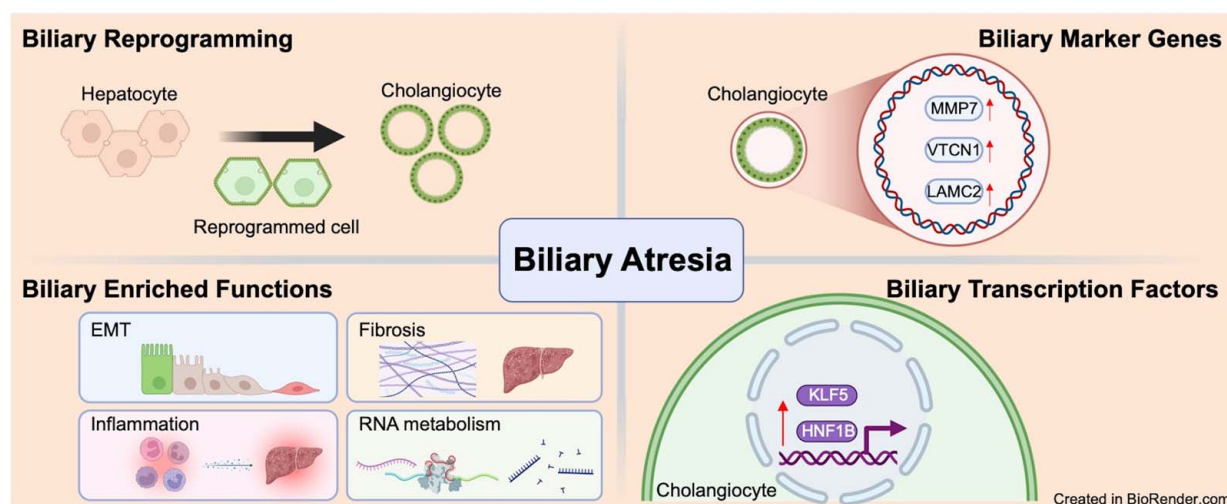


Single-cell transcription reveals hepatocyte-to-cholangiocyte reprogramming and biliary gene profile in biliary atresia

VISUAL ABSTRACT











Single-cell transcription reveals hepatocyte-to-cholangiocyte reprogramming and biliary gene profile in biliary atresia



ORIGINAL ARTICLE

OPEN

Single-cell transcription reveals hepatocyte-to-cholangiocyte reprogramming and biliary gene profile in biliary atresia

Lingdu Meng¹  | Min Du²  | Haodong Li³  | Fanyang Kong¹  |
 Jiajian Yang⁴  | Rui Dong¹  | Shan Zheng¹  | Gong Chen¹  |
 Zhen Shen¹  | Junfeng Wang¹ 

¹Department of Pediatric Surgery, Children's Hospital of Fudan University, Shanghai Key Laboratory of Birth Defect, and Key Laboratory of Neonatal Disease, Ministry of Health, Shanghai, P.R. China

²Department of Pediatric Gastroenterology, Chengdu Women's and Children's Central Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, P.R. China

³Department of Pediatric Orthopedics, Children's Hospital of Fudan University, Shanghai, P.R. China

⁴Department of Pediatric Surgery, Children's Hospital of Fudan University (Xiamen Branch), Xiamen Children's Hospital, Xiamen, P.R. China

Correspondence

Junfeng Wang, Department of Pediatric Surgery, Children's Hospital of Fudan University, Shanghai Key Laboratory of Birth Defect, and Key Laboratory of Neonatal Disease, Ministry of Health, 399 Wan Yuan Road, Shanghai 201102, P.R. China.
 Email: wangjunfeng@fudan.edu.cn

Zhen Shen, Department of Pediatric Surgery, Children's Hospital of Fudan University, Shanghai Key Laboratory of Birth Defect, and Key Laboratory of Neonatal Disease, Ministry of Health, 399 Wan Yuan Road, Shanghai 201102, China.
 Email: szhen0157079@hotmail.com

Gong Chen, Department of Pediatric Surgery, Children's Hospital of Fudan University, Shanghai Key Laboratory of Birth Defect, and Key Laboratory of Neonatal Disease, Ministry of Health, 399 Wan Yuan Road, Shanghai 201102, China.
 Email: chengongzlp@fudan.edu.cn

Abstract

Background: Ductular reaction (DR), characterized by the expansion of biliary epithelial cells in the portal area, is a typical hepatic pathology for biliary atresia (BA). The cellular source and function of DR remain poorly understood. Herein, we performed single-cell RNA sequencing (scRNA-seq) in BA to resolve the complexity of DR in BA.

Methods: A total of 4 BA and 3 normal control livers underwent scRNA-seq. The epithelial cells were extracted from all cells for further analysis. The cell types, functions, and differentiation trajectory of epithelial cells were determined. The biliary markers and transcription factors (TFs) were identified by combining public bulk and scRNA-seq data and validated by immunohistochemistry.

Results: ScRNA-seq identified the existence of biliary reprogramming in BA, and the reprogrammed cells expressed both hepatocyte and cholangiocyte markers. When compared with hepatocytes, genes of epithelial-mesenchymal transition, fibrosis, inflammation, and RNA metabolism were enriched in cholangiocytes and upregulated in BA. Pseudotime analysis depicted a differentiation trajectory from hepatocytes across reprogrammed cells to cholangiocytes in BA. Matrix metalloproteinase 7 (*MMP7*), *VTCN1*, and

Abbreviations: BA, biliary atresia; BEC, biliary epithelial cell; CS, cholestasis; DDC, 3,5-diethoxycarbonyl-1,4-dihydrocollidine; DEG, differentially expressed gene; DR, ductular reaction; EMT, epithelial-mesenchymal transition; IHC, immunohistochemistry; *MMP7*, matrix metalloproteinase 7; NC, normal control; scRNA-seq, single-cell RNA sequencing; TF, transcription factor; TMA, tissue microarray.

Lingdu Meng and Min Du contributed equally to the work.

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.hepcommjournal.com.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Copyright © 2025 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Association for the Study of Liver Diseases.

LAMC2 were identified as the biliary markers, and *KLF5* and *HNF1B* were determined as the biliary TFs in BA. All the biliary markers and TFs were upregulated in BA when compared with controls.

Conclusions: Dissecting the cellular source and function of cholangiocytes is essential to understand the pathological role of DR in BA. The identified specific biliary markers and TFs provide important insights into its potential diagnosis and mechanism exploration for BA in the future.

Keywords: biliary atresia, biliary reprogramming, ductular reaction, single-cell RNA sequencing

INTRODUCTION

Biliary atresia (BA), a dominant etiology of obstructive jaundice in infants, is characterized by fibroinflammatory cholangiopathy, which results in progressive liver fibrosis and even cirrhosis.^[1] However, the pathogenesis of BA is largely undefined. In BA, ductular reaction (DR) is characterized by proliferative bile ductules in the portal area. Experimental studies have demonstrated that hepatic progenitor cells, hepatocytes, and cholangiocytes contribute to DR.^[2] The process whereby hepatocytes transdifferentiate into cholangiocytes is known as “biliary reprogramming”.^[3] What is the resource of DR and whether biliary reprogramming is involved in DR remain largely undefined in BA.

The role of DR in BA is poorly understood. Given that DR is an adaptive response to cholestasis (CS), it may support the drainage of bile.^[4] However, DR is located at the portal area, where substantial collagen is deposited, followed by the formation of hepatic scar. In this way, DR may be linked to liver fibrosis.^[5] Previous studies have demonstrated that the nascent biliary epithelial cells (BECs) in DR secrete inflammatory cytokines (IL-8, osteopontin) and platelet-derived growth factor-BB to create an inflammatory and profibrogenic microenvironment, thereby inducing HSC activation and liver fibrosis.^[6–8] Additionally, these aberrant BECs in DR manifested epithelial-mesenchymal transition (EMT) pathology,^[9,10] which may also contribute to biliary fibrosis. It is essential to understand the biliary gene profile to figure out the role of DR in BA. Up to now, matrix metalloproteinase 7 (MMP7), a unique product of BECs, has been identified as a novel biomarker to diagnose BA with high sensitivity and specificity.^[11–15] However, apart from MMP7, few biliary markers were reported in BA. Therefore, it is urgent to find more biliary biomarkers to help us understand the pathogenesis of BA.

Single-cell RNA sequencing (scRNA-seq) is a technology to dissect cell features at single-cell resolution. In this study, single-cell analysis demonstrated a

differentiation trajectory from hepatocytes, across reprogrammed cells, to cholangiocytes in BA. In this trajectory, the genes in the pathways, including EMT, fibrosis, inflammation, and mRNA metabolism, were upregulated. Bulk and single-cell analyses identified *MMP7*, *VTCN1*, and *LAMC2* as biliary markers, and determined *KLF5* and *HNF1B* as biliary transcription factors (TFs) in BA. This study elucidates the cell source and transcriptional profile of DR, which may provide important insights into the pathological role of DR in BA.

METHODS

Patient samples

A total of 4 BA and 3 normal control livers (NC, normal liver tissue adjacent to the excised hepatoblastoma) were collected for scRNA-seq; the patients' clinical information is shown in Supplemental Table S1, <http://links.lww.com/HC9/B972>. Hepatoblastoma was diagnosed by hepatic histopathology, and intraoperative cholangiography and subsequent liver histological examination were used for BA diagnosis.

