mRNA and IncRNA expression profiles of liver tissues in children with biliary atresia

WENYAN WU1,2*, WEIFANG WU3,4*, YONGQIN YE4,5, TAO LI2 and BIN WANG4

1Medical Laboratory, Shenzhen Luohu People's Hospital, Shenzhen, Guangdong 518001; 2Guangdong Provincial Key Laboratory of Medical Molecular Diagnostics, Guangdong Medical University, Dongguan, Guangdong 523000; 3Medical College, Shantou University Medical College, Shantou, Guangdong 515041; 4Department of General Surgery, Shenzhen Children's Hospital, Shenzhen, Guangdong 518026; 5Faculty of Medicine, Macau University of Science and Technology, Macau SAR 999078, P.R. China

Received May 18, 2022; Accepted August 1, 2022

DOI: 10.3892/etm.2022.11571

Abstract. Progressive liver fibrosis is the most common phenotype in biliary atresia (BA). A number of pathways contribute to the fibrosis process so comprehensive understanding the mechanisms of liver fibrosis in BA will pave the way to improve patient's outcome after operation. In this study, the differentially expressed profiles of mRNAs and long non-coding RNAs from BA and choledochal cyst (CC) liver tissues were investigated and analyzed, which may provide potential clues to clarify hepatofibrosis mechanism in BA. A total of two BA and two CC liver tissue specimens were collected, the expression level of mRNAs and IncRNAs was detected by RNA sequencing. Differentially expressed mRNAs (DEmRNAs) were functionally annotated and protein-protein interaction networks (PPI) was established to predict the biological roles and interactive relationships. Differentially expressed IncRNAs (DEIncRNAs) nearby targeted DEmRNA network and DEIncRNA-DEmRNA co-expression network were constructed to further explore the roles of DEIncRNAs in BA pathogenesis. The expression profiles of significant DEmRNAs were validated in Gene Expression Omnibus database. A total of 2,086 DEmRNAs and 184 DEIncRNAs between BA and CC liver tissues were obtained. DEmRNAs were enriched in 521 Gene Ontology terms and 71 Kyoto Encyclopedia of Genes and Genomes terms which were mainly biological processes and metabolic pathways related to immune response and inflammatory response. A total of five hub proteins (TYRO protein tyrosine kinase binding protein, C-X-C motif chemokine ligand 8, plekstrin, Toll-like receptor 8 and C-C motif chemokine receptor 5) were found in the PPI networks. A total of 31 DEIncRNA-nearby-targeted DEmRNA pairs and 2,337 DEIncRNA-DEmRNA co-expression pairs were obtained. The expression of DEmRNAs obtained from RNA sequencing were verified in GSE46960 dataset, generally. The present study identified key genes and IncRNAs participated in BA associated liver fibrosis, which may present a new avenue for understanding the patho-mechanism for hepatic fibrosis in BA.

Introduction

Biliary atresia (BA) is a rare but severe neonatal disease characterized by an inflammatory and progressive fibrotic obstruction of the extrahepatic bile ducts resulting in cholestasis and subsequent hepatic failure if left untreated (1). Rapid liver fibrosis is a major outcome for children with BA even though they might appear normal at birth (2,3). Knowledge on the pathogenesis of liver fibrosis in BA is still limited. Considering the complicated etiologies, the identify potentially modifiable factors in BA associated liver fibrosis is an urgent need (2). The microarray and RNA-seq technique can help to identify signatures of predominant transcriptions of the liver during fibrosis progressive in BA, which could be beneficial for uncovering new molecular mechanisms to improve the prognosis, diagnosis and treatment of this disease (4). This technique has been successfully applied for liver tissues from BA and CC to investigate the potential underlying mechanism (5-8). A transcriptional analysis of liver tissues from BA and CC identified 877 differentially expressed genes (such as COX1, SCO2, and CYTB, which are involved in oxidative phosphorylation) and several pathways with immune and inflammatory responses related to the Wnt-signaling pathway and TGF-β pathway (5). Previous studies have found that the NF-κB signaling pathway could regulate genes related to immune response and mediate liver fibrosis by initiating fibrosis factors and thus might serve a key role in the development of BA (9-11). Another RNA sequencing of livers from 6 BA and 6 CC found that Foxa3 may exert a role in the progression of hepatic fibrosis in BA, which may

