Title

Biliary atresia is associated with polygenic susceptibility in ciliogenesis and planar polarity effector genes

Journal

Journal of Hepatology, 79(6)

ISSN

0168-8278

Authors

Glessner, Joseph T
Ningappa, Mylarappa B
Ngo, Kim A
et al.

Publication Date

2023-12-01

DOI

10.1016/j.jhep.2023.07.039

Peer reviewed

HHS Public Access

Author manuscript

J Hepatol. Author manuscript; available in PMC 2023 December 19.

Published in final edited form as:

J Hepatol. 2023 December ; 79(6): 1385–1395. doi:10.1016/j.jhep.2023.07.039.

Biliary atresia is associated with polygenic susceptibility in ciliogenesis and planar polarity effector genes

Joseph T. Glessner[1,†], **Mylarappa B. Ningappa**[2,†], **Kim A. Ngo**[3], **Maliha Zahid**[4], **Juhoon So**[4], **Brandon W. Higgs**[2], **Patrick M.A. Sleiman**[1], **Tejaswini Narayanan**[3], **Sarangarajan Ranganathan**[5], **Michael March**[1], **Krishna Prasadan**[6], **Courtney Vaccaro**[1], **Miguel Reyes-Mugica**[7], **Jeremy Velazquez**[8], **Claudia M. Salgado**[7], **Mo R. Ebrahimkhani**[8], **Lori Schmitt**[9], **Dhivyaa Rajasundaram**[10], **Morgan Paul**[2], **Renata Pellegrino**[1], **George K. Gittes**[11], **Dong Li**[1], **Xiang Wang**[1], **Jonathan Billings**[1], **Robert Squires**[12], **Chethan Ashokkumar**[2], **Khalid Sharif**[13], **Deirdre Kelly**[13], **Anil Dhawan**[14], **Simon Horslen**[12], **Cecilia W. Lo**[4], **Donghun Shin**[4], **Shankar Subramaniam**[3,15,*], **Hakon Hakonarson**[16,*], **Rakesh Sindhi**[2,*]

[1]Center for Applied Genomics (CAG), Children's Hospital of Philadelphia, Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[2]Hillman Center for Pediatric Transplantation, UPMC-Children's Hospital of Pittsburgh, and Thomas E Starzl Transplant Institute, University of Pittsburgh, Pittsburgh, PA, USA

[3]Department of Bioengineering, University of California, San Diego, San Diego, La Jolla, CA, USA

[4]Department of Developmental Biology, University of Pittsburgh, Pittsburgh, PA, USA

[5]Division of Pathology and Laboratory Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

[6]Rangos Research Center Animal Imaging Core, UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA, USA

[7]Division of Pediatric Pathology, UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA, USA

[*]Corresponding authors. Addresses: Director, Pediatric Transplant Research, University of Pittsburgh, Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA 15201, USA; Tel.: 412-692-7849; (R. Sindhi), or Director of the Center for Applied Genomics, The Joseph Stokes Jr. Research Institute of Children's Hospital of Philadelphia, 3401 Civic Center Blvd, Philadelphia, PA 19104, USA; Tel.: 267-426-0088; (H. Hakonarson), or University of California San Diego, California, Powell-Focht Bioengineering Hall, 9500 Gilman Drive, La Jolla, CA 92093, USA; Tel.: 858-822-0986; (S. Subramaniam). rakesh.sindhi@chp.edu (R. Sindhi), hakonarson@chop.edu (H. Hakonarson), shsubramaniam@ucsd.edu (S. Subramaniam).

[†]Contributed equally.

Authors' contributions

RS conceived the project, obtained funding, and wrote the manuscript with the support of HH, SS, and DS, co-PIs of this project. AD, DK and KS provided samples for genotyping, interpreted results, and also wrote the manuscript. JG performed all association studies and MBN supervised all sequencing studies and coordinated analysis of all samples with CA. Zebrafish, mouse tracheal ciliogenesis, human liver organoid, and human liver immunostaining studies were performed by JS, MZ and CL, JV and MRE, and MRM and CD, respectively. Supportive association analyses and sequencing was performed by PMAS, MM, CV, RP, DL, XW, JB (CAG investigators). All transcriptional analyses and integrative analyses were performed by BWH, KN, TN. All authors including SH and RS contributed to writing and approved this version of the manuscript.