The samples of liver tissue microarray (TMA) harvested here were the same as those in our previous study.^[16] In this TMA, 30 NC (normal liver tissue adjacent to the excised hepatoblastoma), 34 CS with non-BA causes (disease control), and 93 patients with BA were enrolled. Their clinical information is summarized in Supplemental Table S2, <http://links.lww.com/HC9/B972>. All the samples were subjected to immunohistochemistry (IHC) analysis.

Preparation of single-cell solution

The single-cell solution was prepared as described.^[16] Briefly, the dissected liver tissues were digested with the solution mixed with collagenase IV (Gibco, 17104-19, 2 mg/mL) and DNase I (Sigma, DN25-100MG,

0.1 mg/mL) for 30 minutes in the 37 °C shaker. The suspension was filtered through the 70-µm Nylon cell strainer to remove the undigested tissue residue. Subsequently, red blood cells and dead cells were removed using the red blood cell lysis solution (Miltenyi Biotec, 130-094-183) and a magnetically activated cell sorting kit (Miltenyi Biotec, 130-090-101) according to the manufacturer's instructions, respectively. Finally, the single-cell solution was acquired, and its cell number and viability were determined by the Countess II FL (Invitrogen, AMQAF1000).

Single-cell library construction and sequencing

Shanghai Genengy Inc. performed the single-cell library construction and sequencing. Qualified single-cell suspensions, gel beads with 10× barcodes, and oil were loaded on a Chromium Single Cell Instrument (10× Genomics) to generate single-cell gel bead-in-emulsions for sample and cell barcoding. Single-cell 3' RNA-seq libraries were generated following the manufacturer's instructions (10× Genomics Chromium Single Cell 3' Reagent Kit User Guide v2 Chemistry).

The libraries were pooled and sequenced on Illumina NovaSeq 6000 with a sequencing depth of ~400 M reads per sample. After the sequencing, Illumina bcl2fastq (2.19.1) was used to convert the raw data into a FASTQ file and align the data to the human genome reference sequence (GRCH38). The Cell Ranger (10× Genomics, 2.1.1) analysis pipeline was applied to perform sample demultiplexing and single-cell 3' gene counting to generate a digital gene-cell matrix by using the default parameters. The sequencing information of each sample is shown in Supplemental Table S3, <http://links.lww.com/HC9/B972>.

Seurat analysis

Seurat package (4.0.2) (https://satijalab.org/seurat/articles/pbm3k_tutorial.html) was used for scRNA-seq data processing.^[17] Cells with the number of detected genes per cell in the range of 500–6000, and <25% of mitochondrial genes detected were retained for subsequent analyses, while genes detected (UMI count > 0) in <3 cells were removed. After quality control, a total of 70,538 cells were obtained. Then, the gene expressions of individual cells underwent normalization, identification of highly variable 2000 genes, linear dimensional reduction, PCA reduction, determination of dimensionality of the data set, and cell clustering with resolution = 0.6. UMAP was employed for the nonlinear dimensional reduction of data set visualization. The *FindAllMarkers* function (Wilcoxon statistical test, min.pct = 0.25, logfc.threshold = 0.25) was applied to identify

the differentially expressed genes (DEGs) in a given cluster when compared with all the other clusters, and the DEGs in each cluster were visualized using a heatmap. According to the DEGs of each cluster, the cell type was defined using the *RenamelDents* function. The number of each cell type in every sample is shown in Supplemental Table S3, <http://links.lww.com/HC9/B972>. *VlnPlot* and *Dotplot* were used to visualize one specific gene expression. The epithelial cells were extracted using the *subset* function and reanalyzed as described above. The doublets were removed using the R package DoubletFinder according to the multiplet rate provided by the 10× Genomics. The R package Harmony (<https://github.com/immunogenomics/harmony>) was adopted to merge multiple data sets.^[18]

Single-cell gene function enrichment analysis

Single-cell gene function enrichment analysis was performed using the R package clusterProfiler.^[19] Gene symbols in each cluster were converted into the ENTREZID format in accordance with the human gene annotation package (org.Hs.eg.db). Then, the *enrichGO* function was used to perform GO enrichment analysis, and the *enrichKEGG* function was used for KEGG enrichment analysis. Finally, the results were visualized using the *dotplot* function. Furthermore, gene set variation analysis (GSVA)^[20] was employed to identify the enriched pathways at the single-cell level; the gene sets in GO biological processes (c5.go.bp.v7.5.1.symbols.gmt) were set as the reference, and the results were visualized by heatmaps. Based on the 1184 EMT genes in dbEMT 2.0 (<http://www.dbemt.bioinfo-min-zhao.org/index.html>), the EMT score of each cell type was calculated by the *AddModuleScore* function in the Seurat package.

Pseudotime trajectory analysis

To build single-cell trajectories in epithelial cells, the R package Monocle 2 was used to perform pseudotime analysis.^[21] The highly variable features identified by Seurat analysis were set as the reference to construct single-cell trajectories. Dynamic gene expression patterns in each state were determined in pseudotime.

RNA velocity analysis

The loom file of each sample that contained spliced and unspliced reads was generated with Cell Ranger software. The preprocessed data of each sample that contained cell annotations were extracted from the Seurat analysis. The above 2 data sets were merged to

store the cell information of each sample, and the data of all samples were further merged for the following RNA velocity analysis under the pipeline of the Python scVelo package (version 0.2.4, <https://scvelo.readthedocs.io>).^[22] The RNA velocity values for each gene in each cell were calculated and projected to the low-dimensional space in the UMAP plot. On the basis of the velocity graph, a velocity pseudotime was computed.

Transcription factor analysis

The python package for Single-Cell Regulatory Network Inference and Clustering (SCENIC) pySCENIC (version 0.11.2)^[23] was adopted to analyze the differential TFs among all the cell types with default parameters, and the count matrix of cells from Seurat was set as input. The regulon activity (AUC) scores of TFs were visualized using *Dotplot*, *Vlnplot*, or *heatmap*. To screen the transcription activity of EMT-related TFs in BA epithelial cells, the EMT TF database^[24] (<http://www.medsysbio.org/EMTRegulome>, type 6) was applied as the reference.

Bioinformatic analysis of the public data

A total of 8 public data sets were included for analysis, including 7 bulk RNA data sets (GSE46960, GSE157661, GSE55552, GSE108315, GSE156894, GSE218945, and GSE90047)^[3,25–29] and 1 scRNA-seq data set (GSE157698).^[3] In the bulk RNA data sets, the genes encoding m⁶A readers, EMT, and fibrotic markers, as well as the inflammatory genes, were visualized via heatmaps. To calculate the cell correlation between our scRNA-seq data and the public data sets, the overlapping genes between the variable features of our scRNA-seq data (2000 genes) and the row names of the bulk gene matrix data were identified. If the bulk data were from the mouse samples, the mouse gene symbols were transformed into human gene symbols by biomaRt package. Then, the single-cell matrix was merged with the bulk matrix. Subsequently, the correlations of each cell in our scRNA-seq and the cell types in bulk data were determined by the Spearman analysis. Finally, the correlation was visualized by a density plot or violin plot. All the scRNA-seq data were processed under the analytical workflow of the R package Seurat. Cells that expressed fewer than 200 genes, and genes expressed in <3 cells were removed. In the data set of GSE157698, since sample YFP_w3 had more than 70,000 cells, we randomly selected 5000 cells to match the number of cells in the other samples. After quality control, the number of cells in each data set is shown in Supplemental Table S4, <http://links.lww.com/HC9/B972>. UMAP was adopted to visualize the nonlinear dimensional reduction, and each cell type was identified by its differentially expressed

features (min.pct = 0.25, logfc.threshold = 0.25). To observe the large cells' gene dynamic expressions from one state to another in GSE157698, Monocle 3 was employed to demonstrate the pseudotime trajectory of the cells and plot the genes in pseudotime by using the default parameter settings (<https://cole-trapnell-lab.github.io/monocle3/>).^[30]

IHC and immunofluorescence staining

Paraffin-embedded liver sections in the microarray were first dewaxed and hydrated, followed by antigen retrieval and inactivation of endogenous peroxidases. After antigen blocking with 5% BSA, the sections were incubated overnight at 4 °C with the following primary antibodies: CK19 (Abcam, ab52625, 1:600), HNF4A (Abcam, ab201460, 1:2000), SOX9 (CST, 82630, 1:600), CDH2 (CST, 13116, 1:200), Smad2/3 (CST, 8685, 1:400), FGFR2 (CST, 23328, 1:200), PDGFA (santa cruz, sc-9974, 1:200), CXCL6 (CST, 73201, 1:100), IGF2BP2 (Abcam, ab124930, 1:200), CFTR (Abcam, ab270238, 1:500), MMP7 (Abcam, ab207299, 1:4000), VTCN1 (Abcam, ab209242, 1:1000), LAMC2 (Abcam, ab210959, 1:500), ANKRD1 (Proteintech, 11427-1-AP, 1:200), ELF3 (Sigma, HPA003479, 1:200), KLF5 (Abcam, ab273672, 1:500), and HNF1B (Abcam, ab128912, 1:100). Subsequently, the sections were incubated with secondary antibodies conjugated to horseradish peroxidase (Gene tech, GK500710) for 30 minutes at room temperature. The signals were detected using the DAB substrate and hematoxylin. The detailed procedures for image acquisition and expression calculations were described in our previous study.^[16] Due to tissue detachment in the TMA for FGFR2 IHC staining, a total of 28 BA, 21 CS, and 25 NC samples were included for its analysis. Multiple immunofluorescence staining was performed using a tyramide signal amplification kit (Panovue, 10079100020); the detailed methods can also be found in our previous study.^[16]

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the Children's Hospital of Fudan University. The implementations were conducted in accordance with the Declaration of Helsinki, and written informed consent was obtained from the legal guardians of all subjects.