*Contributed equally

Key words: biliary atresia, liver fibrosis, mRNA, long non-coding RNA, bioinformation analysis
be a potential targeted treatment (6). Fas ligand mRNA (7), MMP7 and PCK1 (8) can be used as potential biomarkers to predict the outcome of BA and the fibrosis progression. A comparison of mRNA expression level from normal, diseased control and end-stage BA livers identified 35 genes involved in cell signaling, transcription regulation, hepatobiliary development and fibrosis process (12). The expression level of genes regulating fibrosis in liver tissues increases in infants who survived within 2 years (13). Furthermore, cDNA array demonstrate two hepatofibrosis-associated genes (4), plasminogen activator inhibitor-1 (PAI-1) and tissue inhibitor of metalloproteinase-1 (TIMP-1).

Long noncoding RNAs (lncRNAs) exert specific regulatory functions in various cellular organization (14), which are involved in hepatic fibrosis by regulating gene or protein expression (15). LncRNAs can act as competing endogenous RNAs (ceRNAs) to participate in hepatic fibrosis (16). The expression level of long noncoding RNA H19 (lncRNA H19) is positively correlated with the severity of fibrotic liver injuries in BA patients, which serves a pivotal role in BA cholangiocyte proliferation and cholestatic liver injury as a sponge for let-7/HMGA2 axes and regulates SIPR2/SphK2 (17). LncRNA-Annexin A2 pseudogene 3 (ANXA2P3) is identified to participate in the process of BA-induced liver fibrosis by regulating Annexin A2 (ANXA2) (18). Our previous study on lncRNA-adducin 3 antisense RNA1 was identified as participating in liver fibrosis by targeting ADD3 in vitro accelerating the proliferation and migration of LX-2 cells, which are the key cells involved in hepatic fibrosis (19). The construction and analysis of regulation network based on microarray and RNA-seq results can improve understanding of the mechanisms underlying BA associated hepatic fibrosis. More evidence will pave the way to fully understand hepatic fibrosis pathogenesis in BA patients (20).

The present study aimed to identify some key DElncRNAs and DEmRNAs in liver tissues of BA patients through RNA-sequencing and also to discover some hub proteins via constructing PPI networks of DEmRNAs and DElncRNA-DEmRNA co-expression networks. Through the aforementioned approaches, it was expected to ascertain some lncRNAs, mRNAs and pathways that serve an important role in the pathogenesis of liver fibrosis with BA.

Materials and methods

Patient samples. The present study was approved by Ethics Committee on Human Research of Shenzhen Children's Hospital (approval no. 202106602) and was conducted in accordance with the principles expressed in the 1975 Declaration of Helsinki. A total of two BA infants and two age-matched choledochal cyst (CC) infants with a normal liver function were included in the present study. All the enrolled infants were confirmed diagnosed intraoperative cholangiography by the same surgical group in Shenzhen Children's Hospital. Liver biopsy tissues were collected with a written informed consent from all infants' guardians. The patient's clinical information is presented in Table I.

RNA isolation and sequencing. Total RNA was extracted from liver tissue samples using TRizol® (Thermo Fisher Scientific, Inc.). Nanodrop (Thermo Fisher Scientific, Inc.) was used for preliminary qualification and quantification of RNA concentration and Agilent 2100 (Agilent Technologies, Inc.) was used for precise detection of quality of RNA library. Agarose gel electrophoresis was used to detect the integrity of retracted RNA.

After the RNA samples were qualified, rRNA was removed with Ribo-zero kit (EpiCentre; Illumina, Inc.) and then the RNA was fragmented under high temperature and metal ions. Using ribosomal-depleted RNA as a template, first-strand cDNA was synthesized with random hexamers. Subsequently, second-stranded cDNA was synthesized by adding buffer, dNTPs (dUTP, dATP, dGTP and dCTP) and enzymes, followed by purification of double-stranded cDNA using VAHTSTM DNA Clean Beads (Vazyme Biotech Co., Ltd.). The final strand-specific cDNA library was obtained through a series of experiments such as end repair, tailing, sorting, digestion of cDNA containing U by using UDG enzyme and PCR enrichment. Illumina HiSeq X Ten platform (Illumina, Inc.) sequencing was performed after pooling different libraries according to the requirements of effective concentration and target data volume. The processing and database construction for RNA-seq are in Fig. S1 and Appendix S1.

