Fate and functional roles of Prominin 1\(^+\) cells in liver injury and cancer

Raymond Wu\(^1\), Stephanie Pan\(^1\), Yibu Chen\(^2\), Yasuhiro Nakano\(^3,5\), Meng Li\(^2\), Steven Balog\(^1\) & Hidekazu Tsukamoto\(^1,4\)*

Prominin 1 (PROM1) is one of a few clinically relevant progenitor markers in human alcoholic hepatitis (AH) and hepatocellular carcinoma (HCC), and mouse liver tumor initiating stem cell-like cells (TICs). However, the origin, fate and functions of PROM1\(^+\) cells in AH and HCC are unknown. Here we show by genetic lineage tracing that PROM1\(^+\) cells are derived in part from hepatocytes in AH and become tumor cells in mice with diethyl nitrosamine (DEN)-initiated, Western alcohol diet-promoted liver tumorigenesis. Our RNA sequencing analysis of mouse PROM1\(^+\) cells, reveals transcriptomic landscapes indicative of their identities as ductular reaction progenitors (DRPs) and TICs. Indeed, single-cell RNA sequencing reveals two subpopulations of Prom1\(^-\)Afp\(^-\)DRPs and Prom1\(^+\)Afp\(^+\)TICs in the DEN-WAD model. Integrated bioinformatic analysis identifies Discodin Domain Receptor 1 (DDR1) as a uniquely upregulated and patient-relevant gene in PROM1\(^+\) cells in AH and HCC. Translational relevance of DDR1 is supported by its marked elevation in HCC which is inversely associated with patient survival. Further, knockdown of Ddr1 suppresses the growth of TICs and TIC-derived tumor growth in mice. These results suggest the importance of PROM1\(^+\) cells in the evolution of liver cancer and DDR1 as a potential driver of this process.

Liver progenitor cells (LPCs) proliferate after massive parenchymal loss or/and suppressed hepatocyte proliferation in different liver diseases. However, the origin and fate of these cells in liver injury and cancer development are unclear. LPCs are identified by immunophenotype markers such as A6, SOX9, KRT19 (CK19), EpCAM and PROM1 and generated from biliary epithelial cells or hepatocytes through hepatocyte-to-ductal metaplasia. Tumor-initiating stem cell-like cells (TICs) also express the overlapping progenitor genes and uniquely possess self-renewing and tumor-initiating properties due to intrinsic TLR4-Nanog oncogenic pathway. Therefore, LPCs and TICs may share the same or similar lineage, but a definite relationship between them is currently unknown.

PROM1 is a well-known marker of TICs in many cancers and used to define cancer stemness and disease severity in patients. PROM1-positive (PROM1\(^+\)) cells also accumulate in mouse liver injury models and livers of alcoholic hepatitis (AH) patients in the pattern commonly characterized as ductular reaction progenitors (DRPs). The expression of PROM1 is also inversely associated with survival of AH patients. These clinical findings underscore the potential importance of PROM1\(^+\) cells in chronic liver disease and cancer. A study using mice bearing Prom1\(^+\) which expresses Cre recombinase and LacZ genes in the Prom1 locus, reports that PROM1\(^+\) cells are generated by proliferating ductular epithelial cells in response to chronic thioacetamide liver damage. Another study using the same genetic mice, reports that PROM1\(^+\) cells from neonatal liver have high regenerative capacity, which is almost undetectable in the adult liver. However, these PROM1\(^+\) cells reappear in adult mice upon feeding a diet containing 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC), and such livers express similar stemness transcriptomes as those in neonatal liver. Further, PROM1\(^+\) cells are transformed to tumor cells and give rise to more aggressive liver tumors upon DDC feeding when they are forced to express oncogenes or mutant tumor suppressors. Although HCC cells are not derived from cytokeratin 19 (Krt19) or osteopontin (Spp1) expressing biliary epithelial cells, they express higher levels of Prom1 compared to non-tumor tissues, suggesting the contribution of hepatocyte-derived PROM1\(^+\) cells in HCC formation.

1 Southern California Research Center for ALPD and Cirrhosis and Department of Pathology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. 2 USC Libraries Bioinformatics Services, University of Southern California, Los Angeles, CA, USA. 3 Center for Matrix Biology and Medicine, Graduate School of Medicine, Tokai University, Isehara, Japan. 4 Greater Los Angeles VA Healthcare System, Los Angeles, CA, USA. 5 Present address: Department of Developmental and Regenerative Biology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan. *email: htsukamo@usc.edu
Alcoholic liver disease (ALD) is one of the significant health-related financial burdens in the world\textsuperscript{11}. AH represents the most severe spectrum of ALD with high mortality\textsuperscript{12} and is presented as acute syndrome superimposed on chronic liver disease characterized by intense polymorphonuclear cell (PMN) inflammation, the appearance of DRPs and jaundice. One study suggests that these DRPs in AH patients reflect cholangiocyte proliferation\textsuperscript{13}. But it is largely unknown how DRPs emerge in AH or if and how they contribute to AH pathogenesis. To address these questions in an animal model, we developed a mouse model which reproduces histologic AH with DRP proliferation\textsuperscript{14}. The model also reproduces clinical features of AH including portal hypertension as judged by splenomegaly, hypoalbuminemia, and increased plasma bilirubin. More importantly, an expansion of PROM1+ cells expressing SPP1, NANOG, and AFP is detected in the model\textsuperscript{14}.

Understanding the mechanisms underlying how PROM1+ cells emerge and expand in AH mouse liver may reveal the potential pathogenic role of these cells in AH and liver cancer susceptibility. Unfortunately, epidemiology of liver cancer incidence in AH patients is lacking due primarily to the high mortality rate of AH patients and difficulties in long-term tracking and assessment of the patients after hospital discharge. However, heavy alcohol intake and Western diet are two top-ranked risk factors for both AH and liver cancer\textsuperscript{15}. In the present study, we aimed to determine the origin and molecular characteristics of PROM1+ cells in AH by genetic lineage tracing and RNA sequencing (RNAseq) analysis and to assess the fate of PROM1+ cells in liver tumor development in mice given Western alcohol diet (WAD) by genetic tracing and single-cell RNA sequencing (scRNAseq).

**Results**

**Transcriptome of PROM1+ cells from AH resembles known DRP and HCC gene signatures.** To characterize a gene expression profile of PROM1+ cells in AH, we isolated PROM1+ cells from AH mouse livers\textsuperscript{14} by FACS using antibodies against PROM1 and CD49f which were used as stem cell markers for mouse liver TICs\textsuperscript{16}. In normal mouse liver, a PROM1+/CD49f-CD45- population accounts for ~0.1% which increases to 1.93% in diethyl nitrosamine (DEN)-liver cancer model\textsuperscript{7}. The yield of PROM1+/CD49f+CD45- cells from AH mouse livers ranged from 1 to 4% of sorted cells (Fig. 1a, suggesting PROM1+ cells are increased in the liver of AH mice. To characterize PROM1+ cells, PROM1+/CD49f+/CD45- (R6) versus PROM1-/CD49f+/CD45- (R5) cells were sorted and total RNA was sequenced as shown in a flow chart in Fig. 1b. We also sorted PROM1-/CD49f- cells present in CD45+ population as hematopoietic PROM1+ cells (R8).