Conflict of interest

M.R.E and J.V. have a patent (WO2019237124) for the organoid technology used in this publication.
Please refer to the accompanying ICMJE disclosure forms for further details.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2023.07.039.

[8]Department of Pathology, School of Medicine, Pittsburgh Liver Research Center, McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, PA, USA

[9]Histology Core Laboratory Manager, UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA, USA

[10]Department of Pediatrics, Division of Health Informatics, Children's Hospital of Pittsburgh, Pittsburgh, PA, USA

[11]Surgeon-in-Chief Emeritus, UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA, USA

[12]Pediatric Gastroenterology, UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA, USA

[13]Paediatric Liver Unit Including Intestinal Transplantation, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK

[14]Paediatric Liver GI and Nutrition Center and MowatLabs, NHS Foundation Trust, King's College Hospital, London, UK

[15]Department of Computer Science and Engineering, and Nanoengineering, University of California, San Diego, San Diego, La Jolla, CA, USA

[16]Divisions of Human Genetics and Pulmonary Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA

Abstract

Background & Aims: Biliary atresia (BA) is poorly understood and leads to liver transplantation (LT), with the requirement for and associated risks of lifelong immunosuppression, in most children. We performed a genome-wide association study (GWAS) to determine the genetic basis of BA.

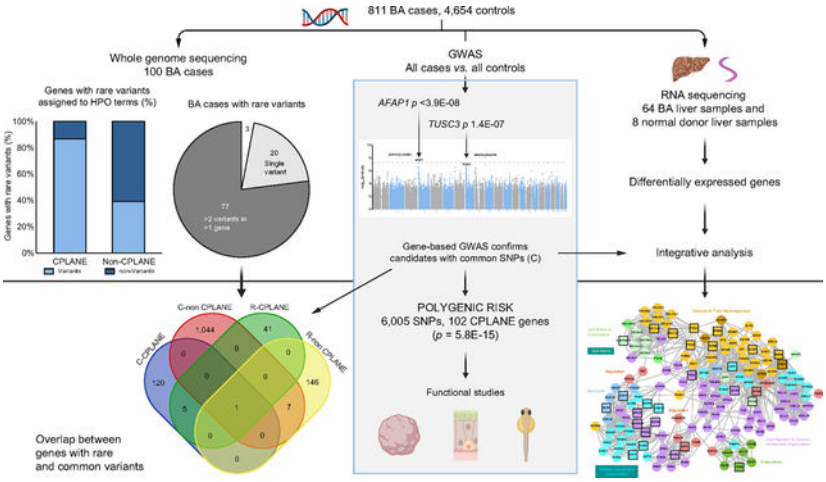
Methods: We performed a GWAS in 811 European BA cases treated with LT in US, Canadian and UK centers, and 4,654 genetically matched controls. Whole-genome sequencing of 100 cases evaluated synthetic association with rare variants. Functional studies included whole liver transcriptome analysis of 64 BA cases and perturbations in experimental models.