Identification of DElncRNAs and DEmRNAs. The reference sequences of the corresponding species were downloaded from the database Ensemble GRCh38.p7 (ftp://ftp.ncbi.nlm.nih.gov/genomes/Homo_sapiens) for comparison with the sequencing data. Expression of mRNAs and lncRNAs was normalized and outputted with StringTie version 1.3.3b (http://ccb.jhu.edu/software/stringtie/), log2FC>1 and P<0.05 were used as the cut-off criteria. Volcano plots and hierarchical clustering were used to visualize the overall distribution of differential transcripts.

Functional annotation and pathway enrichment of DEmRNAs. DAVID (https://david.ncifcrf.gov) is an online database that aggregates large numbers of genes or proteins into corresponding functional annotations and pathways, providing a quick access to various annotation data. P<0.05 was used
as the cutoff criterion for Gene Ontology (GO) functional enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis.

Protein-protein interaction (PPI) networks construction. The STRING protein interaction database (http://string-db.org/) and the R language (http://www.R-project.org/) package STRINGdb were used for Top 100 up- and downregulated DEmRNAs protein interaction network analysis. Cytoscape software (version 3.5.0, http://www.cytoscape.org) was applied to visualize PPI networks.

Nearby-targeted DEmRNAs of DElncRNAs. The functions of lncRNAs may be related to their adjacent protein-coding mRNAs (21,22), so the protein-coding mRNAs (100-kb upstream and downstream) adjacent to lncRNAs were selected as their target mRNAs. The lncRNA-mRNA regulated networks were visualized by Cytoscape software (version 3.5.0, http://www.cytoscape.org) was applied to visualize PPI networks.

DElncRNA-DEmRNA co-expression networks. DElncRNA-DEmRNA co-expression was calculated based on the dynamic changes of lncRNA and mRNA expression signal values, including expression regulation relationship and regulation direction. The DElncRNA-DEmRNA pairs with absolute values of PCC >0.99 and P<0.01 were selected and DElncRNA-DEmRNA co-expression network was constructed and visualized by Cytoscape software (version 3.5.0, http://www.cytoscape.org). Functional annotation and Pathway Enrichment of the DEmRNAs co-expressed with DElncRNAs were performed with DAVID. P<0.05 was used as the criterion.

Validation in the Gene Expression Omnibus (GEO) dataset. The mRNA expression profiles (GSE46960) was downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), which included a collection of 64 liver biopsy sample obtained from a patient with BA (case group) and 7 liver biopsy sample obtained from a deceased-donor adult (normal group). The expression of screened DEmRNAs obtained from our RNA sequencing were verified using the GSE46960 dataset.

Statistical analysis. Sequencing data were analyzed by using fold changes (FC) and Student’s t-test. |log2FC|>1 and P<0.05 were set as the criterion to identify the differentially expressed mRNAs and lncRNAs.

Results

DEmRNAs and DElncRNAs between BA liver tissues and CC liver tissues. There were 2,086 DEmRNAs (1,036 up- and 1,050 downregulated) and 184 DElncRNAs (52 up- and 132 downregulated) in the BA group. The top 10 upregulated and downregulated DEmRNAs and DElncRNAs are in Tables II and III. Volcano maps of DEmRNAs and DElncRNAs are in Fig. 1A and B, respectively. Hierarchical clustering of DEmRNAs and DElncRNAs is shown in Fig. 1C and D. The raw data of the sequencing data of the present study have been uploaded to Sequence Read Archive (SRA) for SUB7923751 at https://www.ncbi.nlm.nih.gov/sra/PRJNA701623.

Table II. Top 10 up- and downregulated DEmRNAs between BA liver tissues and CC liver tissues.