Total RNAseq analysis revealed a unique transcriptomic landscape of R6 PROM1+ cells as compared to R5 PROM1- or R8 hematopoietic PROM1+ cells (Fig. 1c). To identify differentially regulated genes in R6 (PROM1+) cells, we compared gene expression profiles of R6 versus R5 (PROM1-), R6 versus R8 (hematopoietic PROM1+), and R6 versus all CD45+ sorted cells. As shown in the Venn diagrams in Fig. 1d, this analysis identified 613 upregulated and 708 downregulated genes specific for PROM1+ liver cells (Suppl. Table 3 in Excel format). The upregulated gene list included twofold or higher expression of genes considered as markers for progenitors (Prom1, EpCAM, Sox9), biliary epithelial cells (Krt19, Hnf1b, Spf1) and mature hepatocytes (HNF4a, Alb), suggesting hepatic PROM1+ cells in AH are bipotent progenitor cells. Ingenuity Pathway Analysis of the PROM1+ cell transcriptome, also revealed significantly regulated stemness and tumorigenesis pathways such as pluripotency, EMT, tissue factor in cancer, Ephrin receptor, HIPPO pathway, and ILK signaling, suggesting their potential involvement in liver tumorigenesis (Fig. 1c).

**Lineage of PROM1+ cells in AH and their fate in liver cancer.** To determine how PROM1+ cells are generated in AH, we performed lineage tracing in the AH mouse model. To test the contribution of hepatocytes to the genesis of PROM1- cells, we utilized two approaches to label hepatocytes in Rosa26mTmG reporter mice before AH induction: the use of the conventional Albumin-Cre (Alb-Cre) and the more mature hepatocyte specific Transferrin-Cre (Ttr-Cre). Alb-Cre; Rosa26mTmG mice showed strong GFP expression by hepatocytes. Using these cells, we were able to establish the gating for GFP positive cells.

To generate Ttr-Cre mice, we injected AAV8-Ttr-Cre virus to Rosa26mTmG mice via tail vein and allowed 1 month for recombination to occur. The labeling efficiency was 98.8% for Alb-Cre mice and 99.2% for Ttr-Cre mice according to GFP positivity (Suppl. Figure 1b and 1c). To trace GFP+ hepatocytes to PROM1+ DRPs in the AH mice, we isolated CD45- nonparenchymal liver cells by a low speed centrifugation to remove hepatocytes before AH induction: the use of the conventional Alb-Cre and magnetic depletion of CD45+ cells. The cells from C57/BL6 mice after the AH protocol were used as a negative control because the Western diet regimen contributed to high background fluorescence (Suppl. Figure 1d). The cells were sorted from Alb-Cre; Rosa26mTmG mice (Suppl. Figure 1g) and Rosa26mTmG mice injected with AAV8-Ttr-Cre (Suppl. Figure 1j) for PROM1 and GFP. From the cells from AH liver of Alb-Cre; Rosa26mTmG mice, PROM1+/GFP+ cells were detected at 0.118% of frequency versus PROM1-/GFP- cells at 0.172%, suggesting 40.7% of PROM1+ cells are GFP-traced cells from Alb-Cre-based labeling in this representative mouse (Suppl. Figure 1g). Using Ttr-Cre labeling method which is considered more specific for hepatocytes\textsuperscript{17}, the frequency of GFP+ PROM1+ cells was shown to be 0.195% as compared to 0.691% of total PROM1+ cells with or without GFP expression, indicating 28.2% of the PROM1+ cells are hepatocyte-derived in this particular Ttr-Cre; Rosa26mTmG mouse subjected to AH (Suppl. Figure 1j, right panel). After repeating the same analysis in 4 AH mice, the average percentage of GFP+ PROM1+ cells was determined to be 24% (Suppl. Figure 1k). These results suggest that PROM1+ cells are generated in part from mature hepatocytes in AH and the origin appears heterogeneous. The latter notion is also consistent with the recent finding of the heterogeneity of PROM1+ cells in HCC\textsuperscript{19}.

A subpopulation of PROM1+ cells observed in HCC is suggested to possess tumor initiating properties\textsuperscript{4}. Indeed, the DRP markers such as Prom1, EpCAM, Krt19, and Sox9 detected in AH, were also induced in non-tumor liver tissue (NTL) of mice subjected to DEN injection and tumor promotion with WAD as compared to normal liver or tumor tissues, suggesting a more selective increase of DRPs in NTL (Fig. 2a). Indeed, immunofluorescent microscopy detected proliferating DRPs were concentrated in NTL around the peripheral border.
of the tumor mass as Ki67+ SOX9+ cells (purple arrows, Fig. 2b), revealing a close spatial association of DRPs with tumor cells.

To determine the fate and role of PROM1+ cells in liver tumorigenesis, we traced PROM1+ cells using the DEN-initiated and WAD-promoted liver tumor model in $Prom1^{C-L}$ mice, a strain engineered to have CreERT2 and LacZ genes in the $Prom1$ locus. While this mouse serves to express CreERT2 under the $Prom1$ promoter for labeling of PROM1+ cells, the homozygous strain of $Prom1^{C-L.C-L}$ can also be used as $Prom1$ knockout

Figure 1. Progenitor and tumorigenic characteristic of PROM1+CD49f+ cells isolated from mouse model of alcoholic hepatitis (AH). (a) FACS plots of PROM1+CD49f+ cells isolated from normal mice, AH mice and DEN-initiated liver tumor mice. For DEN mice, 2-week old C57/Bl6 mice were injected with 25 mg/kg of DEN intraperitoneally and fed with WAD for 5 months. (b) The flow chart shows how PROM1+CD49f+CD45− cells were isolated from mouse liver. (c) Hierarchical clustering of PROM1+CD49f+CD45− (R6), PROM1− CD49f+CD45− (R5), PROM1+CD49f+CD45− (R8) cells and CD45− liver cells, is shown using a differentially expressed gene (DEG) set specific for R6 versus R5. (d) Venn diagrams show the number of upregulated and downregulated genes specific for PROM1+ liver cells (CD45−). (e) Ingenuity pathway analysis of DEG specific for PROM1+ liver cells, reveals known stemness and tumorigenic pathways highlighted in red, are significantly regulated.
because it completely ablates the Prom1 expression due to the biallelic knock-in of the CreERT2-LacZ cassette. We crossed Prom1<sup>C-L/L</sup> mice with Rosa26mTmG mice to generate Prom1<sup>C-L/L</sup>; Rosa26mTmG (Prom1WT) and Prom1<sup>C-L/L</sup>;Rosa26mTmG (Prom1KO) mice, injected DEN at 2-week age and fed W AD from 5-week age for 6 months. Tamoxifen (Tam) was injected to label PROM1<sup>+</sup> cells at 3 weeks after DEN injection (Fig. 2c).