Results: A GWAS of common single nucleotide polymorphisms (SNPs), *i.e.* allele frequencies >1%, identified intronic SNPs rs6446628 in *AFAP1* with genome-wide significance (p = 3.93E-8) and rs34599046 in *TUSC3* at sub-threshold genome-wide significance (p = 1.34E-7), both supported by credible peaks of neighboring SNPs. Like other previously reported BA-associated genes, *AFAP1* and *TUSC3* are ciliogenesis and planar polarity effectors (CPLANE). In gene-set-based GWAS, BA was associated with 6,005 SNPs in 102 CPLANE genes (p = 5.84E-15). Compared with non-CPLANE genes, more CPLANE genes harbored rare variants (allele frequency <1%) that were assigned Human Phenotype Ontology terms related to hepatobiliary anomalies by predictive algorithms, 87% *vs.* 40%, p <0.0001. Rare variants were present in multiple genes distinct from those with BA-associated common variants in most BA cases. AFAP1 and TUSC3 knockdown blocked ciliogenesis in mouse tracheal cells. Inhibition of ciliogenesis caused biliary dysgenesis in zebrafish. AFAP1 and TUSC3 were expressed in fetal liver organoids, as well as fetal and BA livers, but not in normal or disease-control livers. Integrative analysis of

BA-associated variants and liver transcripts revealed abnormal vasculogenesis and epithelial tube formation, explaining portal vein anomalies that co-exist with BA.

Conclusions: BA is associated with polygenic susceptibility in CPLANE genes. Rare variants contribute to polygenic risk in vulnerable pathways via unique genes.

Graphical Abstract



Keywords

Biliary Atresia; Polygenic Susceptibility; Ciliogenesis; Portal Vein; Vascular Development; Tube Morphogenesis

Introduction

Biliary atresia (BA) affects 1 in 18,000 children of European ancestry and is the largest single indication for liver transplantation (LT) in children.[1,2] Post-transplant immunosuppression incurs a lifelong risk of life-threatening infections and malignancies. Therefore, understanding the pathogenesis of BA that requires LT is essential to develop strategies that prevent liver disease and its progression, and enhance surgical management. Genetic susceptibility and a developmental basis are important components of this pathogenesis. Unremitting jaundice due to loss of extrahepatic bile ducts is evident shortly after birth. Despite surgical drainage, three-fourths of all children develop cirrhosis and liver failure and require LT. Up to a fifth of all BA cases also manifest extrahepatic anomalies of other abdominal organs, the genitourinary or the cardiovascular systems.[3] Extrahepatic involvement also includes syndromic laterality defects in a tenth of BA cases. These defects range from asplenia to complete heterotaxy. Independent of extrahepatic involvement, up to half of all BA cases have portal vein anomalies which include poor portal venous hemodynamics and atresia or pre-duodenal location.[4–6] As a result, up to a fifth of BA cases may require portal venous reconstruction at the time of LT.[7] Patients with BA also experience a higher incidence of post-transplant portal venous thrombosis, compared with LT for other indications in children.[8] Thus, BA which progresses to LT could yield novel insights to advance our understanding of disease pathogenesis. To date, susceptibility loci

identified from genome-wide association studies (GWASs) show that exact replication can be challenging (Table S1).[9–25] Examples include the association, in cases of European ancestry, of BA with mannosidase alpha class 1A member 2 (*MAN1A2*), which has very few single nucleotide polymorphisms (SNPs) in eastern populations. An SNP downstream of X-prolyl aminopeptidase 1 (*XPNPEP1*) on 10q24.2 showed the strongest association with BA in Chinese BA cases.[16] Polymorphisms in this locus were subsequently found to affect the *ADD3* gene. Toward understanding genetic susceptibility to BA, we have performed a GWAS of the largest multicenter cohort of BA cases of self-reported European ancestry. We have also corroborated our findings using supportive analyses and functional studies.

Patients and methods

The study design is summarized in Fig. 1. Additional details are provided in the supplementary methods.