Availability of data and materials

The raw sequencing data of this study were deposited to the Genome Sequence Archive for Human of National Genomics Data Center under accession number HRA000019, and the processed count matrix

of scRNA-seq was deposited to OMIX of National Genomics Data Center under accession number OMIX008679.

Statistical analysis

Statistical analysis was performed using R (version 4.0.3). The results are presented as median and quartiles. For the data that does not conform to normality, the Kruskal-Wallis test was used to calculate the differences in gene expression in the immunostained liver TMA and the clinical indicators among 3 groups, where the Dunn multiple comparisons test was employed to determine the difference between every 2 groups for post hoc analysis. For the normalized data in GSE46960, one-way ANOVA was adopted to analyze the differences among 3 groups, where Tukey multiple comparison test was applied to compare every 2 groups for post hoc analysis. The correlation between the cell types in our data and those in the public data was analyzed using the Spearman test. The Wilcoxon signed-rank test was employed to compare the difference between the 2 groups in the scRNA-seq data. Differences were considered significant when $p < 0.05$.

RESULTS

Single-cell RNA sequencing identified the existence of biliary reprogramming in BA

DR is a response to cholestatic liver injury, and the number of CK19-positive bile ductules in BA and CS was significantly higher than that in NC (Figures 1A, B, BA and CS vs. NC: $p < 0.0001$). To dissect the gene profile of DR in BA, scRNA-seq was performed. A total of 70,538 cells were acquired after quality control; their UMAP plot, cell composition in each sample and group, and top 3 marker genes are presented in Supplemental Figures S1A–D, <http://links.lww.com/HC9/B972>. Among all cells, 1282 epithelial cells were isolated for further analysis after removing the contaminating *PTPRC*⁺/*CD3D*⁺/*MS4A1*⁺/*CD68*⁺ immune cells and *PECAM1*⁺ endothelial cells (Figure 1C, Supplemental Figures S1E, F, <http://links.lww.com/HC9/B972>). Based on the expressions of epithelial markers, these epithelial cells were clustered into hepatocytes, cholangiocytes, and reprogrammed cells (Figure 1D, Supplemental Table S5, <http://links.lww.com/HC9/B972>). These reprogrammed cells expressed both hepatocyte and cholangiocyte markers (Figure 1E) and were only found in patients with BA (Figure 1F). The heatmap of the top 10 genes also demonstrated that the programmed cells

shared common gene features of hepatocytes and cholangiocytes (Figure 1G).

To validate the reprogrammed cell existence, we compared the cell types in our data and those in GSE157661,^[3] GSE157698,^[3] GSE156894,^[3] GSE55552,^[26] and GSE218945.^[28] In GSE157661, human hepatocytes were transplanted into mouse livers and then biliary reprogramming was induced by 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC) treatment. The hepatocytes, cholangiocytes, and reprogrammed cells in our data were consistent with those in GSE157661 (Supplemental Figure S1G, <http://links.lww.com/HC9/B972>). In GSE157698, RosaYFP mice were treated with DDC to induce hepatocyte-to-biliary reprogramming, and the hepatocytes, cholangiocytes, and reprogrammed cells were collected for scRNA-seq. The reprogrammed cells in our data resembled the intermediate state between hepatocytes and cholangiocytes in GSE157698 (Supplemental Figure S1H, <http://links.lww.com/HC9/B972>). In GSE156894, wild-type hepatocytes, wild-type biliary cells, DDC-treated hepatocytes, DDC-treated biliary cells, and DDC-treated reprogrammed cells were isolated by fluorescence-activated cell sorting and sequenced in DDC-treated mice. The reprogrammed cells in our data were kept consistent with DDC-treated reprogrammed cells and wild-type or DDC-treated biliary cells in GSE156894 (Supplemental Figure S1I, <http://links.lww.com/HC9/B972>). In GSE55552, hepatocytes, hepatocyte-derived oval cells (reprogrammed cells), and cholangiocyte-derived oval cells (cholangiocytes) underwent RNA-seq in DDC-treated mice. Reprogrammed cells in our data were also correlated with reprogrammed cells and hepatocytes in GSE55552 (Supplemental Figure S1J, <http://links.lww.com/HC9/B972>). In GSE218945, varied stages of reprogrammed cells and cholangiocytes were collected for sequencing in DDC-treated mice. The reprogrammed cells in our data also manifested an intermediate state between early reprogrammed cells and cholangiocytes (Supplemental Figure S1K, <http://links.lww.com/HC9/B972>). To sum up, the reprogrammed cells in our data demonstrated a similar transcriptome profile with those in the DDC-induced CS, and all these results reinforced the accuracy of our cell definition.

To further validate the findings in scRNA-seq, liver TMA's immunofluorescence staining was performed. The percentage of cholangiocytes (SOX9⁺) and reprogrammed (HNF4A⁺ SOX9⁺) cells had a remarkable elevation in BA and CS when compared with NC (Figures 1H–K, BA vs. NC: $p < 0.0001$, CS vs. NC: $p < 0.001$). More specifically, the reprogrammed cells were located at the intermediate zone between the parenchyma and portal area (Figure 1H).