Liver tumor development in Prom1<sup>WT</sup> and Prom1<sup>KO</sup> mice was assessed via gross and histologic examination.
As predicted, Prom1 expression in tumors of Prom1<sup>−/−</sup>;Rosa26mTmG (Prom1KO) mice were almost undetectable by qPCR analysis, validating their knockout status (Suppl. Figure 2a).

Surprisingly, total liver tumor volumes calculated in these two genetic groups, were not different (Fig. 2d), indicating Prom1 is not essential in liver tumorigenesis. To determine the fate of PROM1<sup>+</sup> cells in tumor development, we stained the liver sections with visible tumors with anti-GFP and anti-KRT19 antibodies (Fig. 2f,g). While DRPs were detected as GFP<sup>+</sup> KRT19<sup>+</sup> cells in NTL region, tumor cells were also labeled with GFP but not for KRT19. These results suggest that PROM1<sup>+</sup> cells give rise to liver tumor cells even though PROM1 is not functionally required for tumor development.

**scRNAseq analysis reveals Prom1<sup>+</sup> TIC and DRP subpopulations in liver tumor.** To understand the molecular characteristics of Prom1<sup>+</sup> cells in the single cell resolution, we performed scRNAseq analysis. To enrich the LPCs and TICs, we have developed a unique approach of sorting Col1a1-expressing cells by FACS. The rationale for this approach is based on: (1) LPCs were previously reported to express high levels of Col1a1 expression<sup>1</sup>; (2) our own results confirmed 23-fold and 90-fold higher expression of Col1a1 in LPC line (PIL4) and TICs as compared to mouse primary hepatocytes, respectively (Suppl. Figure 3a); and (3) our additional data validated higher expression of Col1a1 in PROM1<sup>+</sup> cells versus PROM1<sup>−</sup> cells isolated via MACS from the DEN-WAD mouse livers (Suppl. Figure 3b). We used Col1a1-driven Cre recombinase expression to label LPCs and TICs in Col1a1Cre;Rosa26mTmG mice subjected to the DEN-WAD protocol. Tumor bearing livers were perfused and digested with collagenase to collect liver cells. Mature hepatocytes were removed by low speed centrifugation, and the remaining liver cells were subjected to FACS sorting based on ultraviolet (UV) and GFP. UV-excited gating was used to exclude vitamin A-storing hepatic stellate cells (HSC), the major Col1a1-expressing cell type in the liver. The UV-GFP<sup>+</sup> population was sorted as tumor-associated Col1a1-expressing cells and submitted for scRNAseq analysis (Fig. 3a). Barcode rank plot of the scRNAseq sample showed an excellent separation between the cell-associated barcodes and barcodes associated with empty partitions (Suppl. Figure 4). From a single cell library with the target cell number of 10,000, 8578 cells passed this quality control assessment. By this scRNAseq, we assessed co-expression of these genes with the cell-associated barcodes and barcodes associated with empty partitions (Suppl. Figure 4). From a single cell library, we identified two subpopulations with Prom1<sup>+</sup> cells (Cluster 1 and 2) (Fig. 3b) with no or minimal expression of the HSC marker Lrat, portal fibroblast marker Thy-1, hematopoietic marker Cd45, and endothelial cell markers Cd31 and Cd65 (Suppl. Figure 5a). As predicted, Cluster 1 and 2 cells expressed the biliary or/and progenitor markers such as Sox9, Spp1, Hnf1b, Krt19, Krtn23, Epica3 (Fig. 3c). To further validate these results, we assessed co-expression of these genes with Prom1 at the single cell level. As shown in Suppl Fig. 5b, Cluster 1 and 2 are enriched with cells which co-expressed Prom1 and Sox9, Spp1, Hnf1b, or Krtn23 and most Prom1<sup>+</sup> cells in both clusters did not co-express the HSC, portal fibroblast, hematopoietic, and endothelial cell markers, validating these Prom1<sup>+</sup> Cluster 1 and 2 cells are LPCs. The cells co-expressing Prom1 and Afp, the known marker of liver tumor cells, were detected primarily in Cluster 1 (Suppl Fig. 5a and 3a). In fact, Cluster 1 contained both Prom1<sup>+</sup>Afp<sup>+</sup> and Prom1<sup>+</sup>Afp<sup>−</sup> cells, suggesting TICs and tumor cells were enriched in this Cluster. Prom1<sup>+</sup>Afp<sup>+</sup> cells enriched in Cluster 2 are DRPs as the biliary/DRP markers such as Hnf1b, Krt19, Epica3, Mmp7 and Cfr were more selectively expressed in Cluster 2 cells (Fig. 3c).

We next assessed whether the known TIC/HCC transcripts are expressed in the scRNAseq profiles of the two clusters. For this, we assessed a set of genes reported to characterize cancer progenitor cells with tumor-initiating properties which were isolated from DEN-induced tumor bearing livers<sup>26</sup>. Many of the genes upregulated in these tumor initiating cells were expressed in our Cluster 1 cells including Tgf3, Tspan8, Gpc3, Ly6d, Cldn2 (Fig. 3c). These genes are involved in malignancies and cancer stemness<sup>21–23</sup>, further supporting the notion that Cluster 1 cells are TICs and tumor cells. We then acquired gene expression profiles unique to each of the three populations of Prom1<sup>+</sup>Afp<sup>+</sup>, Prom1<sup>+</sup>Afp<sup>−</sup>, and Prom1<sup>+</sup>Afp<sup>−</sup> cells by comparing among them by the locally distinguishing comparison feature of the Loupe Browser as depicted by heatmaps and volcano plots (Figs. 4a–d). As shown in the heatmaps, Prom1<sup>+</sup>Afp<sup>+</sup> and Prom1<sup>+</sup>Afp<sup>−</sup> cells share coordinately upregulated genes most likely reflecting their common tumorigenic function yet with some discordant gene expression which may reflect genetic differences between TICs with stemness property and tumor cells which have lost the stemness. In contrast, Prom1<sup>+</sup>Afp<sup>−</sup> cells which we consider as DRPs have a clearly distinct gene expression pattern from other two Afp<sup>+</sup> cell populations.