Cases

The diagnosis of BA in each participant was confirmed by review of pertinent diagnostic liver biopsy, radiologic and surgical reports. BA cases of self-reported European ancestry included 556 participants enrolled in either of two prospective observational cohort studies, A Prospective Database of Infants with Cholestasis [PROBE; NCT00061828] or Biliary Atresia Study in Infants and Children [BASIC; NCT00345553]) within the NIDDK-funded multi-center Childhood Liver Disease Research Network (ChiLDReN, http://childrennetwork.org). Additional DNA samples from BA cases included 152 from Children's Hospital of Pittsburgh (CHP) under Institutional Review Board approvals #19030279 (NCT#01163578), 63 from King's College Hospital, London, UK, under National Health Services REC ref 18/WA/0009, and 40 from Birmingham Children's Hospital (BCH), Birmingham, UK, under National Health Services, REC ref 19/LO/1103. CHP also participates in the ChiLDReN network. DNA samples from CHP were unique and not present among those provided from the ChiLDReN network. DNA was extracted at a core laboratory from blood samples collected under institution-specific IRB-approved protocols at participating centers in the US and Canada. RNA was extracted from snap frozen liver samples obtained at the time of transplant from 38 of 556 BA cases from ChiLDReN, snap frozen liver samples from 5 of 40 BA cases from BCH and from 21 liver samples collected in RNA-later in a subset of 152 BA cases from CHP and eight snap frozen normal liver samples under corresponding IRB-approved protocols. RNA was extracted as described.[10]

Controls

All 4,654 controls of self-reported European ancestry were recruited, and DNA was extracted as described, at the CAG-CHOP (Center for Applied Genomics of the Children's Hospital of Philadelphia) under IRB#:06–004886.[26]

Genotyping and data analysis

All genotyping was performed at CAG-CHOP. The Omni2.5–8v1–1 chip subversion was used for 695 or 85.7% cases, and 4,252 or 91.4% controls. The remainder were genotyped

with the Omni2.5–8v1–3 chip subversion. Genome-wide significance was set at $p < 5 \times 10–8$, and suggestive association was defined as an average of one false positive association per GWAS in European populations, or $p < 1 \times 10–5$. No deviation from the expected p values was observed in the Q-Q plot (Fig. 2). The genomic inflation factor λ was estimated as 1.02079. MiniMac3 as implemented in the Michigan Imputation Server was used for imputation of all SNPs and indel variants annotated in the 1000 Genomes Project Phase 3 v5 in BA cases and their respective controls.[27] Variants that were imputed with low confidence (R2 <0.4), or with minor allele frequency (MAF) <0.01 were removed from subsequent association testing. Testing for association with BA under an additive genetic effect model, adjusting for the top twenty principal components (PCs), as well as age at transplantation, gender and chip version, was performed using PLINK for 811 European cases *vs.* 4,654 European controls (Fig. 2, Table S2). Regional association plots were obtained with LocusZoom (Fig. 2).[28]

Gene-based association testing and polygenic risk

The p values of each marker within 20 kb upstream and downstream of a ciliogenesis and planar polarity effector (CPLANE) gene (Table S3–S4) were used for gene-based association analysis in the BA cohort using the fastBAT function in the GCTA (Genome-wide Complex Trait Analysis) tool version 1.92.3 beta3 Linux.[29,30] fastBAT performs a fast set-based association analysis for human complex traits using summary-level data from GWASs and linkage disequilibrium (LD) data from a reference sample with individual-level genotypes.[30] We also conducted gene-based association testing using VEGAS2: software for more flexible gene-based association testing (Table S5).[31]

Polygenic risk score

Polygenic risk score (PRS) values for all cases and controls were calculated using Bayesian Regression and Continuous Shrinkage Priors as implemented in PRS-CS, a Python tool that uses inputs GWAS summary statistics and an external LD reference panel to infer posterior SNP effect sizes under continuous shrinkage priors.[32] The training (base) GWAS dataset was from Chen Y *et al.*[13] and the test sample consists of independent samples in our current study.

Whole-genome sequencing

Whole-genome sequencing (WGS) of DNA from a test sub-cohort of 100 BA cases, all recruited at CHP and analysis of rare variants was performed using the GD Cross variant ranking and prioritization algorithm as described previously by us and others[33–35] (see supplementary methods). The input diagnosis for each BA case was biliary atresia. Variants that were assigned Human Phenotype Ontology (HPO) terms by the variant prioritization algorithm were evaluated further.