<table>
<thead>
<tr>
<th>ID</th>
<th>Symbol</th>
<th>log2FC</th>
<th>P-value</th>
<th>Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSG00000244734</td>
<td>HBB</td>
<td>6.99799</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000147257</td>
<td>GPC3</td>
<td>5.86627</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000159217</td>
<td>IGF2BP1</td>
<td>5.61870</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000081051</td>
<td>AFP</td>
<td>5.41660</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000068366</td>
<td>ACSL4</td>
<td>5.21616</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000276368</td>
<td>HIST1H2AJ</td>
<td>4.71279</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000169213</td>
<td>RAB3B</td>
<td>4.47424</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000274267</td>
<td>HIST1H3B</td>
<td>4.37440</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000150099</td>
<td>CCL20</td>
<td>4.33093</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000274641</td>
<td>HIST1H2BO</td>
<td>4.25590</td>
<td>5.00x10^-5</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000140505</td>
<td>CYP1A2</td>
<td>-7.78389</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000117594</td>
<td>HSD11B1</td>
<td>-5.94395</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000122787</td>
<td>AKR1D1</td>
<td>-4.81138</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000124713</td>
<td>GNMT</td>
<td>-4.80215</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000180875</td>
<td>GREM2</td>
<td>-4.63967</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000160868</td>
<td>CYP3A4</td>
<td>-4.56139</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000205362</td>
<td>MT1A</td>
<td>-4.36469</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000134760</td>
<td>DSG1</td>
<td>-4.21303</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000171234</td>
<td>UGT2B7</td>
<td>-4.15848</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000166415</td>
<td>WDR72</td>
<td>-4.1164</td>
<td>5.00x10^-5</td>
<td>Down</td>
</tr>
</tbody>
</table>

DEmRNAs, differentially expressed mRNAs; FC, fold change.
Functional annotation of DEmRNAs. To improve understanding of the biological functions of DEmRNAs, GO and KEGG enrichment analyses were performed. Oxidation-reduction process (P=1.52x10\(^{-21}\)), immune response (P=3.10x10\(^{-17}\)), inflammatory response (P=4.12x10\(^{-16}\)) and metabolic process (P=1.09x10\(^{-11}\)) were significantly enriched GO biological process in BA (Fig. 2A). Extracellular exosome (P=3.47x10\(^{-45}\)) and mitochondrial matrix (P=1.35x10\(^{-18}\)) were significantly enriched GO cellular component in BA (Fig. 2B). Electron carrier activity (P=5.30x10\(^{-16}\)) and pyridoxal phosphate binding (P=1.14x10\(^{-12}\)) were significantly enriched GO molecular function in BA (Fig. 2C). Metabolic pathways (P=4.63x10\(^{-18}\)), Biosynthesis of antibiotics (P=5.88x10\(^{-17}\)), Glycine, serine and threonine metabolism (P=2.19x10\(^{-13}\)) and complement and coagulation cascades (P=3.19x10\(^{-12}\)) were top 4 enriched KEGG pathways in BA (Fig. 2D).

Protein-protein interaction (PPI) networks. The PPI networks contained 176 nodes and 665 edges. TYRO protein tyrosine kinase binding protein (degree=29), C-X-C motif chemokine ligand 8 (degree=27), Toll-like receptor (TLR) 8 (degree=24), pleckstrin (degree=23), C-C motif chemokine receptor (CCR)5 (degree=22), CCR1 (degree=22), lysosomal-associated protein transmembrane 5 (degree=21), TLR7 (degree=21), cytochrome P450 family 3 subfamily A member 4 (degree=21) and CD53 molecule (degree=20) were the top 10 hub proteins of the PPI networks (Fig. 3).

DElncRNAs-nearby-targeted DEmRNAs. There were 31 DElncRNAs-nearby-targeted DEmRNA pairs in total, of which 31 DElncRNAs and 30 DEmRNA were detected (Fig. 4). CTD-2256P15.2 and XLOC_045318 all nearby CMBL.

DElncRNA-DEmRNA co-expression networks. There were 2,337 DElncRNA-DEmRNA co-expression pairs in total, of which 54 DElncRNAs and 848 DEmRNAs were identified with an absolute value of PCC >0.99 and P<0.01 (Fig. 5). The positive regulated network is shown in Fig. 5A and the negative regulated network shown in Fig. 5B and C.