**Discoidin domain receptor 1 is a potential oncogenic driver of Prom1<sup>+</sup> cells.** To address the relevance of the transcriptomic data from Prom1<sup>+</sup> cells to AH livers in mice and patients, we compared our Prom1<sup>+</sup> cell RNAseq profile with a RNAseq profile of AH mouse liver and a previously reported proteomic profile of human AH liver<sup>24</sup>. This integrated analysis identified 7 upregulated and 7 downregulated genes that were common in Prom1<sup>+</sup> cells in both clusters did not co-express the HSC, portal fibroblast, hematopoietic, and endothelial cell markers, validating these Prom1<sup>+</sup> Cluster 1 and 2 cells are LPCs. The cells co-expressing Prom1 and Afp, the known marker of liver tumor cells, were detected primarily in Cluster 1 (Suppl Fig. 5a and 3a). In fact, Cluster 1 contained both Prom1<sup>+</sup>Afp<sup>+</sup> and Prom1<sup>+</sup>Afp<sup>−</sup> cells, suggesting TICs and tumor cells were enriched in this Cluster. Prom1<sup>+</sup>Afp<sup>+</sup> cells which we consider as DRPs have a clearly distinct gene expression pattern from other two Afp<sup>+</sup> cell populations.

Expression of 4 top genes uniquely expressed in each population are shown in Fig. 4e–g, and complete lists of the genes are provided in Suppl. Tables 4–6 in Excel format.

**Discoidin domain receptor 1 is a potential oncogenic driver of Prom1<sup>+</sup> cells.** To address the relevance of the transcriptomic data from Prom1<sup>+</sup> cells to AH livers in mice and patients, we compared our Prom1<sup>+</sup> cell RNAseq profile with a RNAseq profile of AH mouse liver and a previously reported proteomic profile of human AH liver<sup>24</sup>. This integrated analysis identified 7 upregulated and 7 downregulated genes that are common in Prom1<sup>+</sup> cells, mouse and human AH livers (Fig. 5a). Since our results showed Prom1<sup>+</sup> cells gave rise to liver tumor cells, we looked for potential oncogenic drivers unique to Prom1<sup>+</sup> cells by first screening the expression of the 7 upregulated genes in NTL tissues of DEN/WAD mice compared to normal liver. This analysis revealed that Prom1<sup>+</sup> Ddr1, Ddr2, Lamc2, Lamc3 and Pkhd1 were induced significantly in DEN-WAD livers (Fig. 5b). Next, we examined the expression of these genes in cell clusters revealed by our scRNAseq analysis. Of these genes examined, Ddr1 and Pkhd1 expressions were detected in Cluster 1 and 2 (Fig. 5c). DDR1 is known as a receptor for collagen<sup>27</sup> and implicated in oncogenesis, metastasis and chemoresistance of cancers of pancreas, lung, prostate, ovary, and colon<sup>28–29</sup>. Ddr2, in contrast, is expressed in clusters of activate Lrat<sup>+</sup> HSC and Thy1<sup>+</sup> portal fibroblasts (Fig. 5c and Suppl. Figure 5). Sptn1 and Elyf were expressed more selectively in DRPs in Cluster 2 (Fig. 5c), suggesting their importance in the biology of Prom1<sup>+</sup> DRPs.

To confirm Ddr1 upregulation in AH in patients in our hands, we analyzed Ddr1 mRNA levels in liver tissues of AH patients versus healthy subjects received from the explant repository program of John Hopkins University. We found that Ddr1 mRNA levels were significantly higher in AH patients compared to healthy controls (Fig. 5d). These results suggest that Ddr1 is a potential oncogenic driver of Prom1<sup>+</sup> cells in liver tumors.
University. Indeed, this analysis revealed at least 20-fold higher DDR1 expression in AH livers compared to the normal liver (Fig. 6a). DDR2 level was also statistically higher in AH tissues than normal liver, but the fold induction was much lower than DDR1. To address the role of DDR1 in human liver cancer, we performed the data mining approach. This bioinformatic analysis showed that DDR1 expression was significantly elevated in both HCC30 (Fig. 6b) and cirrhotic livers31 (Fig. 6c) compared to normal livers.

Patients with combined HCC-intrahepatic cholangiocarcinoma (ICC) have poor prognosis33. ICC is characterized by higher expression of liver progenitor markers such as EpCAM and KRT19, suggestive of expanded TICs and/or DRPs. Indeed, DDR1 expression is higher in patients with ICC or combined HCC-ICC compared to those with HCC alone (Fig. 6d)34.

Figure 3. Identification of Prom1+ TICs and Prom1+ DRPs subpopulations from DEN mouse liver by scRNAseq. (a) The FACS plot depicting the sorting of UV- GFP+ cells from DEN-WAD mouse liver cells. Non-parenchymal cells were isolated by collagenase perfusion from Col1a1Cre;Rosa26mTmG mice subjected the DEN-WAD regimen. The cells were sorted by FACS to remove UV+ hepatic stellate cells and to collect UV- GFP+ cells as Col1a1 expressing and vitamin A negative cells. The scRNAseq library was prepared immediately after sorting. (b) The analysis detected Prom1‘Afp’, Prom1‘Afp+’ and Prom1‘Afp+’ populations. The plots were created in the t-SNE format using filters of Prom1 and Afp transcripts. (c) Co-expression analysis of Prom1+ cells with known markers of DRPs, TICs, tumor. The plots were created in the t-SNE format with gene/feature expression in Loupe browser. The scale bar represents log2-transformed normalized expression.
Next we assessed the relationship of DDR1 and PROM1 expression with the HCC patient survival, by analyzing in TCGA HCC cohort data using UCSC Xena database. We stratified the cohort into DDR1 high and low expression groups. The Kaplan–Meier analysis showed that DDR1 high patient group had lower survival than the DDR1 low patients ($p < 0.039$) (Fig. 6e). When PROM1 was used to stratify the cohort, PROM1 high patients had a significantly worse outcome in their survival compared to the PROM1 low group ($p < 0.021$) (Fig. 6f). When PROM1 and DDR1 co-expression was analyzed, 23 patients with high expressions of both PROM1 and DDR1 had a worst outcome compared to 69 patients with both genes expressed at the low level ($p < 0.014$) (Fig. 6g). The difference in survival in these patients was greatest (> 1825 days). In summary, DDR1 is induced in human cirrhosis, HCC, and ICC and inversely associated with the survival of HCC patients. In addition, the expression of DDR1 is correlated with PROM1 expression and predicts the survival of HCC patients in this TCGA cohort. These data suggest that DDR1 expressing PROM1+ DRPs and TICs may determine the clinical course of HCC patients.