Functional studies

Gene and protein expression of AFAP1 and TUSC3 was evaluated in fetal liver organoids, immunohistochemistry of residual formalin-fixed paraffin-embedded liver tissue explanted from children with BA and other cholestatic diseases as disease controls. The effects of

AFAP1 and *TUSC3* knockdown on ciliogenesis were evaluated in mouse tracheal cells. Effects of disrupting ciliogenesis were evaluated with ciliobrevin D administration to zebrafish embryos (see supplementary methods).

Total RNA-sequencing and reverse-transcription quantitative PCR

RNA was extracted from 38 snap frozen BA liver samples (ChiLDReN), 5 snap frozen BA liver samples (BCH), 21 RNA-later preserved BA liver samples (CHP) and eight snap frozen normal liver samples (CHP). Total RNA-sequencing libraries were generated and sequenced using Illumina NovaSeq 6000 to get up to 40 M paired end reads at 100bp, as described in supplementary methods. Reverse-transcription quantitative PCR was performed using TaqMan gene expression assays targeting the *TUSC3* and *AFAP1* genes and the housekeeping gene GAPDH (see supplementary methods).

Integrative network analysis

An integrative analysis using protein-protein interactions (PPIs) was performed with STRING (version 11.0b, https://version-11-0b.string-db.org/) then visualized in Cytoscape (v3.8.2), a network visualization tool to identify clusters of genes that were well-connected within the functional enrichment results[36,37] A PPI network was based on significantly upregulated genes in the BA liver transcriptome, and included those that were among the 102 CPLANE genes[38] which conferred polygenic susceptibility, those in proximity to significant SNPs with p <0.0001 from GWAS results in Table S2, and the first neighbors of these genes (see supplementary methods).

Results

Samples and participants

The 811 cases included 321 or 40% males, while the 4,654 controls included 2,927 or 63% males. The median age at transplantation for cases (who are primarily pediatric) was 0.9 years (range 0.151–36.5 and standard deviation 4.17), while the median age of controls was 6 years (range 0–38 and standard deviation 5.78). The Omni2.5–8v1–1 chip subversion was used for genotyping 695 or 85.7% cases, and 4,252 or 91.4% controls. The remainder were genotyped with the Omni2.5–8v1–3 chip subversion.

GWAS implicates ciliary dysgenesis

This, the largest ever GWAS of BA, compared 811 cases with this rare disease, and 4,654 healthy controls, both of self-reported European ancestry (see supplementary methods). We selected 1,433,639 common SNPs with MAFs >1% for imputation after quality control. High confidence (R2 >0.4) imputed SNPs, with <0.1 genotype missingness, MAF >0.01, and Hardy-Weinberg equilibrium p >0.00005 numbering 9,261,934 were tested for association using logistic regression in a generalized linear model, adjusted by covariates consisting of 20 PCs, sex, age at transplantation, and chip subversion. PC analysis, using 1000 genomes as a reference for clustering, confirmed the European clustering consistent with the self-reported ethnicity. Among these cases *vs.* controls, the additive genetic effects model using the abovementioned covariates identified one BA-associated intronic SNP, rs6446628, chr4:7831935: A:C (hg19), MAF 0.348 *vs.* 0.285 in cases *vs.* controls with

genome-wide significance (p = 3.93E-8) and odds ratio = 1.4148 in the actin filament associated protein 1 (*AFAP1*) gene (Fig. 2, Table S2). We also found just below genome-wide significance, another BA-associated SNP rs34599046, position chr8:15544872:C:G (hg19), MAF 0.324 *vs.* 0.263 in cases *vs.* controls, p = 1.34E-7 and odds ratio = 1.40273, in the tumor suppressor candidate 3 (*TUSC3*) gene (Fig. 2, Table S2). Both SNPs were associated with credible peaks of SNPs in LD in corresponding genes. Gene-based association testing using VEGAS2v2 confirmed the ranking of *AFAP1* (p = 1E-06) and *TUSC3* (1E-05) as the first and second coding genes (Table S5).