Functional annotation of DEmRNAs co-expressed with DElncRNAs. Inflammatory response (P=1.04x10\(^{-12}\)), immune response (P=1.64x10\(^{-09}\)), neutrophil chemotaxis (P=3.03x10\(^{-08}\)) and oxidation-reduction process (P=1.85x10\(^{-07}\)) were significantly enriched GO molecular function in BA (Fig. 6C). Oxidoreductase activity (P=3.77x10\(^{-05}\)) and aldehyde dehydrogenase activity (P=6.74x10\(^{-05}\)) were significantly enriched GO molecular function in BA (Fig. 6C). Complement and coagulation cascades (P=2.57x10\(^{-07}\)), Fatty acid degradation (P=2.85x10\(^{-06}\)), Drug metabolism-cytochrome P450 (P=1.20x10\(^{-04}\)) and Metabolic pathways (P=2.68x10\(^{-04}\)) were top 4 significantly enriched KEGG pathways (Fig. 6D).

Validation of selected DEmRNAs in GEO database. An mRNA expression profile was downloaded from GEO database and the expression patterns of DEmRNAs obtained from our RNA sequencing were verified. As shown in Fig. 7A and B, 111 mRNAs in upregulated DEmRNAs and

<table>
<thead>
<tr>
<th>ID</th>
<th>Symbol</th>
<th>log2FC</th>
<th>P-value</th>
<th>Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSG00000260604</td>
<td>RP1-140K8.5</td>
<td>2.61436</td>
<td>5.00x10(^{-5})</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000281508</td>
<td>CDR1-AS</td>
<td>2.19782</td>
<td>5.00x10(^{-5})</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000270507</td>
<td>CTA-21C21.1</td>
<td>3.44985</td>
<td>0.00010</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000232931</td>
<td>LINCO0342</td>
<td>1.56982</td>
<td>0.00435</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000259834</td>
<td>RP11-284N8.3</td>
<td>1.59598</td>
<td>0.00665</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000225470</td>
<td>JPX</td>
<td>25.47320</td>
<td>0.00705</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000277152</td>
<td>RP11-622C24.2</td>
<td>1.55379</td>
<td>0.00765</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000214049</td>
<td>UCA1</td>
<td>3.00640</td>
<td>0.00990</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000267731</td>
<td>RP11-147L13.8</td>
<td>1.79259</td>
<td>0.01435</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000172965</td>
<td>MIR4435-2HG</td>
<td>1.87816</td>
<td>0.01525</td>
<td>Up</td>
</tr>
<tr>
<td>ENSG00000230645</td>
<td>AC016682.1</td>
<td>-4.03424</td>
<td>5.00x10(^{-5})</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000261578</td>
<td>RP11-21L23.2</td>
<td>-3.94999</td>
<td>5.00x10(^{-5})</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000250056</td>
<td>LINCO1018</td>
<td>-2.43565</td>
<td>0.00070</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000248709</td>
<td>CTC-50503.2</td>
<td>-3.32413</td>
<td>0.00100</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000236378</td>
<td>RP11-394G3.2</td>
<td>-8.03214</td>
<td>0.00115</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000234509</td>
<td>AP000253.1</td>
<td>-3.18490</td>
<td>0.00145</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000251637</td>
<td>RP11-119D9.1</td>
<td>-3.81996</td>
<td>0.00275</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000261012</td>
<td>RP11-116D2.1</td>
<td>-1.07168</td>
<td>0.00285</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000260855</td>
<td>RP11-439E19.10</td>
<td>-1.57691</td>
<td>0.00395</td>
<td>Down</td>
</tr>
<tr>
<td>ENSG00000234456</td>
<td>MAG12-AS3</td>
<td>-2.02662</td>
<td>0.00420</td>
<td>Down</td>
</tr>
</tbody>
</table>

DElncRNAs, differentially expressed lncRNAs; FC, fold change.