Figure 4. Novel transcript profiles specific for Prom1+Af+ , Prom1+Af− and Prom1−Af+ populations. (a) The graph-based clustering of DEG sets for Prom1+Af+ , Prom1+Af− and Prom1−Af+ cells. (b–d) volcano plots of transcript profiles from the three subpopulations. (e–g) Top representative novel transcripts identified by scRNAseq for each subpopulation are shown in t-SNE format as generated by the Loupe browser. The scale bar represents log2 transformed normalized expression.
Functional significance of DDR1 in TIC-derived tumor development. To determine the functional significance of DDR1 expressed in PROM1+ cells in liver oncogenesis, we first measured Ddr1 expression in TICs, PIL4 liver progenitors and mouse primary hepatocytes. TICs were isolated from a mouse model of alcohol-promoted liver cancer. PIL4 cells were previously isolated from p53-null mouse fed with choline-deficient, ethionine-supplemented diet. TICs have self renewing and tumor-initiating activities while PIL4 cells lack these properties but mimic DRPs. As expected, albumin is expressed exclusively by primary hepatocytes but not by TICs and PIL4 cells while Krt19 (CK19) is expressed by PIL4 cells. Tlr4 which is induced in stem cell-like

![Figure 5](https://www.nature.com/scientificreports/) Unique upregulation of Discoidin Domain Deceptor 1 (Ddr1) in PROM1+ cells. (a) Integrated analysis of a RNAseq profile of PROM1+ cells from mouse AH, a RNAseq profile of the liver of mouse AH, and a proteomic profile of human AH versus control liver, reveals 7 upregulated and 7 downregulated genes common in all three comparisons. (b) Quantitative PCR analysis of DEN-WAD mouse livers versus normal livers, reveals that some of the upregulated genes shown by the integrated bioinformatic analysis are also induced in the liver of the mouse liver tumor model including Ddr1. *p < 0.05 based on non-parametric student's t-test. (c) Expression of Ddr1, Ddr2, Spin1, Ehf, Lamlc3 and Pkhd1 at the single cell level are shown in the t-SNE format generated by Loupe browser. The scale bar represents log2 transformed normalized expression.
cells, is expressed moderately by PIL4 cells and most conspicuously by TICs along with Nanog, Sox2 as previously reported (Suppl. Figure 6). A pattern of Ddr1 expression followed that of Tlr4: undetectable in primary hepatocytes, moderately increased in PIL4 cells, and highest expression (~40-fold) in TICs. When Ddr1 but not Ddr2 was selectively knocked down 80% in TICs by using two separate shRNAs (Ddr1-sh1 and Ddr1-sh2) versus scrambled shRNA (Scr-sh) (Fig. 7c), TIC growth was significantly impaired as shown by crystal violet staining and cell count (Fig. 7a,b). Further, the treatment with the DDR1 inhibitor (DDR1in7rh), suppressed the growth of TICs at IC50 of 0.95 µM and that of the liver cancer cell line DIHXY at IC50 of 1.55 µM while PIL4 cells were four–fivefold less sensitive (Fig. 7d). These results suggested that DDR1 has a functional significance in TIC and liver cancer cell growth. To test this potential oncogenic role of DDR1 in vivo, we performed a xenograft experiment in nude mice transplanted with TICs transduced with Ddr1-sh1 versus Scr-sh. Ddr1 knockdown in TICs significantly attenuated TIC-derived tumor growth to one third of that with TICs with Scr-sh, supporting the functional role of DDR1 in tumor development (Fig. 7e).

Figure 6. DDR1 upregulation in liver cancer associated with poor patient survival. (a) DDR1 and DDR2 expression levels were measured by qPCR in normal human livers versus alcoholic hepatitis (AH) livers. * p < 0.05 based on non-parametric student's t-test. (b,c) DDR1 expression in HCC and cirrhotic human livers were compared to normal livers using the existing microarray data from Oncomine database. (d) DDR1 expression was compared in intrahepatic cholangiocarcinoma (ICC), combined HCC and ICC (HCC-ICC) and HCC in patients. * p < 0.05 and ** p < 0.01 based on the Oncomine database. (e–g) Kaplan–Meier analysis of TCGA HCC cohort data using UCSC Xena, reveals an inverse relationship between the expression of DDR1, PROM1 or both with patient survival. p values were calculated by log-rank test.
Discussion

Our study provided several new findings on the origin, fate and functional role of PROM1+ cells in alcohol associated liver injury and tumorigenesis. Firstly, using lineage tracing, we showed hepatocytes are in part the origin of PROM1+ cells in AH. Secondly, our RNAseq analysis revealed these PROM1+ cells have a unique gene expression profile consistent with their characteristics as LPCs/DRPs or TICs/tumor cells. Thirdly, in the DEN-initiated, WAD-promoted liver tumor model, Prom1+ cells gave rise to DRPs and more importantly tumor cells. Yet, biallelic Prom1 knock-in mice which lack Prom1 expression, still developed liver tumors as WT mice, suggesting that PROM1 is a pivotal marker for DRPs/TICs which generate tumor cells while PROM1 itself is not required for tumor development. Fourthly, we have developed a unique method of enriching LPCs and TICs using Col1a1-based labeling and FACS. This allowed us to identify two Prom1+ subpopulations by scRNAseq.
analysis which we believe represent DRPs and TICs based on their similar yet distinct gene expression profiles. Fifthly, our integrated bioinformatic analysis of RNAseq and scRNAseq data disclosed Ddr1 as a gene uniquely upregulated in PROM1+ cells, and our in vitro and in vivo loss-of-function approaches revealed the oncogenic role of Ddr1 in TICs. Finally, the clinical relevance and importance of Ddr1 was supported by Ddr1 induction in patient HCC and its significant association with low patient survival.

The fate of LPCs has actively been investigated in different mouse models and generated conflicting or mixed results. Lineage tracing studies showed that cholangiocytes defined by KRT19+ or SOX9+ cells rarely gave rise to hepatocytes. In line with this observation, KRT19+ or SOX9+ cells did not develop HCC in mice, further suggesting that cholangiocyte and hepatocyte lineages may not merge. However, recent studies using dual lineage tracing of cholangiocytes and hepatocytes revealed that a significant portion of SOX9+ LPCs differentiated into both cholangiocytes and hepatocytes. Moreover, KRT19+ cholangiocytes became hepatocytes when hepatocyte proliferation was impaired by p21 overexpression or feeding mice with choline deficient ethionine supplemented diet, methionine and choline deficient diet or DDC containing diet. In contrast to these studies which utilized Sox9, Krt19 and Spp1 as a driver of LPC specific labeling, we used Prom1-Cre to determine the fate of LPCs/DRPs and TICs in liver tumor development based on their shared expression of Prom1. Our results revealed that PROM1+ cells which were derived in part from hepatocytes in AH, gave rise to both DRPs and liver tumor cells in the DEN-WAD mouse model. Because of TICs known tumor-initiating activity, labeled TICs were presumed to have generated tumor cells while labeled DRPs must have persisted or expanded in the model. However, we did not observe labeled hepatocytes, suggesting no generation of hepatocytes from PROM1+ cells in this model. Our method could not allow selective fate mapping of DRPs versus TICs to clarify their relative contribution to generation of tumor cells. Thus, we cannot rule out a possibility of PROM1+ DRPs giving rise to tumor cells either directly or through TICs. From a list of genes shown to be uniquely expressed in Prom1+ Afp+ cells versus Prom1+ Afp- cells, we may be able to devise a method of selective TIC or DRP labeling in future.