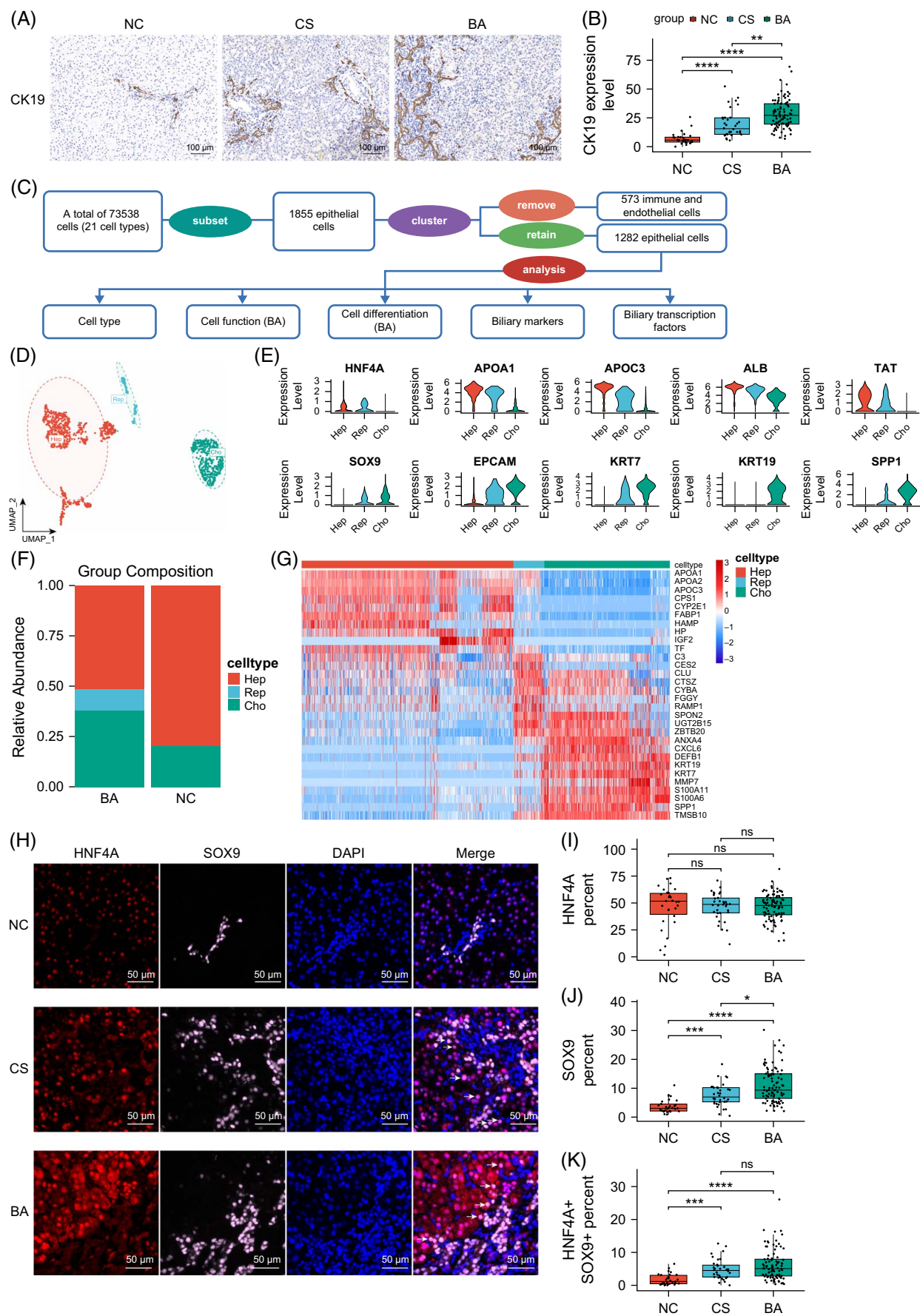


FIGURE 1 Single-cell RNA sequencing identified the existence of biliary reprogramming in BA. (A) Representative immunohistochemistry of CK19 in a tissue microarray (TMA). (B) Box plots showing the expression levels of CK19 in TMA, colored by different groups. ns, not significant; $^{**}p < 0.01$, $^{****}p < 0.0001$. (C) Flowchart of obtaining the final 1282 epithelial cells from the whole-cell population and the whole study design. (D) UMAP visualization of hep, rep, and cho. (E) Violin plots showing the expression levels of the marker genes of hep, rep cells, and chol, colored by cell types. (F) Bar plots displaying the composition of each epithelial cell type in different groups. (G) Heatmap demonstrating the top 10 genes highly expressed in hep, rep cells, and chol. (H) Immunofluorescence staining showing the expression of HNF4A and SOX9 in different groups. (I–K). Box plots showing the percentage of HNF4A⁺, SOX9⁺, and HNF4A⁺ SOX9⁺ cells in different groups, respectively. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$. Abbreviations: BA, biliary atresia; CHO, cholangiocytes; CS, cholestasis; HEP, hepatocytes; NC, normal control; REP, reprogrammed cells.

Genes in the pathways of EMT, fibrosis, inflammation, and RNA metabolism were enriched in cholangiocytes and upregulated in BA

To understand the function of epithelial cells in the BA group, we extracted the BA epithelial cells from the merged groups and defined the cell types based on their classical markers (Figures 2A–C, Supplemental Table S5, <http://links.lww.com/HC9/B972>). As expected, the reprogrammed cells in BA also demonstrated a similar transcriptomic profile to those in the public data sets (GSE157661, GSE157698, GSE156894, GSE55552, and GSE218945) (Supplemental Figure S2A, <http://links.lww.com/HC9/B972>). GO analysis demonstrated that metabolic processes were enriched in hepatocytes, whereas reprogrammed cells were linked to immune response, and RNA splicing and focal adhesion were enriched in cholangiocytes (Supplemental Figures S2B, <http://links.lww.com/HC9/B972>). KEGG supported the results of GO analysis (Supplemental Figure S2C, <http://links.lww.com/HC9/B972>). GSVA analysis indicated the enrichment of EMT, fibrosis, inflammation, and mRNA metabolism in cholangiocytes (Figure 2D). To better understand the EMT states of these 3 cell types, the EMT gene database was set as the reference gene set to evaluate the EMT score of each cell type. The results indicated that epithelial cells exhibited a graded increase in EMT score across reprogrammed cells, with low-scoring cells in hepatocytes and high-scoring cells in cholangiocytes (Figure 2E). The SCENIC analysis further revealed that 80 EMT-related TFs had elevated transcription activity in cholangiocytes (Supplemental Figure S2D, <http://links.lww.com/HC9/B972>). Average gene expression level in each cell type demonstrated that EMT and fibrotic markers, inflammatory genes, and m⁶A readers were upregulated in cholangiocytes and reprogrammed cells when compared with the levels in hepatocytes (Figure 2F). Given the expanding number of BECs in cholestatic diseases, the expression levels of CDH2, Smad2/3, FGFR2, CXCL6, and IGF2BP2 were elevated in BA and CS when compared with the control levels (Figures 2G–I, S2E, <http://links.lww.com/HC9/B972>). Whereas PDGFA had a higher expression in endothelial cells than BECs and was excluded from the calculation.

To validate the alternation of selected gene expressions in BA, we reanalyzed the data set of GSE157661.^[3] The expression levels of EMT and fibrotic markers, inflammatory genes, and m⁶A readers were also increased in the order of hepatocytes, reprogrammed cells, and cholangiocytes (Supplemental Figure S2F, <http://links.lww.com/HC9/B972>). Similarly, the expression levels of these genes were validated in another data set of DDC-induced CS (GSE55552)^[26] (Supplemental Figures S2G, <http://links.lww.com/HC9/B972>) and in a data set derived from a mouse model of Alagille syndrome (GSE108315)^[27] (Supplemental Figure S2H, <http://links.lww.com/HC9/B972>).

Pseudotime analysis depicted a differentiation trajectory from hepatocytes across reprogrammed cells to cholangiocytes in BA

To predict the underlying development trajectory in BA epithelial cells, RNA velocity and Monocle 2 analyses were performed. The results demonstrated a differentiation trajectory in which hepatocytes can transform into cholangiocytes through reprogrammed cells (Figures 3A–D). In this trajectory, the hepatocyte signals were gradually replaced by the biliary signals (Figure 3E), and EMT and fibrotic markers, inflammatory genes, and m⁶A readers were upregulated (Figure 3F).

To better figure out the dynamic expression of the above genes in biliary reprogramming, we used the scRNA-seq data set in GSE157698.^[3] Their cell composition and DEGs in each cell type are presented in Figure 3G and Supplemental Figures S3A–D, <http://links.lww.com/HC9/B972>. Monocle 3 demonstrated a pseudotime manner from hepatocytes, across reprogrammed cells to cholangiocytes (Figure 3H). Co-expressed gene module analysis indicated that the reprogrammed cells were different from hepatocytes, but shared more common features with the cholangiocytes (Figure 3I). These results proved the existence of biliary reprogramming in DDC-induced CS. The expression levels of EMT and fibrotic markers, inflammatory genes, and m⁶A readers were also increased in the order of hepatocytes, reprogrammed cells, and