Table III. Top 10 up- and downregulated DElncRNAs between BA liver tissues and CC liver tissues.
102 mRNAs in downregulated DEmRNAs were verified. Then when the present study focused on the top 10 upregulated DEmRNAs in BA from the results (Table II), five of them (HBB, GPC3, IGF2BP1, AFP and CCL20) were consistent with the GEO database in terms of expression patterns of mRNA (Fig. 7C). Similarly, mRNAs of CYP1A2, HSD11B1,
Figure 2. Top 20 significantly enriched GO terms and top 20 KEGG pathways of DEmRNAs between BA liver tissues and CC liver tissues. The x-axis is p-value for each GO term or pathway as -log10 and the y-axis indicates (A) biological process GO term, (B) cellular component GO term, (C) molecular function GO term or (D) KEGG pathway. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEmRNAs, differentially expressed mRNAs; BA, biliary atresia; CC, choledochal cyst.

Figure 3. PPI networks. Proteins encoded by up- and downregulated DEmRNAs between BA liver tissues and CC liver tissues were displayed as red and green ellipses. DEmRNAs derived from top 10 up- and downregulated DEmRNAs between BA liver tissues and CC liver tissues were displayed in ellipses with a black border. PPI, protein-protein interaction; DEmRNAs, differentially expressed mRNAs; BA, biliary atresia; CC, choledochal cyst.
AKR1D1, GNMT, MT1A and WDR72 were shown to be lower-expressed in the GEO database, which were consistent with the top 10 downregulated DEmRNAs from the results (Table II and Fig. 7D).

Discussion

Hepatic fibrosis usually develops in a progressive pattern in infants with BA. Even though these patients received Kasai portoenterostomy (KPE), the majority of them still have to undergo liver transplant due to severe fibrosis, cirrhosis and liver failure. However, the molecular network underpinning this expeditious fibrogenic process remains to be elucidated (23). In this regard, the present study performed RNA sequencing on liver samples of BA and CC to identify abundant DElncRNAs and DEmRNAs between two groups.

First, the results were analyzed for mRNA and lncRNA that were significantly upregulated or downregulated, including GPC3, AFP, CYP1A2, CCL20, 11βHSD1, AKR1D1, GNMT, JPX and MIR4435-2HG.

Glypican-3 (GPC3), a heparan sulfate proteoglycan attached to the cell membrane, serves a role in the regulation of the signaling activity involved in numerous growth factors (24), including Wnt/β-catenin (25), Hedgehogs (26), bone morphogenetic proteins and fibroblast growth factors (27). The Wnt/β-catenin signaling pathway has a key role in the regulation of cellular functions such as biliary cell fate (28), thereby raising a potential participation in BA etiology via GPC3-mediated...
Wnt/β-catenin signaling. Sirisomboonlarp et al (29) found that circulating GPC3 protein is significantly elevated in BA patients and positively correlated with liver stiffness, indicating that serum GPC3 could be a biomarker to evaluate hepatic function and to predict prognosis for BA patients following Kasai surgery. Notably, from the results of the present study, GPC3 was also identified to be the second most over-expressed DEmRNAs in BA liver tissues.

Serum α-fetoprotein (AFP) was strongly associated with epithelial liver tumors (30). It has a sensitivity of >90% and >70%, for hepatocellular carcinoma and hepatoblastoma (HB), respectively (31). Amir et al (32) found that patients with BA are far more prone to suffer from HB and in such cases, serum AFP levels might not be specific to diagnose HB in such cases, though remaining sensitive. A previous study (33) identified that ~8% of BA children were affected by liver tumors and AFP level screening and regular imaging examinations were crucial for early diagnosis of malignant liver tumors in BA patients.

Distinguishing BA from other non-BA neonatal cholestasis is challenging (34). Shteyer et al (35) show that Continuous Breath 13C-methacetin breath test might be useful to distinguish BA from non-BA. Methacetin is a substrate that can evaluate liver metabolic function by being metabolized by cytochrome P450 1A2 (CYP1A2) (36), which is negatively
correlated with hepatic fibrosis in patients with chronic hepatitis C viral infection (37,38) and is also found to be involved in other forms of hepatic diseases (39). Crawford et al (40) discovered that downregulated CYP1A2 could aggravate inflammatory response due to an increase of proinflammatory cytokines, such as TNF-α and IL-6. CYP1A2 is the most significant downregulated DEmRNAs in BA liver tissues, indicating that it might participate in this inflammatory fibrosclerosing disease and thus can be an indicator to differentiate BA. Therefore, further investigations into the intrinsic pathway of CYP1A2 on the etiology of progressive fibrosis in BA are of importance.