Intratumoral heterogeneity presents a unique challenge for understanding of HCC development and therapy resistance. Like other cancers, HCC comprises diverse tumor clones that develop via clonal evolution. HCC in patients with expression of tumor stem cell markers such as EpCAM, PROM1 and KRT19 tend to be aggressive and chemoresistant. Recent single cell genomics has identified an EpCAM + subpopulation with the stemness property with CTSE as an oncogene uniquely induced in this subpopulation. In our study, Ddr1 was identified as a tumor driver gene in PROM1+ subpopulation relevant to both mouse and patient liver cancer. In particular, patients with combined HCC-ICC having co-overexpression of Ddr1 and PROM1 had the worst survival, suggesting the contribution of PROM1+ TICs to this clinical phenotype. We also identified a group of genes uniquely upregulated in Prom1+ Afp+ cells which obviously deserve further attention. Of interest, three of the 4 top genes expressed in this subpopulation (Fig. 4e), are metabolic genes: Gulo, beta-catenin responsive gene which encodes the limiting enzyme in the final step in ascorbic acid biosynthesis, Cypr2c70, a key enzyme in generation of the murine specific bile acid, a-muricholic acid, or Ellov2, an enzyme involved in elongation of long chain w-3 and w-6 fatty acids which was shown to confer stemness of glioma cells. Prom1+ Afp+ cells may constitute therapy-resistant TICs in liver cancer and are a logical and attractive therapeutic target. In fact, in cancer of other organs such as lung, breast, colon, and prostate, DDR1 is already considered as a therapeutic target. A strategy to selectively direct interventions to key driver genes such as Ddr1 in PROM1+ TICs may be desired for the therapy for liver cancer. Finding a molecular link between liver injury and liver carcinogenesis has been a topic of interest. Recently, the DAMP high mobility group box 1 (HMGB1) is shown to underlie hepatocyte-to-ductal metaplasia in liver injury and HCC development in mice. However, the molecular mechanism involved in the potential transition of LPCs to TICs or tumor cells are elusive. Our findings suggest that Ddr1 may at least in part facilitate a potential tumorigenic link of PROM1+ cells to liver cancer (Fig. 7f). In fact, our study is first to show that Ddr1 is essential for TIC growth and TIC-derived tumor formation. Future studies are obviously needed to determine how Ddr1 participates in the transition of PROM1+ cells to liver tumor cells in vivo. In summary, our study provides the evidence for hepatocyte-origin of PROM1+ cells in AH and the functional role of such cells in liver tumorigenesis (Fig. 7f).

Methods

Animals and animal models. All animal experiments were performed by the Animal Core of the Southern California Research Center for ALPD and Cirrhosis in accordance to an approved Institutional Animal Care and Use Committee (IACUC) protocol of University of Southern California (USC). C56/BL6, Prom1C-L (Prom1tm1(Z/ErZ2Giy)) and Rosa26mTmG (Gt(ROSA)26Sortm4(CAG-EGFP)Jaro) were purchased from the Jackson Laboratory and crossed to generate Prom1C-L; Rosa26mTmG. For a model of alcoholic liver injury and liver carcinogenesis, we used Donryu rats (Dyres) which were also superimposed with weekly binge as previously described. For DEN-initiated, WAD-promoted liver tumorigenesis, two-week old male mice were injected with 25 mg/kg of DEN intraperitoneally (Sigma Aldrich, St Louis). For lineage tracing in Prom1C-L; Rosa26mTmG mouse, the mice were intraperitoneally injected with two doses of 75 mg/kg tamoxifen (Sigma Aldrich, St Louis) at the 3-day interval at 3 weeks after DEN injection. Western alcohol liquid diet (WAD) (#710362; Dyets Inc.) was introduced to the mice at 5 weeks old by incrementally increasing alcohol concentration in the diet, reaching the maximal concentration of 3.5% (v/v). After 5–6 months of WAD feeding, the mice were euthanized, the livers were carefully excised,
weighted and visible tumors were counted. Each tumor length (L) and width (W) were measured by a caliper and the tumor volume was calculated using the formula of $(L \times W^2)/2$. For TIC xenograft transplantation model, NCr nude mice (Taconic Biosciences) were subcutaneously transplanted with $1 \times 10^5$ TICs transduced with shRNA against Ddr1 or scrambled shRNA, and TIC-derived tumor growth was monitored for 2 weeks by measuring the tumor length and width for assessment of tumor volume as described above.

**Cell culture.** TICs were provided by the Integrative Liver Cell Core (ILCC) of USC and cultured in 1:1 mixture of Dulbecco’s modified Eagle medium (DMEM) and Nutrient Mixture F-12 media (Life Technologies, Camarillo, CA) with 10% fetal bovine serum (FBS), 20 ng/mL murine endothelial growth factor (EGF; Peprotech, Rocky Hill, NJ), Embryomax nucleosides (MilliporeSigma, Burlington, MA), 100 nM dexamethasone (Sigma–Aldrich, St. Louis, MO), and antibiotic-antimycotic (Life Technologies) as described before. DHX5 cells were cultured in DMEM medium with 10% fetal bovine serum (FBS) and antibiotic-antimycotic (Life Technologies). PIL-4 cells were obtained from Dr. George Yeoh of University of Western Australia and cultured in Williams’ E medium supplemented with 10% FBS, EGF (Peprotech), IGF-II (Peprotech) and human Insulin (Sigma–Aldrich, St. Louis). The DDR1 inhibitor 7rh was purchased from Sigma–Aldrich for the treatment of TICs in culture. MTT assay and crystal violet assay were performed as described before.

**Alcoholic hepatitis patient and healthy liver tissues.** AH patient and control human livers were obtained from the Clinical Resources for Alcoholic Hepatitis Investigators (1R24AA025017) of Johns Hopkins University as previously described. The patient samples were obtained with informed consent in accordance with Johns Hopkins Hospital Institutional Review Board as described before. These tissues were excised from explanted livers in patients with severe AH during liver transplantation, or wedge biopsies from the donor livers (normal control). Suppl. Table 1 summarizes basic clinical information of 4 AH patients and 4 healthy donors from which the samples we analyzed were collected.