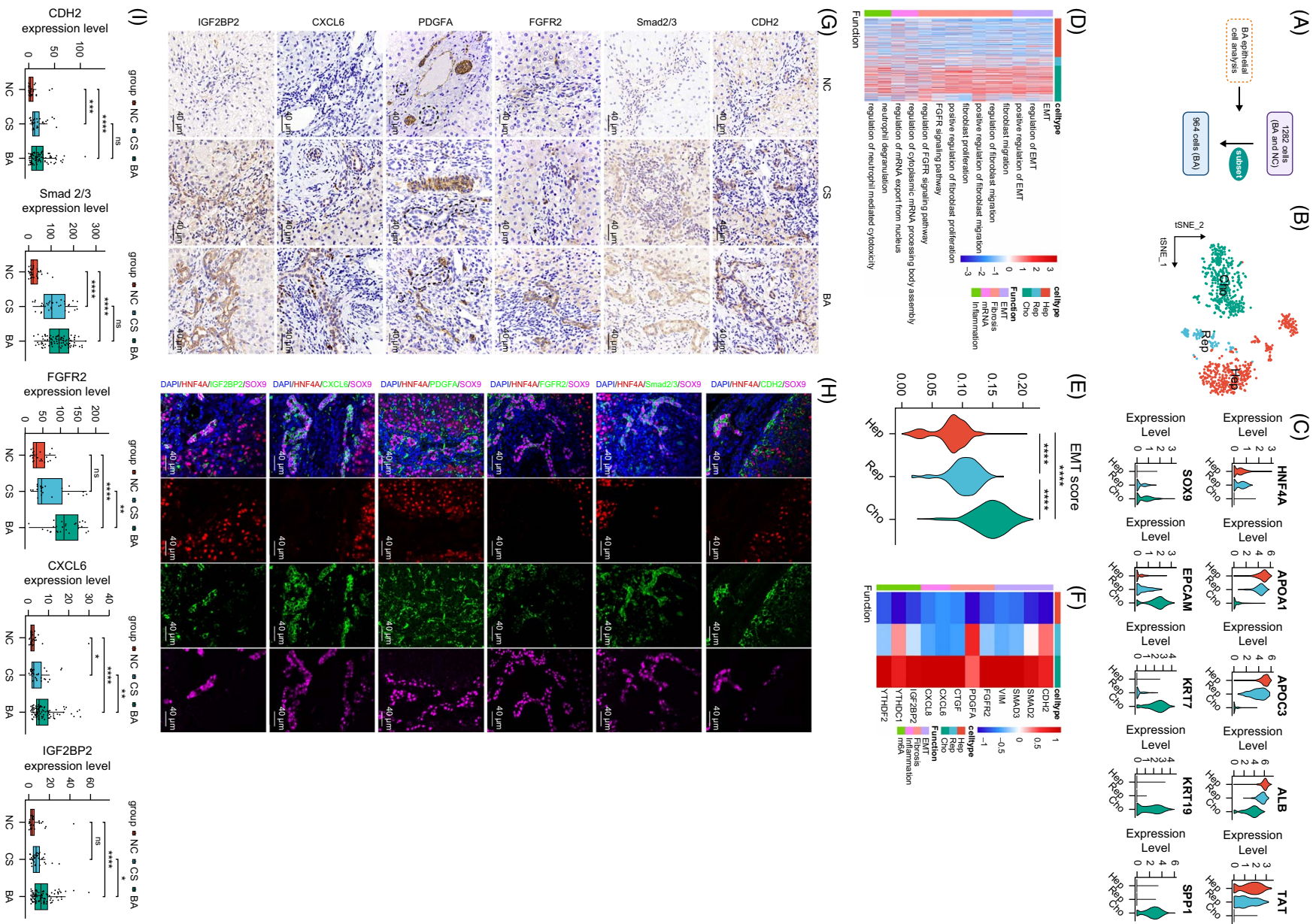


FIGURE 2 Enriched pathways and genes in the cholangiocytes of BA. (A) Flowchart of isolating BA epithelial cells from the merged groups. (B) t-SNE visualization of hep, rep, and cho in BA. (C) Violin plots showing the expression levels of the marker genes of hep, rep, and cho in BA. (D) Heatmap displaying enriched pathways of cholangiocytes in BA by GSVA analysis. (E) Violin plots displayed the EMT score among the 3 types of epithelial cells in BA. (F) Heatmap displaying the average expression level of EMT markers, fibrotic markers, inflammatory genes, and m⁶A readers in BA hep, rep, and cho. (G) Representative immunohistochemistry sections of CDH2, smad2/3, FGFR2, PDGFA, CXCL6, and IGF2BP2 in a TMA. Dashed circles in PDGFA were recognized as cholangiocytes. H. Representative immunofluorescence staining of CDH2, Smad2/3, FGFR2, PDGFA, CXCL6, and IGF2BP2 in TMA of BA group. I. Box plots showing the expression levels of the indicated genes in TMA, colored by different groups. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$. Abbreviations: BA, biliary atresia; CHO, cholangiocytes; CS, cholestasis; EMT, epithelial-mesenchymal transition; HEP, hepatocytes; NC, normal control; REP, reprogrammed cells; TMA, tissue microarray.

cholangiocytes (Figure 3J). Taken together, the gene profile demonstrated a similar alternation in BA and CS-induced biliary reprogramming.

Bulk and single-cell analyses identified upregulated biliary genes in BA

Given the significance of DR in hepatic pathology, we assumed that the genes that were uniquely expressed in cholangiocytes may be the potential targets for the diagnosis and therapy in BA. An RNA microarray data (GSE46960) that consists of 64 BA, 14 CS, and 7 NC samples was adopted for analysis. A total of 485 DEGs were screened out between BA and NC, 30 DEGs were filtered out between BA and CS, and 440 DEGs were determined between CS and NC. Venn diagram identified 16 overlapping genes between the DEGs of BA versus NC and BA versus CS, but not the DEGs between CS and NC (Figure 4A). Single-cell analysis further demonstrated that *MMP7*, *ANKRD1*, *VTCN1*, and *LAMC2* were highly expressed in cholangiocytes (Figures 4B, C). At the single-cell level, all the biliary markers were upregulated in BA when compared with NC (Supplemental Figure S4A, <http://links.lww.com/HC9/B972>), which was supported by the bulk expressions (Supplemental Figure S4B, <http://links.lww.com/HC9/B972>). Furthermore, the IHC and immunofluorescence analysis of liver TMA further validated the higher expressions of the above biliary markers except *ANKRD1* in BA (Figures 4D–F, Supplemental Figure S4C, <http://links.lww.com/HC9/B972>). *ANKRD1* was mainly expressed in hepatocytes and was excluded from the biliary markers, which may be attributed to the nonspecificity of the present antibody.

Single-cell analysis determined the upregulated biliary transcriptional factors in BA

The SCENIC analysis identified 365 differential TFs in epithelial cells; their relative transcriptional activity was presented in Supplemental Figure S5A, <http://links.lww.com/HC9/B972>. To screen out the biliary unique TFs in the liver, 21 TFs with high activity in cholangiocytes

were primarily determined (Figure 5A), and then their cellular expressions in the whole liver were shown in Figure 5B. Among 21 TFs, *ELF3* was expressed in cholangiocytes and reprogrammed cells, and *KLF5* and *HNF1B* were expressed in cholangiocytes (Figure 5C). The mRNA and protein levels of *KLF5* and *HNF1B* were significantly elevated in BA when compared with those in NC (Figures 5D–F, Supplemental Figures S5B, C, <http://links.lww.com/HC9/B972>; IHC expressions: *KLF5*, BA versus NC: $p < 0.01$; *HNF1B*, BA versus NC: $p < 0.0001$), whereas *ELF3* expression level was comparable in BA and NC (Figure 5F, Supplemental Figure S5C, <http://links.lww.com/HC9/B972>). Taken together, *KLF5* and *HNF1B* were determined as the biliary TFs in BA.

DISCUSSION

DR, characterized by the expansion of BECs in the portal area, is a typical hepatic pathology for BA. However, the source and function of these proliferative BECs remain elusive in BA. DR can be identified in various biliary injuries and is driven not only by cholangiocyte self-proliferation but also by hepatocyte transdifferentiation.^[26,31] Inversely, hepatocyte regeneration can also be driven by cholangiocyte transdifferentiation in severe hepatocyte loss.^[32] Therefore, the direction between hepatocytes and cholangiocytes transdifferentiation may be based on the epithelial damage types. Accumulating evidence suggests that hepatocyte-derived biliary cells were the source of DR in cholestatic diseases.^[3,26,27,33] It is worth noting that most of the studies were carried out in animal models; whether this transdifferentiation occurs in human diseases remains to be undefined. In this study, our data indicated that hepatocytes can transdifferentiate into cholangiocytes through biliary reprogramming in BA. The unchangeable number of hepatocytes and elevated number of reprogrammed cells and cholangiocytes also implied the hepatocyte-to-cholangiocyte reprogramming in BA, which can be supported by our recent finding that the number of SOX9-positive progenitor-like cells was increased from periportal parenchyma to the portal region with the progression of BA.^[34]

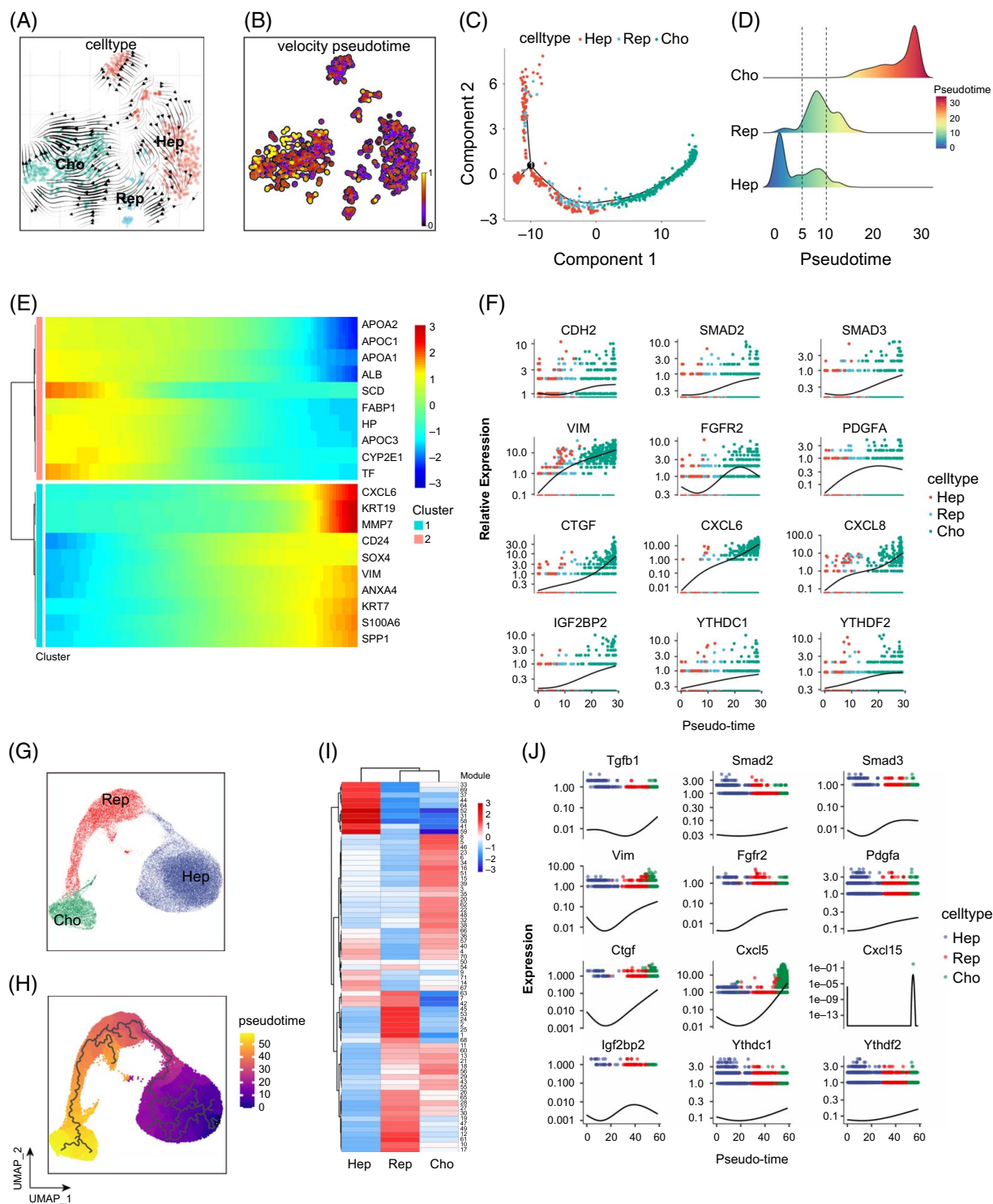


FIGURE 3 Pseudotime analysis depicted a differentiation trajectory from hepatocytes across reprogrammed cells to cholangiocytes in BA. (A) Inferred developmental trajectory of epithelial cells in BA by RNA velocity. (B) t-SNE visualization of BA epithelial cells colored by velocity pseudotime. (C) The distribution of BA epithelial cells constructed by Monocle 2 and colored by cell types. (D) Differentiation trajectory of BA epithelial cells in pseudotime order constructed by Monocle 2. (E) Heatmap showing the dynamic changes of gene expressions along the pseudotime in BA epithelial cells. (F) Plots showing the relative expression of indicated genes in pseudotime order in BA epithelial cells. (G, H) UMAP visualization of epithelial cells colored by cell types (G) and pseudotime trajectory constructed by Monocle 3 (H) in GSE157698. (I) Heatmap showing the expression of gene modules in different cell types in GSE157698. (J) Plots showing the relative expression of indicated genes in pseudotime order in GSE157698. Abbreviations: BA, biliary atresia; CHO, cholangiocytes; HEP, hepatocytes; REP, reprogrammed cells; TMA, tissue microarray.

