | 101.8 (31.8, 196.9)             | 0.51   |
| UDCA                             | 20.6 (5.8, 58.2)                | 19.3 (10.4, 66.6)               | 0.91   |
| TCDCA                            | 223 (33.3, 838)                 | 182 (42.7, 765.5)               | 0.77   |
| GCDCA                            | 719 (185, 1540)                 | 610 (198, 1,795)                | 0.81   |
| CA                               | 15 (3.3, 18.3)                  | 12.2 (6.0, 18.3)                | 1.00   |
| TCA                              | 110 (13.3, 850)                 | 152.5 (18.3, 1,424)             | 0.99   |
| GCA                              | 511 (80.2, 2,220)               | 449.5 (113.6, 2,410)            | 1.00   |
| LCA                              | 2.7 (2.5\(^{†}\), 18.2)        | 2.5\(^{†}\) (2.5\(^{†}\), 14)  | 0.62   |
| TLCA                             | 5\(^{†}\) (5\(^{†}\), 8.3)      | 5\(^{†}\) (5\(^{†}\), 9.31)     | 0.71   |
| GLCA                             | 11 (2.5\(^{†}\), 67)           | 9.2 (2.5\(^{†}\), 21.9)        | 0.53   |
| DCA                              | 53.3 (5\(^{†}\), 136)          | 45.7 (5\(^{†}\), 97.4)         | 0.76   |
| TDC6                             | 38.3 (5\(^{†}\), 76)           | 12.7 (5\(^{†}\), 57.4)         | 0.53   |
| GDCA                             | 78.2 (5.5, 307)                | 92.1 (2.5\(^{†}\), 156.5)      | 0.55   |
| C4, ng/mL                        | 6.7 (4.1, 24.7)                | 12.8 (4.6, 37.4)               | 0.36   |
| FGF19, pg/mL                     | 147.3 (59.7, 222.2)            | 93.3 (60.7, 191.7)             | 0.58   |

Note: Data are presented as median (IQR).

*UDCA and UDCA conjugates removed from the BA pools.

†Bile acid levels below the lower limit of quantification (LLOQ) of the Metabolon assay were assigned to the value corresponding to LLOQ.}

Abbreviations: CA, cholic acid; CDCA, chenodeoxycholic acid; DCA, deoxycholic acid; GCA, glycocholic acid; GCDCA, glycochenodeoxycholic acid; GDCA, glycodeoxycholic acid; GLCA, glycolithocholic acid; LCA, lithocholic acid; TCA, taurocholic acid; TCDCA, taurochenodeoxycholic acid; TDCA, taurodeoxycholic acid; TLCA, taurolithocholic acid.

FIG. 4. Hepatic expression and serum levels of IL-8 according to the subgroup of patients with PSC. Hepatic IL-8 expression was normalized by median log2 expression levels of IL-8 in the low-risk population. The boxplots were presented as median (IQR).
clinical therapeutics that improve outcomes in PSC and evaluating the effects on UPR responses and age acceleration.

The bioinformatics analysis approach in the present study led to several clinical insights including the primacy of fibrogenic pathways in PSC and the suppression of key constituents of the UPR, the homeostatic response to ER stress, in those patients at higher risk of PSC-related complications. Given the chronicity of PSC and the patient-specific genetic and immune-mediated factors that influence the PSC phenotype, we hypothesized that removing the influence of fibrosis would uncover potential mediators in the pathogenesis of PSC. In fact, the UPR has been postulated to be involved in the pathogenesis of injury in multiple preclinical models of cholestasis. Several limitations of the study including the single time-point analysis of gene expression, small sample size and number of clinical events, and the need for further validation in translational studies do not obviate, we believe, the central finding that the UPR is dysregulated in clinical PSC.

In conclusion, the current bioinformatics analysis of hepatic transcriptomic data from liver biopsies of patients with PSC removed the influence of fibrosis-related genes and identified two distinct patient subgroups with marked imbalance in clinical complications. Differentially expressed components of the UPR, including ATF6 in the high-risk patient subgroup, highlight potential deficiencies in the response to ER stress in those patients with PSC with a poor prognosis. These findings illustrate the potential of hepatic transcriptomic data to identify biologic drivers of PSC pathogenesis.

Author Contributions: Y.G., Z.J., C.C., and R.P.M. were responsible for the study design and data analysis. Y.G., C.C., R.P.M, and M.T. were responsible for drafting the manuscript. All authors reviewed the data and approved the final manuscript.