**Immunohistochemistry and immunofluorescence staining.** The liver slices were fixed in 4% paraformaldehyde (PFA) at 4°C overnight and transferred to 70% ethanol on the next day. The fixed tissues were paraffin-embedded, and paraffin sections were stained with hematoxylin & eosin (H&E). For immunofluorescence staining (IF), the liver slices were fixed in 4% PFA at 4°C for 2 h, incubated in 10% sucrose in PBS for 2 h and 30% sucrose in PBS for overnight. The tissues were then frozen in the freezing medium and sectioned with a Cryostat (CM1900; Leica, Buffalo Grove, IL) to prepare 7 µm thickness cryosections. The sections were air-dried overnight, treated with 3% H$_2$O$_2$ in methanol and permeabilized with proteinase K prior to blocking with the solution containing 5% donkey serum and 0.2% bovine serum albumin for 30 min. The sections were then incubated with primary antibodies against Green Fluorescent Protein (GFP; Thermo Fisher Scientific) for 30 min. The sections were counterstained with AlexaFluor 550 (400-fold dilutions, Thermo Fisher Scientific) for 30 min. The sections were counterstained with DAPI. Images were captured with a Nikon 90i microscope and DS-Fi1 digital camera (Nikon, Melville, NY). For IF shown in Fig. 3b, the liver section was stained for HNF4a (50-fold dilution, #6556, Santa Cruz Biotechnology), K67 (50-fold dilution, #652,402, Biolegend), SOX9 (500-fold dilution, #AB5535, Millipore) as previously described. Images were captured with a confocal microscope.

**FACS isolation.** Isolation of PROM1+ cells was performed by the ILCC of USC. The liver was perfused with EMEM medium, minced on a 10 cm dish and resuspended in collagenase IV containing digestion medium (25 mg of collagenase type IV (C5138-Sigma Aldrich) in 50 mL of serum free DMEM-F12, pH 7.35–7.4). Hepatocytes were removed by low speed centrifugation. CD45+ cells were depleted with CD45 microbeads (Miltenyi Biotec). The cells were blocked with FcR blocking reagent and stained with the corresponding antibodies. CD133-APC, CD49f-PE and CD45-efluor450 (eBioscience). The single controls and Fluorescent Minus One (FMO) were added in each experiment to gate the positive population. PI was added to separate live cells from dead cells. The cells were sorted in MoFlo Cell Sorter (Beckman Coulter). Isolation of Col1a1 expressing cells was carried out by using the protocol previously described. A non-parenchymal liver cell fraction from Col1a1-Cre;Rosa26tm1Mcg mice was collected by pronase-collagenase perfusion and gradient ultracentrifugation. Gating for UV and GFP positivity were set by primary hepatic stellate cells isolated from wild type mouse and Col1a1Cre;Rosa26tm1Mcg mouse, respectively, with wild type liver myofibroblasts as a UV and GFP negative cells. UV-GFP+ fraction was collected for scRNAseq analysis.

**RNA sequencing.** The cells were centrifuged and total RNA was isolated with PicoPure RNA isolation kit (Thermo Fisher Scientific). Prior to library construction RNA integrity was verified by Experion analysis (Bio-Rad). PolyA selection was carried out using Illumina Truseq V2 polyA beads. Sequencing was performed on a NextSeq 500 using V2 chemistry. RNA-seq data was analyzed using the RNA-seq workflow in Partek Flow software (v3. Partek Inc., St. Louis, MO, USA). Briefly, the raw sequencing reads were first trimmed based on the quality score (Phred QC > 20, min read length = 25 nt) before mapped to mouse genome build mm10 using TopHat version 2.0.8 with default parameter settings and using Gencode M3 annotation as guidance. Gencode M3 was then used to quantify the aligned reads to genes using Partek E/M method. The gene level aligned read counts were normalized for all samples using Upper Quartile normalization before subjected to differential expression analysis using Partek Gene Specific Analysis method (genes with fewer than 10 aligned reads in any sample among a data set were excluded). The differentially expressed gene (DEG) lists were generated using the cutoff of FDR < 0.05 and fold changes greater than 2 either direction.
Single-cell RNA sequencing. Single cell was prepared using the 10 × Single Gene Expression v3 with the 10 × Chromium technology following manufacturer protocol as described before31,33 (cat# 1000154 and 1000074, 10 × Genomics). The library was prepared immediately after cell sorting with the target cell number of 10,000. The library was validated on the Agilent TapeStation (Agilent Technologies, Palo Alto, CA, USA), and quantified by using Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA) as well as by quantitative PCR (KAPA Biosystems, Wilmington, MA, USA). The sequencing libraries were clustered on 8 single lanes of a flowcell. After clustering, the flowcell was loaded on the Illumina HiSeq instrument according to manufacturer's instructions. The samples were sequenced using a 2 × 150 Paired End (PE) configuration. Raw sequence data (.bcl files) generated from Illumina HiSeq was converted into fastq files and de-multiplexed using the 10 × Genomics’ cellranger version 3.1.0 by the Cell Ranger mkfastq command with default parameters. Subsequently, UMI and cell barcode de-convolution and mapping to the mm10 reference genome was performed with 10 × Genomics’ Cell Ranger software package (version 3.1.0) to generate the final digital gene expression matrices and cloupe files using the Cell Ranger count command with default parameters. The quality of data is provided as a barcode rank graph (Suppl. Figure 4). Of targeted 10,000 cells, 8578 cells were subjected for sequencing. After normalization, 38,901 mean reads per cell and 2776 median genes per cell were detected. Data analysis was performed with 10 × Genomics’ Loupe Browser software. Raw data files were uploaded to GEO database with accession number GSE152748.

Quantitative PCR. RNA extraction was performed using Quick RNA miniprep kit (Zymo Research). Tissues were homogenized in RNA lysis buffer. RNA was extracted according to the manufacturer’s protocol. 500–1000 ng of RNA was used as input to synthesize cDNA using high-capacity cDNA reverse transcription kit (ThermoFisher Scientific). Q-PCR analysis was performed using the primers shown in Suppl. Table 1 and Sybr Green reagents (ThermoFisher Scientific) in Viia7 instrument (Applied Biosystem).

Lentiviral mediated Ddr1 knockdown. Ddr1 mission shRNA were purchased from Sigma Aldrich. Two shRNAs with the following sequences were chosen: Ddr1shRNA1 (CCGGAGATCCATCTAGGTGA-TAT CTCGAGATCCATCTAGGTGAATCATTCTTTT) and Ddr1shRNA2 (CCGGGC-CAGCCTGAACCTTGTG GCATCTCGAGATGCCAAAAGTTCAGCCTGCGTTTTT). The shRNAs were packaged into lentivirus with psPAX2 and pMD2.G in 293FT cell line. The lentiviruses were purified using speedy lentivirus purification kit (LV999, Applied Biological Materials Inc., Canada) and titrated using qPCR lentivirus titration kit (LV900, Applied Biological Materials Inc., Canada). The cells were infected with 10 MOI of virus and cultured in 10 μM of puromycin (Invitrogen) after 48 h of virus infection.

Bioinformatic analysis. Bioinformatic analysis was performed by using University of California Santa Cruz (UCSC) Xena program. TCGA liver cancer cohort was selected with the filters: primary tumors and hepatocellular carcinoma as primary disease. DDR1 and PROM1 expression were extracted from the gene expression generated by IlluminaHiSeq.