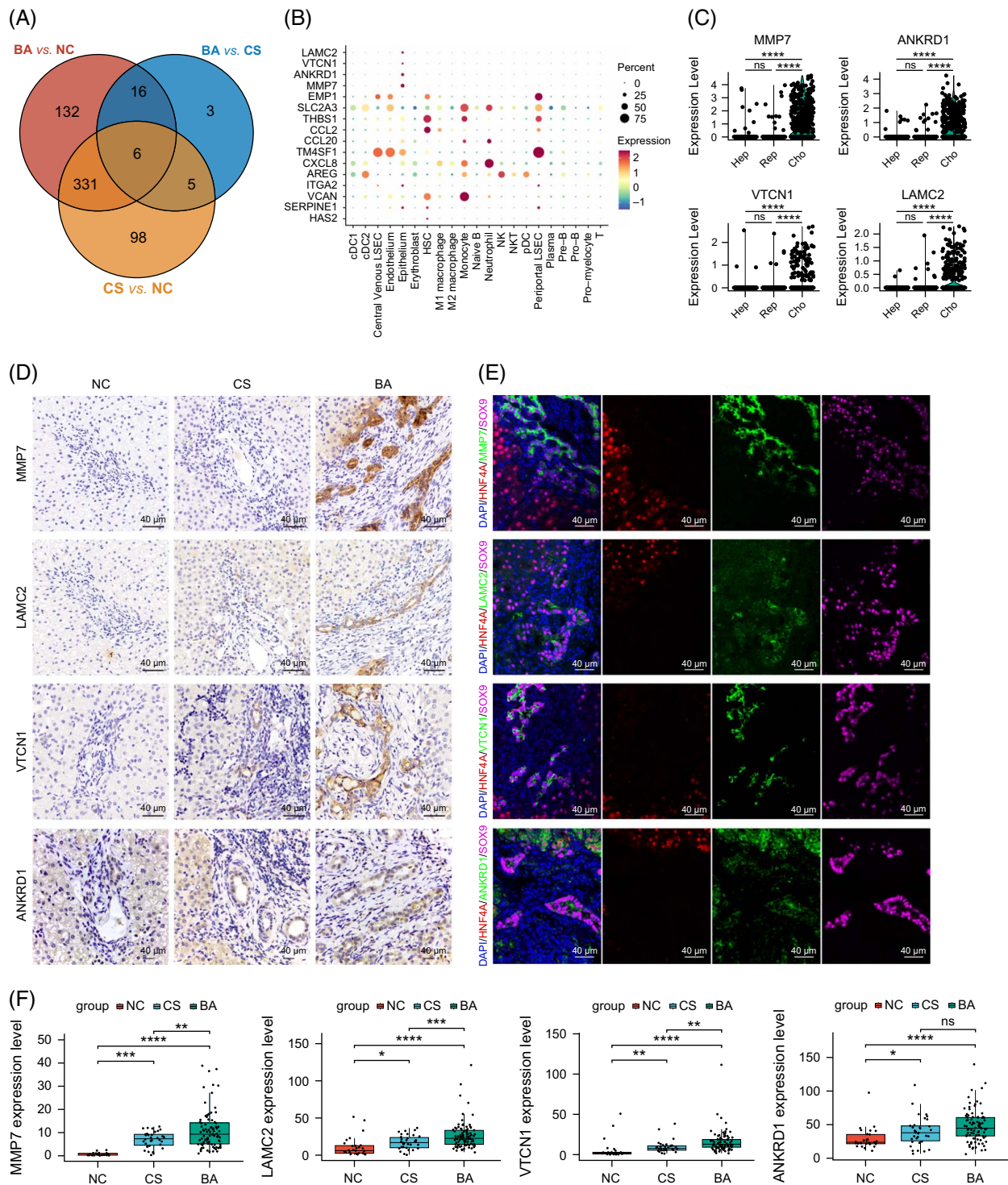


FIGURE 4 Upregulated biliary genes in BA. (A) Venn diagram showing 16 overlapping genes between the DEGs of BA versus NC and BA versus CS, but not the DEGs in CS versus NC. (B) Dot plots displaying the expression of the 16 overlapping genes in different cell types of BA and controls by single-cell RNA sequencing. (C) Violin plot showing the expressions of biliary genes in the epithelial cells of BA and controls. (D) Representative immunohistochemistry sections of MMP7, LAMC2, VTCN1, and ANKRD1 in TMA. (E) Representative immunofluorescence staining of MMP7, LAMC2, VTCN1, and ANKRD1 in TMA of the BA group. (F) Box plots showing the expression levels of the indicated genes in TMA, colored by different groups. ns, not significant; * p < 0.05; ** p < 0.01; *** p < 0.001; **** p < 0.0001. Abbreviations: BA, biliary atresia; CHO, cholangiocytes; CS, cholestasis; HEP, hepatocytes; DEG, differentially expressed gene; NC, normal control; REP, reprogrammed cells; TMA, tissue microarray.