AFAP1 and *TUSC3* are members of the recently described CPLANE network (Table S3), as are five previously described BA-associated genes, *GPC1, ARF6, ADD3, EFEMP1*, and *MAN1A2* (Table S1). Thus, we tested whether several CPLANE genes together conferred polygenic susceptibility to BA.

Polygenic susceptibility in CPLANE genes

We performed gene-set-based GWAS of self-reported cases and controls of European ancestry using all nominally significant (p <0.005) BA-associated SNPs in or 100 kb upstream and downstream of 2,436 known CPLANE genes (Table S3). A SNP-gene set of 6,005 SNPs in 102 genes (Table S4) was strongly associated with BA (p = 5.48E-15) confirming polygenic susceptibility. Liver and bile duct development is also facilitated by non-CPLANE genes, as evident from a suggestive association of rs75969446 in SOX17 (p = 9.4E-07, Table S2), and nominal association of the *SOX17* gene with BA in SNP-based and gene-based GWAS (p = 0.015, Table S5). Given extensive experimental support for this non-CPLANE gene in BA pathogenesis, which is further addressed in the discussion, we hypothesized and subsequently found significantly higher genome-wide PRSs in BA cases (p <2.2E-16) compared with corresponding controls (Table S6, Fig. 2).

GWAS SNPs can reveal common causal variants in LD or synthetic association with rare causal variants.[39] An efficient search for such common variants near 6,005 CPLANE SNPs or genome-wide SNPs would entail WGS of all 811 cases, greatly exceeding the scope of our ongoing study.

Rare variants associated with BA are enriched among CPLANE genes

To test for synthetic association with rare variants MAF <1%, we performed an exploratory genome-wide search with WGS of DNA from 100 BA cases, who were recruited at our center. After excluding false positive variant calls due to sequencing errors among the top-ranked variants of potential pathogenic contribution for each BA case, 517 rare variants were identified in 442 genes including 389 non-CPLANE and 53 CPLANE genes (Table S7). For the input diagnosis of biliary atresia, the GD Cross algorithm assigned 275 of these variants to seven HPO terms, biliary atresia, biliary tract abnormalities, abnormal biliary tract morphology, abnormality of the biliary system, abnormality of the abdominal organs, abnormality of the liver, and abnormalities of the digestive tract (HP: 0005912, HP:0001080, HP:0012440, HP:0004297, HP:0002012, HP:0001392, HP:0025031, respectively). These annotated variants were present in a larger fraction of CPLANE genes, 46 of 53 or 87%, compared with non-CPLANE genes, 153 of 389 or 40%, p <0.0001, Fisher's test (Fig. 3).

The distribution of rare variants was 1.4 per CPLANE and 1.36 per non-CPLANE genes (p = n.s.). The variant type, whether non-synonymous, splice region or splice site, frameshift or inframe, stop gain, or intronic, was also comparable in CPLANE and non-CPLANE genes.

CPLANE gene variants were present in 47 BA cases, of whom 14 harbored 2–3 variants in 1–3 genes, while the remaining harbored one variant in one gene. Noteworthy examples of CPLANE genes with ≥2 were *HTT, PLEC, DYSF, CC2D2A*, and *DNAH5*. Non-CPLANE gene variants were present in 91 of 100 BA cases, of whom 58 harbored 2–7 variants in 1–5 genes. Noteworthy examples of non-CPLANE genes with multiple variants were *DNAH9, GFI1, CDON, NEB*, and *GNA11*.