Data analysis. RNAseq data was analyzed by Partek Flow (St. Louis, MO) and performed by USC libraries Bioinformatics Services. Pathway analysis was performed by Ingenuity Pathway Analysis (Qiagen, Redwood City, CA). Statistical Analysis was performed by Graphpad Prism (San Diego, CA) and Microsoft Excel.

Received: 1 June 2020; Accepted: 18 October 2020
Published online: 10 November 2020

References
1. Miya1jima, A., Tanaka, M. & Itoh, T. Stem/progenitor cells in liver development, homeostasis, regeneration, and reprogramming. Cell Stem Cell 14, 561–574. https://doi.org/10.1016/j.stem.2014.04.010 (2014).
2. Michalopoulos, G. K. & Khan, Z. Liver stem cells: experimental findings and implications for human liver disease. Gastroenterology 149, 876–882. https://doi.org/10.1053/j.gastro.2015.08.004 (2015).
3. Tarlow, B. D. et al. Bipotent adult liver progenitors are derived from chronically injured mature hepatocytes. Cell Stem Cell 15, 665–618. https://doi.org/10.1016/j.stem.2014.09.008 (2014).
4. Chen, C. L. et al. Reciprocal regulation by TLR4 and TGF-β in tumor-initiating stem-like cells. J. Clin. Invest. 123, 2832–2849. https://doi.org/10.1172/JCI65859 (2013).
5. Ma, S. et al. miR-130b Promotes CD133(+) liver tumor-initiating cell growth and self-renewal via tumor protein 53-induced nuclear protein 1. Cell Stem Cell 7, 694–707. https://doi.org/10.1016/j.stem.2010.11.010 (2010).
6. Zhao, Q. et al. Prognostic value of the expression of cancer stem cell-related markers CD133 and CD44 in hepatocellular carcinoma: from patients to patient-derived tumor xenograft models. Oncotarget 7, 47431–47443. https://doi.org/10.18632/oncotarget.10164 (2016).
7. Sancho-Bru, P. et al. Liver progenitor cell markers correlate with liver damage and predict short-term mortality in patients with alcoholic hepatitis. Hepatology 55, 1931–1941. https://doi.org/10.1002/hep.25614 (2012).
8. Kaminoto, K. et al. Heterogeneity and stochastic growth regulation of biliary epithelial cells dictate dynamic epithelial tissue remodeling. Elife. https://doi.org/10.7554/eLife.15034 (2016).
9. Zha, L. et al. Multi-organ mapping of cancer risk. Cell 166, 1132–1146.e1137. https://doi.org/10.1016/j.cell.2016.07.045 (2016).
10. Mu, X. et al. Hepatocellular carcinoma originates from hepatocytes and not from the progenitor/biliary compartment. J. Clin. Invest. 125, 3891–3903. https://doi.org/10.1172/JCI77995 (2015).
11. Seitz, H. K. et al. Alcoholic liver disease. Nat. Rev. Dis. Primers 4, 16. https://doi.org/10.1038/nrdp0014-7 (2018).
12. Lucey, M. R., Mathurin, P. & Morgan, T. R. Alcoholic hepatitis. N. Engl. J. Med. 360, 2758–2769. https://doi.org/10.1056/NEJMoa0805786 (2009).
24. Khanova, E.
25. Vogel, W., Gish, G. D., Alves, F. & Pawson, T. The discoidin domain receptor tyrosine kinases are activated by collagen. 
26. Yang, J. C.
28. Azizi, R., Salemi, Z., Fallahian, F. & Aghaei, M. Inhibition of discoidin domain receptor 1 reduces epithelial-mesenchymal transition and induce cell-cycle arrest and apoptosis in prostate cancer cell lines. 
17. Malato, Y.
16. Machida, K.
34. Woo, H. G.
35. Dumble, M. L., Croager, E. J., Yeoh, G. C. & Quail, E. A. Generation and characterization of p53 null transformed hepatic progenitors and cancer.
14. Lazaro, R.
13. Dubuquoy, L.
33. Xue, R.
36. Tarlow, B. D., Finegold, M. J. & Grompe, M. Clonal tracing of Sox9+ liver progenitors in mouse oval cell injury.
38. Font-Burgada, J.
39. Han, X.
49. Page, A.
46. Ambrogio, C.
47. Hernandez, C.
45. Wu, R.
129–140. https://doi.org/10.1002/hep.27978 (2018).
10.1016/j.hep.2014.09.033 (2015).
10.1016/j.jhep.2014.09.033 (2015).
2436–2451. https://doi.org/10.1172/JCI91786 (2018).
Hepatol. Commun. 6, 388–397. https://doi.org/10.1038/ncomms14049 (2017).
9. (9), 1248–1267. https://doi.org/10.1158/2159-8290.CD-19-0061 (2020).
274, 155, 131 (6), 1190–1203. https://doi.org/10.1038/s41598-020-76458-8
Acknowledgements
The authors thank for the services rendered by the Animal Core of the Southern California Research Center for ALPD and Cirrhosis funded by the NIAAA/NIH (P50AA011999). We also thank Sonal Sinha for DEN and Tamoxifen injections. Paraffin embedding and H&E staining were performed by the Liver Histology Core of USC Liver Center (P30DK048522). Flow cytometry was performed in the USC Flow Cytometry Facility that is supported in part by the National Cancer Institute Cancer Center Shared Grant Award (P30CA014089) and the USC Office of the Provost, Dean's Development Funds, Keck School of Medicine of USC. PROM1+ AH mRNA-seq was performed by USC Norris Molecular Genomics Core. The bioinformatics software and computing resources used in the analysis of PROM1+ AH mRNA-seq are funded by the USC Office of Research and the Norris Medical Library. 10× Single cell RNA-seq libraries were also prepared by USC Norris Molecular Genomics Core. Sequencing of single cell RNA-seq libraries and cell ranger analysis were performed by Genewiz. The study was supported by Grants: U01AA027681, U01AA018663, P50AA011999 (pilot project award to R.W.), and R24 AA012885 (Integrative Liver Cell Core) from NIAAA/NIH, and IK6BX004205 (BLR&D Research Career Scientist Award) and I01 BX001991 (VA Merit Review) from Department of Veterans Affairs.

Author contributions
The project was originally and conceptually developed and supervised until its completion by H.T. R.W. performed the experiments and analyses except those described below. Y.C. performed Partek flow analysis of PROM1+ AH RNA-seq data in Fig. 1. Y.N. performed IF shown in Fig. 2b. S.P. isolated liver cells used in Figs. 1, 3, 4 and 5. H.T. and S.P. generated the TIC transplantation data in Fig. 7e. M.L. provided bioinformatic consultation for Fig. 6b-g. S.B. generated Suppl. Figure 3a-b and Col1a1Cre;Rosa-26mTmG mice. R.W. wrote the manuscript draft which was edited, revised and submitted by H.T.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-76458-8.

Correspondence and requests for materials should be addressed to H.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2020