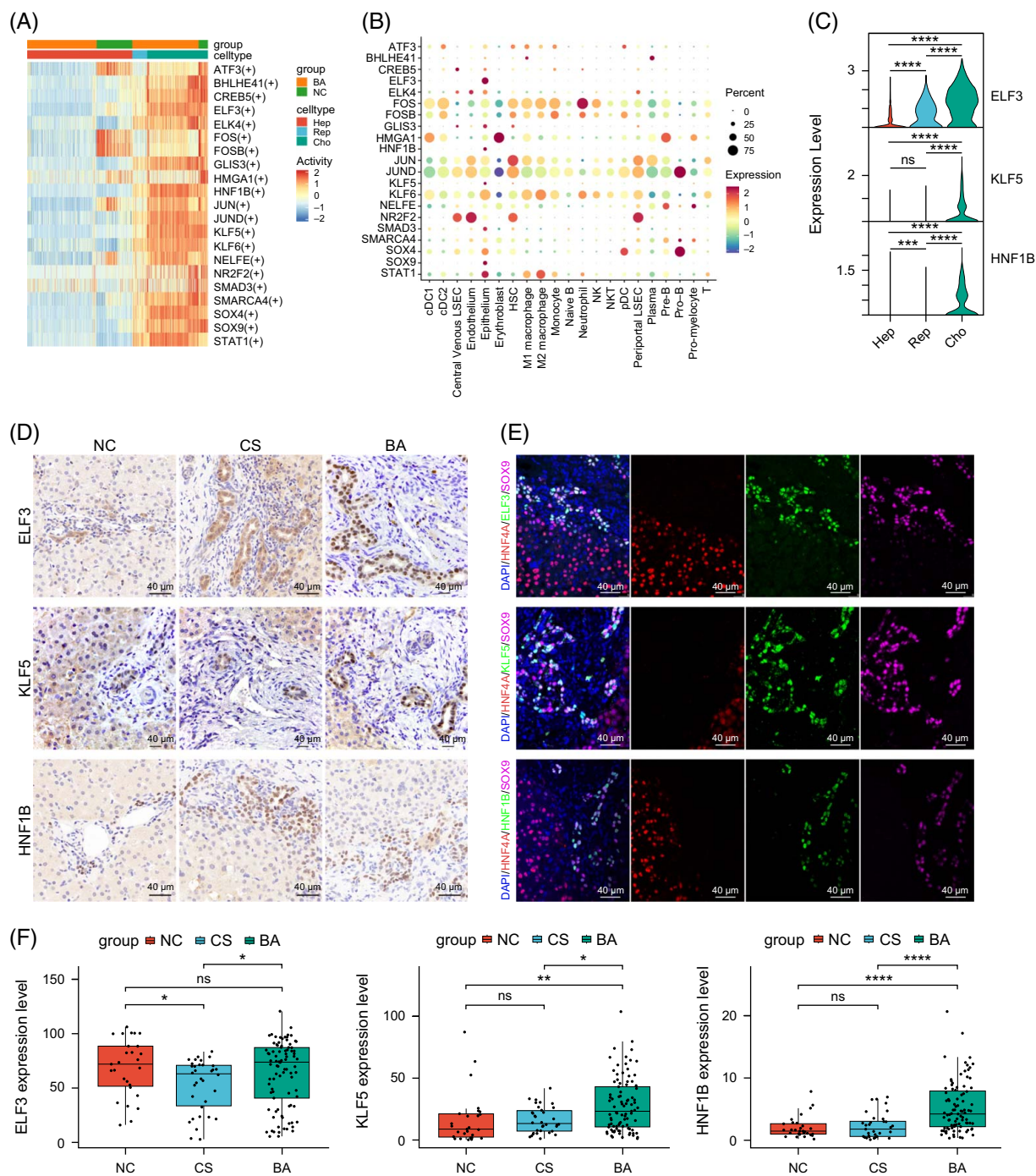


FIGURE 5 Upregulated biliary TFs in BA. (A) Heatmap illustrating the 21 TFs with high transcription activity in cholangiocytes. (B) Dot plots displaying the expression of 21 TFs in different cell types of BA and controls. (C) Violin plot demonstrated the expressions of biliary TFs in the epithelial cells of BA and controls. (D) Representative immunohistochemistry sections of ELF3, KLF5, and HNF1B in TMA. (E) Representative immunofluorescence staining of ELF3, KLF5, and HNF1B in TMA of the BA group. (F) Box plots showing the expression level of the indicated TFs in TMA, colored by different groups. * $p < 0.05$; ** $p < 0.01$; **** $p < 0.0001$. Abbreviations: BA, biliary atresia; CDC, conventional dendritic cell; CHO, cholangiocytes; CS, cholestasis; HEP, hepatocytes; DEG, differentially expressed gene; HSC, hepatic stellate cell; LSEC, liver sinusoidal endothelial cell; NC, normal control; NK, natural killer; pDC, plasmacytoid dendritic cell; REP, reprogrammed cells; TF, transcriptional factor; TMA, tissue microarray.

High-grade DR has been reported to be associated with faster postoperative progression of fibrosis and native liver survival in BA,^[35,36] indicating the detrimental role of DR in the pathogenesis of BA. Our study

dissected the functions of hepatocytes, reprogrammed cells, and cholangiocytes. As expected, reprogrammed cells and cholangiocytes demonstrated distinct functions when compared with hepatocytes. During biliary

reprogramming, the enrichment of EMT, fibrosis, inflammation, and mRNA metabolism was gradually enhanced and reached the highest levels in cholangiocytes. EMT pathology evidenced by higher mesenchymal markers and EMT TFs may be linked to biliary dysmorphogenesis and delayed epithelial development in BA.^[10,37] As expected, our data suggested that DR exerted the fibroinflammatory role in BA. The proliferating BECs in portal area can produce pro-inflammatory factors, chemokines and osteopontin to promote HSC activation in other cholestatic diseases.^[6,7] Apart from the pathological role of CXCL8 in BA,^[25] CXCL6 appeared to be a novel chemokine that was exclusively secreted by cholangiocytes. CXCL6 can recruit neutrophils to the inflammation sites by interacting with its receptors CXCR1 and CXCR2,^[38] and was associated with liver fibrosis.^[39] Our data suggested that PDGFA may be the profibrotic proteins that were excreted by cholangiocytes and endothelial cells in BA. Our recent study highlighted the significance of m⁶A methylation in liver fibrosis in BA.^[16] In this study, the higher expressions of m⁶A readers in cholangiocytes may extend the role of m⁶A modification in BA cholangiopathy.

Identification of DR markers may assist the diagnosis and disease evaluation in BA. In most studies, CK7 was the marker of DR, BA patients with the advanced stages of liver fibrosis and non-native liver survival state had higher levels of CK7 expressions when compared with the corresponding controls.^[35,36] In this study, CK7 was upregulated in reprogrammed cells and cholangiocytes and was the representative of all transdifferentiated cells in DR. Our data demonstrated that MMP7 was the alternative marker of DR. We and others reported that MMP7 can specifically discriminate BA from other causes of neonatal CS,^[12–15] and was also correlated with liver fibrosis and native liver survival in BA.^[11,14,15,40] LAMC2 has been proposed to be a specific biomarker for BA by RNA measurement.^[25] Herein, we first demonstrated the protein expression of LAMC2 in the liver and found that it was specifically produced by cholangiocytes. The most striking finding is the novel DR marker VTCN1. Our data has demonstrated that it was upregulated in patients with BA^[41] and biliary atresia-induced BA mouse model,^[42] and may be related to the inhibition of T-cell infiltration.^[43]

Screening out the biliary TFs may be conducive to figuring out the molecular mechanism of biliary pathology in BA. HNF1B is known to participate in biliary system organogenesis, and deletion of HNF1B leads to severe jaundice and biliary hypoplasia in mice^[44] and ciliary defects in human cholangiocytes.^[45] The higher expression of HNF1B may contribute to the expansion of bile ducts to relieve CS in BA. KLF5, a YAP downstream gene,^[46] plays a critical role in DR by inducing BEC proliferation.^[47] DR is a biliary tissue-remodeling process for liver injury; the elevated levels of

KLF5 may maintain the biliary epithelial tissue expansion to alleviate the liver injury in BA. ELF3 has been linked with the biliary tract cancer or gallbladder cancer.^[48,49] The role of ELF3 in DR remains unknown, it may be involved in biliary inflammation according to a scRNA-seq study.^[50]

In conclusion, to our knowledge, this is the first study that comprehensively depicts the epithelial transcription profile in BA from a single-cell perspective. We elucidated the origin, function, gene markers, and TFs of cholangiocytes in BA. All these findings may assist in the understanding of the DR in the pathogenesis of BA, and the screened gene markers and TFs may shed new light on the diagnosis of BA and the exploitation of target interventions to block BA-induced liver injury in the future.

AUTHOR CONTRIBUTIONS

Junfeng Wang, Zhen Shen, and Gong Chen designed and supervised the study. Lingdu Meng and Min Du conducted the experiments, performed the data analysis, and wrote the manuscript. Haodong Li provided assistance in data analysis. Fanyang Kong collected the clinical information of all patients. Jiajian Yang collected the liver samples. Rui Dong and Shan Zheng reviewed and edited the manuscript. All authors read and approved the manuscript.

FUNDING INFORMATION

This study received financial support from the National Natural Science Foundation of China (no. 82201915, no. 82270541), the National Clinical Key Specialty Construction Project (no. 10000015Z155080000004), Shanghai Municipal Health Commission (2024ZZ2035), Shanghai Municipal Key Clinical Specialty (no. shslc zdzk05703), and Xiamen Key Laboratory of Pediatric Surgical Disease (CHP-2023-XKL-005).

CONFLICTS OF INTEREST


The authors have no conflicts to report.

ORCID

Lingdu Meng  <https://orcid.org/0000-0002-0048-0687>

Min Du  <https://orcid.org/0000-0001-7776-3987>

Haodong Li  <https://orcid.org/0009-0000-5822-7353>

Fanyang Kong  <https://orcid.org/0009-0001-2923-3268>

Jiajian Yang  <https://orcid.org/0000-0001-9175-6572>

Rui Dong  <https://orcid.org/0000-0001-9645-4768>

Shan Zheng  <https://orcid.org/0000-0002-9712-4573>

Gong Chen  <https://orcid.org/0000-0003-3351-0819>

Zhen Shen  <https://orcid.org/0000-0002-6287-6169>

Junfeng Wang  <https://orcid.org/0000-0001-5477-1239>

REFERENCES

- Asai A, Miethke A, Bezerra JA. Pathogenesis of biliary atresia: Defining biology to understand clinical phenotypes. *Nat Rev Gastroenterol Hepatol*. 2015;12:342–52.