It is hypothesized that the discovery of more valuable biomarkers to monitor the progression of hepatic fibrosis following KPE may help improve clinical outcomes in patients and delay the need for liver transplantation (41), including IncRNA APTR (42), miR-21 (43) and mRNA FN14 (44). In the present study, the results also identified a number of valuable biomarkers related to liver fibrosis. Chemokine ligand-20 (CCL20) acted as a potent chemoattractant for immature dendritic cells, which can mediate the strong inflammatory responses that drive liver fibrosis (45). In vivo experiments show that a deficiency of 11β-hydroxysteroid dehydrogenase-1 (11βHSD1) could lead to the activation of hepatic myofibroblast and thus exacerbate liver fibrosis in mice (46). Expression level of steroid 5β-reductase (AKR1D1) in the liver tissue is negatively correlated with liver fibrosis and inflammation (47). Glycine N-methyltransferase (GNMT) is the most abundant

Figure 5. DElncRNA-DEmRNA co-expression networks, including (A) positive regulatory and negative regulated networks. (B and C). DElncRNAs and their nearby DEmRNAs between BA liver tissues and CC liver tissues were shown in triangles and ellipses, respectively. Up- and downregulation in BA liver tissues compared with CC liver tissues were displayed in red and green color, respectively. DElncRNAs/DEmRNAs derived from top 10 up- and downregulated DElncRNAs/DEmRNAs were displayed in triangles and ellipses with a black border. DElncRNAs, differentially expressed lncRNAs; DEmRNAs, differentially expressed mRNAs; BA, biliary atresia; CC, choledochal cyst.
methyltransferase in the liver and its reduction has a negative effect on the maintenance of normal liver functions progressing to fibrosis, cirrhosis and hepatocellular carcinoma (48).

Liver fibrosis is strongly associated with the activation of hepatic stellate cells. LncRNA XIST enhances eol-induced autophagy process and subsequently activates hepatic stellate cells through the miR‑29b/HMGB1 signaling pathway (49). LncRNA JPX, a non‑protein coding RNA transcribed by the X deactivated central gene, activates XIST transcription via an interaction with CCCTC‑binding factor (50). From the sequencing data, the present study identified that JPX was an upregulated DElncRNA. Although the regulatory relationship between JPX and XIST was established, the role it serves in hepatic fibrosis and BA remains to be explored in the future.

Kerola et al (51) found a decrease of the expression level of TGF‑β1 after successful HPE, indicating that TGF‑β1 might be involved in exacerbating liver fibrosis. LncRNA miR4435‑2HG can promote prostate carcinoma (52), non‑small‑cell lung carcinoma (53) and ovarian carcinoma (54) by interacting with TGF‑β signaling and thus lncRNA miR4435‑2HG‑mediated TGF‑β1 signaling pathway might be involved in liver fibrosis of BA.

Second, Cytoscape was used to select the top 10 hub genes of PPI network for further analysis and it revealed that CXCL8, TLR8, CCR5, TLR7 and CYP3A4 were associated with BA, revealing that these hub genes may be involved in the occurrence of BA.

CXCL8 regulates inflammation and immune response via chemotactance to neutrophils (55). Studies show that an elevation of serum CXCL8 together with its increase in liver may participate in inflammation and liver fibrosis in patients with BA and thus CXCL8 may be a potential biomarker for diagnosis, severity evaluation and outcome of BA (56,57).

The major inappropriate host immunological reactions against unknown ligands via the TLR cascades may trigger progressive inflammatory biliary destruction that manifests as BA in newborns or infants. TLR cascades may induce an excessive immune reaction against multiple ligands in the bile duct system, leading to a progressive destruction and inflammatory fibro‑obstruction of biliary ducts seen in BA. miR expression level of TLR8 is significantly elevated in liver tissues in the early BA group (58). The activation of TLR7 signaling pathway is crucial for cholangiocytes proliferation in rhesus rotavirus (RRV)‑infected mice models (59). In addition, TLR7 can recognize infectious pathogens, activated type 1 interferon and induce the expression of inflammatory response genes, including IL‑8 (60).