REFERENCES

1) Lazaridis KN, LaRusso NF. Primary sclerosing cholangitis. N Engl J Med 2016;375:2501-2502.
2) Karlsen TH, Fosleras T, Thorburn D, Vesterhus M. Primary sclerosing cholangitis—a comprehensive review. J Hepatol 2017;67:1298-1323.
3) de Vries EM, de Kruijger M, Farkkila M, Arola J, Schirmacher P, Gotthardt D, et al. Validation of the prognostic value of histologic scoring systems in primary sclerosing cholangitis: an international cohort study. Hepatology 2017;65:907-919.
4) Muir AJ, Levy C, Janssen HLA, Montano-Loza AJ, Shiffman ML, Caldwell S, et al. Sunitinib for primary sclerosing cholangitis: phase 2 study results with insights on the natural history of the disease. Hepatology 2019;69:684-698.
5) Karlsen TH, Boberg KM. Update on primary sclerosing cholangitis. J Hepatol 2013;59:571-582.
6) Karlsen TH. A lecture on the genetics of primary sclerosing cholangitis. Dig Dis 2012;30(Suppl. 1):32-38.
7) Lynch KD, Chapman RW, Keshav S, Montano-Loza AJ, Mason AL, Kremers AE, et al. Effects of vedolizumab in patients with primary sclerosing cholangitis and inflammatory bowel diseases. Clin Gastroenterol Hepatol 2020;18:179-187.e176.
8) Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 2017;14:417-419.
9) Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
10) Xu T, Le TD, Liu L, Su N, Wang R, Sun B, et al. CancerSubtypes: an R/Bioconductor package for molecular cancer subtype identification, validation and visualization. Bioinformatics 2017;33:3131-3133.
11) Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 2010;26:1572-1573.
12) Law CW, Chen Y, Shi W, Smyth GK. Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 2014;15:R29.
13) Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York, NY: Springer; 2000.
14) Liu X, Guo GL, Kong B, Hilburn DB, Hubchak SC, Park S, et al. Farnesoid X receptor signaling activates the hepatic X-box binding protein 1 pathway in vitro and in mice. Hepatology 2018;68:304-316.
15) Willy JA, Young SK, Stevens JL, Masouka HC, Wek RC. CHOP links endoplasmic reticulum stress to NF-kappB activation in the pathogenesis of nonalcoholic steatohepatitis. Mol Biol Cell 2015;26:2190-2204.
16) Vesterhus M, Holm A, Hvov JR, Nygaard S, Schrumpf E, Melm E, et al. Novel serum and bile protein markers predict primary sclerosing cholangitis disease severity and prognosis. J Hepatol 2017;66:1214-1222.
17) Zweers SJ, Shiryaev A, Komuta M, Vesterhus M, Hvov JR, Perugoria MJ, et al. Elevated interleukin-8 in bile of patients with primary sclerosing cholangitis. Liver Int 2016;36:1370-1377.
18) Trauner MH, Gindin Y, Zhou JZ, Hsu R, Chung C, Subramanian M, et al. RNA-Seq from patients with primary sclerosing cholangitis identifies a fibrosis-independent gene expression signature that predicts clinical events. Hepatology 2019;70:758A-759A.
19) Yoshida H. ER stress and diseases. FEBBS J 2007;274:630-638.
20) Dara L, Ji C, Kaplowitz N. The contribution of endoplasmic reticulum stress to liver diseases. Hepatology 2011;53:1752-1763.
21) Henkel AS, LeCuyer B, Olivesas S, Green RM. Endoplasmic reticulum stress regulates hepatic bile acid metabolism in mice. Cell Mol Gastroenterol Hepatol 2017;3:261-271.
22) Bochis IM, Rubins NE, White P, Furth EE, Friedman JR, Kaestner KH. Hepatocyte-specific ablation of Foxa2 alters bile acid homeostasis and results in endoplasmic reticulum stress. Nat Med 2008;14:828-836.
23) Liu X, Henkel AS, LeCuyer BE, Hubchak SC, Schipma MJ, Zhang E, et al. Hepatic deletion of X-box binding protein 1 impairs bile acid metabolism in mice. J Lipid Res 2017;58:504-511.
24) Ozcan U, Yilmaz E, Ozcan L, Furushashi M, Vaillancourt E, Smith RO, et al. Chemical chaperones reduce ER stress and
restore glucose homeostasis in a mouse model of type 2 diabetes. Science 2006;313:1137-1140.

25) Lawless MW, Greene CM, Mulgrew A, Taggart CC, O’Neill SJ, McElvaney NG. Activation of endoplasmic reticulum-specific stress responses associated with the conformational disease Z alpha 1-antitrypsin deficiency. J Immunol 2004;172:5722-5726.

26) Trauner M, Gindin Y, Jiang Z, Chung C, Subramanian GM, Myers RP, et al. Methylation signatures in peripheral blood are associated with marked age acceleration and disease progression in patients with primary sclerosing cholangitis. JHEP Rep 2020;2:100060.

27) Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14:R115.

28) Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. J Infect Dis 2015;212:1563-1573.

29) Loomba R, Gindin Y, Jiang Z, Lawitz E, Caldwell S, Djedjos CS, et al. DNA methylation signatures reflect aging in patients with nonalcoholic steatohepatitis. JCI Insight 2018;3:96685.

30) Brown MK, Naidoo N. The endoplasmic reticulum stress response in aging and age-related diseases. Front Physiol 2012;3:263.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.31488/suppinfo.