2. Sato K, Marzioni M, Meng F, Francis H, Glaser S, Alpini G. Ductular reaction in liver diseases: Pathological mechanisms and translational significances. *Hepatology*. 2019;69:420–30.
3. Merrell AJ, Peng T, Li J, Sun K, Li B, Katsuda T, et al. Dynamic transcriptional and epigenetic changes drive cellular plasticity in the liver. *Hepatology*. 2021;74:444–57.
4. Vartak N, Damle-Vartak A, Richter B, Dirsch O, Dahmen U, Hammad S, et al. Cholestasis-induced adaptive remodeling of interlobular bile ducts. *Hepatology*. 2016;63:951–64.
5. Glaser SS, Gaudio E, Miller T, Alvaro D, Alpini G. Cholangiocyte proliferation and liver fibrosis. *Expert Rev Mol Med*. 2009;11:e7.
6. Lazaridis KN, Strazzabosco M, Larusso NF. The cholangiopathies: Disorders of biliary epithelia. *Gastroenterology*. 2004;127:1565–77.
7. Wang X, Lopategi A, Ge X, Lu Y, Kitamura N, Urtasun R, et al. Osteopontin induces ductular reaction contributing to liver fibrosis. *Gut*. 2014;63:1805–18.
8. Azad AI, Krishnan A, Troop L, Li Y, Katsumi T, Pavelko K, et al. Targeted apoptosis of ductular reactive cells reduces hepatic fibrosis in a mouse model of cholestasis. *Hepatology*. 2020;72:1013–28.
9. Harada K, Sato Y, Ikeda H, Isse K, Ozaki S, Enomae M, et al. Epithelial-mesenchymal transition induced by biliary innate immunity contributes to the sclerosing cholangiopathy of biliary atresia. *J Pathol*. 2009;217:654–64.
10. Omenetti A, Bass LM, Anders RA, Clemente MG, Francis H, Guy CD, et al. Hedgehog activity, epithelial-mesenchymal transitions, and biliary dysmorphogenesis in biliary atresia. *Hepatology*. 2011;53:1246–58.
11. Huang CC, Chuang JH, Chou MH, Wu CL, Chen CM, Wang CC, et al. Matrilysin (MMP-7) is a major matrix metalloproteinase upregulated in biliary atresia-associated liver fibrosis. *Mod Pathol*. 2005;18:941–50.
12. Lertudomphonwanit C, Mourya R, Fei L, Zhang Y, Gutta S, Yang L, et al. Large-scale proteomics identifies MMP-7 as a sentinel of epithelial injury and of biliary atresia. *Sci Transl Med*. 2017;9:e20190902.
13. Yang L, Zhou Y, Xu P, Mourya R, Lei H, Cao G, et al. Diagnostic accuracy of serum matrix metalloproteinase-7 for biliary atresia. *Hepatology*. 2018;68:2069–77.
14. Jiang J, Wang J, Shen Z, Lu X, Chen G, Huang Y, et al. Serum MMP-7 in the diagnosis of biliary atresia. *Pediatrics*. 2019;144:e20190902.
15. Wu JF, Jeng YM, Chen HL, Ni YH, Hsu HY, Chang MH. Quantification of serum matrix metalloproteinase 7 levels may assist in the diagnosis and predict the outcome for patients with biliary atresia. *Journal of Pediatrics*. 2019;208:30–37.
16. Wang J, Du M, Meng L, Yang Y, He S, Zhu Y, et al. Integrative analysis implicates the significance of m6A in the liver fibrosis of biliary atresia by regulating THY1. *Hepatol Commun*. 2023;7:e0004.
17. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184:3573–587.e29.
18. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*. 2019;16:1289–96.
19. Yu G, Wang LG, Han Y, He QY. clusterProfiler: An R package for comparing biological themes among gene clusters. *Omics*. 2012;16:284–7.
20. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform*. 2013;14:7.
21. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017;14:979–82.
22. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol*. 2020;38:1408–14.
23. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: Single-cell regulatory network inference and clustering. *Nat Methods*. 2017;14:1083–6.
24. Zhao Z, Zhou W, Han Y, Peng F, Wang R, Yu R, et al. EMT-Regulome: A database for EMT-related regulatory interactions, motifs and network. *Cell Death & Disease*. 2017;8:e2872.
25. Bessho K, Mourya R, Shivakumar P, Walters S, Magee JC, Rao M, et al. Gene expression signature for biliary atresia and a role for interleukin-8 in pathogenesis of experimental disease. *Hepatology*. 2014;60:211–23.
26. Tarlow BD, Pelz C, Naugler WE, Wakefield L, Wilson EM, Finegold MJ, et al. Bipotential adult liver progenitors are derived from chronically injured mature hepatocytes. *Cell Stem Cell*. 2014;15:605–18.
27. Schaub JR, Huppert KA, Kurial SNT, Hsu BY, Cast AE, Donnelly B, et al. De novo formation of the biliary system by TGF β -mediated hepatocyte transdifferentiation. *Nature*. 2018;557:247–51.
28. Katsuda T, Sussman J, Li J, Merrell AJ, Vostrejs W, Secreto A, et al. Evidence for in vitro extensive proliferation of adult hepatocytes and biliary epithelial cells. *Stem Cell Reports*. 2023;18:1436–50.
29. Yang L, Wang WH, Qiu WL, Guo Z, Bi E, Xu CR. A single-cell transcriptomic analysis reveals precise pathways and regulatory mechanisms underlying hepatoblast differentiation. *Hepatology*. 2017;66:1387–401.
30. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566:496–502.
31. Renzi A, DeMorrow S, Onori P, Carpino G, Mancinelli R, Meng F, et al. Modulation of the biliary expression of arylalkylamine N-acetyltransferase alters the autocrine proliferative responses of cholangiocytes in rats. *Hepatology*. 2013;57:1130–41.
32. Raven A, Lu WY, Man TY, Ferreira-Gonzalez S, O'Duibhir E, Dwyer BJ, et al. Cholangiocytes act as facultative liver stem cells during impaired hepatocyte regeneration. *Nature*. 2017;547:350–4.
33. Michalopoulos GK, Barua L, Bowen WC. Transdifferentiation of rat hepatocytes into biliary cells after bile duct ligation and toxic biliary injury. *Hepatology*. 2005;41:535–44.
34. Lin Y, Zhang F, Zhang L, Chen L, Zheng S. Characteristics of SOX9-positive progenitor-like cells during cholestatic liver regeneration in biliary atresia. *Stem Cell Research & Therapy*. 2022;13:114.
35. Santos JL, Kielsing CO, Meurer L, Vieira S, Ferreira CT, Lorentz A, et al. The extent of biliary proliferation in liver biopsies from patients with biliary atresia at portoenterostomy is associated with the postoperative prognosis. *J Pediatr Surg*. 2009;44:695–701.
36. Hukkinen M, Kerola A, Lohi J, Heikkilä P, Merras-Salmio L, Jahnukainen T, et al. Treatment policy and liver histopathology predict biliary atresia outcomes: Results after National Centralization and Protocol Biopsies. *J Am Coll Surg*. 2018;226:46–57.
37. Amarachintha SP, Mourya R, Ayabe H, Yang L, Luo Z, Li X, et al. Biliary organoids uncover delayed epithelial development and barrier function in biliary atresia. *Hepatology*. 2022;75:89–103.
38. Wolf M, Delgado MB, Jones SA, Dewald B, Clark-Lewis I, Baggiolini M. Granulocyte chemotactic protein 2 acts via both IL-8 receptors, CXCR1 and CXCR2. *Eur J Immunol*. 1998;28:164–70.
39. Cai X, Li Z, Zhang Q, Qu Y, Xu M, Wan X, et al. CXCL6-EGFR-induced Kupffer cells secrete TGF- β 1 promoting hepatic stellate cell activation via the SMAD2/BRD4/C-MYC/EZH2 pathway in liver fibrosis. *J Cell Mol Med*. 2018;22:5050–61.
40. Kerola A, Lampela H, Lohi J, Heikkilä P, Mutanen A, Hagström J, et al. Increased MMP-7 expression in biliary epithelium and serum underpins native liver fibrosis after successful portoenterostomy in biliary atresia. *J Pathol Clin Res*. 2016;2:187–98.

41. Dong R, Yang Y, Shen Z, Zheng C, Jin Z, Huang Y, et al. Forkhead box A3 attenuated the progression of fibrosis in a rat model of biliary atresia. *Cell Death Dis.* 2017;8:e2719.
42. Yang Y, Wang J, Zhan Y, Chen G, Shen Z, Zheng S, et al. The synthetic toxin biliary atresia causes biliary atresia in mice. *Lab Invest.* 2020;100:1425–35.
43. Zhu C, Ma J, Zhu K, Yu L, Zheng B, Rao D, et al. Spatial immunophenotypes predict clinical outcome in intrahepatic cholangiocarcinoma. *JHEP Rep.* 2023;5:100762.
44. Coffinier C, Gresh L, Fiette L, Tronche F, Schütz G, Babinet C, et al. Bile system morphogenesis defects and liver dysfunction upon targeted deletion of HNF1beta. *Development.* 2002;129:1829–38.
45. Roelandt P, Antoniou A, Libbrecht L, Van Steenberghe W, Laleman W, Verslype C, et al. HNF1B deficiency causes ciliary defects in human cholangiocytes. *Hepatology.* 2012;56:1178–81.
46. Zheng C, Luo J, Yang Y, Dong R, Yu F-X, Zheng S. YAP activation and implications in patients and a mouse model of biliary atresia. *Front Pediatr.* 2021;8:618226.
47. Okada H, Yamada M, Kamimoto K, Kok CYY, Kaneko K, Ema M, et al. The transcription factor Klf5 is essential for intrahepatic biliary epithelial tissue remodeling after cholestatic liver injury. *J Biol Chem.* 2018;293:6214–29.
48. Suzuki M, Saito-Adachi M, Arai Y, Fujiwara Y, Takai E, Shibata S, et al. E74-like factor 3 is a key regulator of epithelial integrity and immune response genes in biliary tract cancer. *Cancer Res.* 2021;81:489–500.
49. Pandey A, Stawiski EW, Durinck S, Gowda H, Goldstein LD, Barbhuiya MA, et al. Integrated genomic analysis reveals mutated ELF3 as a potential gallbladder cancer vaccine candidate. *Nat Commun.* 2020;11:4225.
50. MacParland SA, Liu JC, Ma XZ, Innes BT, Bartczak AM, Gage BK, et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat Commun.* 2018;9:4383.

How to cite this article: Meng L, Du M, Li H, Kong F, Yang J, Dong R, et al. Single-cell transcription reveals hepatocyte-to-cholangiocyte reprogramming and biliary gene profile in biliary atresia. *Hepatol Commun.* 2025;9:e0710. <https://doi.org/10.1097/HC9.0000000000000710>