A previous study has shown that CC chemokines, including CCL2, CCL3, CCL5, stimulate THP‑1 cells to increase MMP‑9 protein levels in a dose‑dependent manner, indicating that the activation of immune cells and MMP are closely related to the occurrence of tissue inflammation (61). Leonhardt et al (62) identified >40 significantly differentially‑expressed genes in BA mice. Most of the upregulated genes in BA‑positive mice...
were involved in immune response, such as CCL2, CCL5, CCR5, CXCL9, CXCL10 and IL1F5. To clarify the pathogenesis of BA, the immune reaction during the process of liver fibrosis should be further examined.

Finally, lncRNA-targets were analyzed. In order to fully illustrate how lncRNAs participate in a disease, it is widely accepted to perform a lncRNA-mRNA co-expression analysis to explore the biological functions of lncRNAs by evaluating...
and analyzing their co-expressed mRNAs. In the co-expression networks of the present study, CYP2E1 was a target for H19, H19 was upregulated and CYP2E1 was downregulated. Xiao et al (17) discovered that the levels of H19 in both liver and serum exosome increase in line with the severity of fibrosis in patients with BA. De Bock et al (63) found that CYP1A2, 2C19, 2E1 and 3A4 were at low activity levels in BA patients, which is consistent with the results of the present study. As for how they regulate each other in hepatic fibrosis and BA, this remains to be explored in the future.

A total of 2,086 DEmRNAs and 184 DElncRNAs were identified between liver tissues in BA and CC. DEmRNA and DElncRNA including GPC3, AFP, CYP1A2, CYP3A4, JPX and miR4435-2HG, might have an important role in the pathogenesis of liver fibrosis in BA. Furthermore, the hub genes of PPI network and the relation pairs in the lncRNA-target network were analyzed, including hub gene CXCL8, TLR7, TLR8, CCR5 and lncRNA-target pair H19 CYP2E1.

The mechanisms of fibrogenesis in BA remain to be elucidated, multiple pathways are simultaneous activated under cholestasis background. It is hoped that the RNA-sequencing results and bioinformation analysis in the present study could contribute some indications for description of this pathological process. There are some limitation to the present study; first, as in any statistical analysis, a reliable result depends on large-scale samples. The sample size in the present study was relatively small, due to the rare incidence of this progressive and severe neonatal fibro-obstructive cholangiopathy. Second, two age-unmatched patients with choledochal cyst under normal liver function test was as control instead of normal liver tissues due to ethical issues. It is hoped that further studies can be performed to validate the findings of the present study by enrolling more clinical samples and then the possible mechanisms can be clarified more accurately. Third, experiments on these candidate hub genes and pathways of the potential relation pair and network crosstalking to elucidate the molecular mechanism of BA are required.

Acknowledgements
Not applicable.

Funding
The present study was supported by the National Natural Science Foundation of China under (grant no. 81770512), Science, Technology and Innovation Commission of Shenzhen Municipality under (grant no. JCYJ20210324134202007) and Samming Project of Medicine in Shenzhen under Grant (grant no. SZSM201812055).

Availability of data and materials
The data generated and analyzed in the present study may be found in the Sequence Read Archive (SRA) under accession number (SUB7923751) at the following URL: https://www.ncbi.nlm.nih.gov/sra/PRJNA701623 and the GSE46960 dataset obtained from GEO at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46960. Other data used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions
YY and BW confirm the authenticity of all the raw data. TL and BW conceived and designed the present study. YY collected liver tissues and clinical data. YY and WeiW performed the experiments. WenW and WeiW analyzed the data. YY and WenW interpreted the data. WenW and WeiW drafted the manuscript. WenW and TL prepared Fig. 1-7 and Tables II and III. YY and WeiW prepared Table I. YY, TL, and BW reviewed and revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
The study was approved by Ethics Committee on Human Research of the Faculty of the Shenzhen Children's Hospital (approval no. 202106602). All tissue samples were collected with written informed consent from all participants or parents/guardians in the case of children under 18 and was conducted based on the principles expressed in the 1975 Declaration of Helsinki.

Patient consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.