Rare variants are independently associated with BA

We looked for overlap between a) BA-associated genes which were significant at p values of ≤0.05 in gene-based GWAS of all 811 BA cases (Table S5), and b) 46 CPLANE genes and 153 non-CPLANE genes with rare variants that were associated with seven HPO terms as described above. We found limited overlap between genes with common and rare variants (Fig. 3). Five of 46 CPLANE genes with rare variants, *CNOT1*, *TSC2, BRIP1, PACS1* and *TUBB6* achieved nominal significance in gene-based GWAS with p values of 0.032, 0.018, 0.021, 0.037, and 0.009, respectively. Seven of 153 non-CPLANE genes, *PRKCG, TBL2, SERPINA1, LYST, COL5A2, SLC1A3*, and *RNF13* with rare variants achieved nominal significance in gene-based GWAS with p values of 0.024, 0.043, 0.026, 0.012, 0.038, 0.031 and 0.039, respectively. Overall, rare variants were present in 97 of 100 BA cases, 77 of whom had multiple rare variants. Seven BA cases harbored variants in CPLANE genes, 50 in non-CPLANE genes, and 40 cases had variants in CPLANE and non-CPLANE genes. Rare variants were not identified in three BA cases. While the prioritized BA-related rare variants uncovered with WGS do not explain the common GWAS signals through synthetic association, these rare variants contribute independently from the common variants to disease risk in BA, as previously described in other diseases.[40]

Ciliary dysgenesis in liver development

In a fetal liver organoid model derived from human inducible pluripotent stem cells, single-cell transcriptomics revealed expression of *AFAP1* and *TUSC3* in biliary epithelial cell (BEC) progenitors (Fig. 4).[41] Immunostaining of *AFAP1* in BECs was demonstrable in 18-week fetal livers and BA livers explanted at LT, but not in the normal postnatal liver or a disease-control liver with PFIC1 (primary familial cholestasis type 1) (Fig. 4). *TUSC3* demonstrated similar staining patterns. Consistent with abnormal liver development in BA, we found that Arl13b immunostaining of cholangiocyte cilia was present in fetal liver, normal liver and livers with PFIC, but not BA (Fig. 4). Small-interfering RNA-mediated inhibition of *Afap1* and *Tusc3* in mouse tracheal cell cultures blocked formation of motile cilia (Video S1–S3). Ciliobrevin D-induced chemical inhibition of ciliogenesis led to biliary dysgenesis and abnormal cardiac looping in zebrafish, further confirming the role of cilia in BA (Fig. 5 and Fig. S1). *AFAP1* and *TUSC3* are poorly expressed in zebrafish liver precluding reliable *in vivo* perturbation studies.

Supplementary video related to this article can be found at https://doi.org/10.1016/j.jhep.2023.07.039

Vascular dysgenesis in liver development

To identify potential pathogenic mechanisms, we first identified 5,313 differentially expressed genes (DEGs, false discovery rate <0.05, log2 fold-change ≥0.58) by RNA-sequencing of 64 cirrhotic BA livers removed at LT (in participants from our study) compared with eight healthy donor livers. To minimize the effect of the late disease stage in our cases, all of whom had developed cirrhosis, we compared these genes (Table S8) with those reported previously with microarray analysis of 64 BA liver biopsies obtained at the time of surgical drainage, when liver scarring was not advanced.[42] The 345 common upregulated genes were enriched (adjusted *p* value <0.01) for Gene Ontology Biological Processes (GOBP) terms such as extracellular structure organization, supramolecular fiber organization, vasculature development, and tube morphogenesis (Table S9). Other enriched processes were related to cilia function, *e.g.* cell motility, regulation of cellular component movement, or inflammatory cytokine-mediated signaling. Downregulated genes were enriched for metabolic processes and reflected loss of function with progressive liver disease. These enrichments were also mirrored in transcription factors associated with upregulated and downregulated genes.

To develop an overview of possible mechanisms, we constructed a PPI network using 1,389 DEGs in the BA liver transcriptome (Table S10). These DEGs corresponded with 102 BA-associated CPLANE genes and their first neighbors, and 1,337 genes in proximity to SNPs with *p* <0.0001 from GWAS and their first neighbors in the STRING database. We found enrichment of 17 GOBPs among the 954 upregulated genes (adjusted *p* value <0.01, Table S11). The top-ranked GOBPs were cell migration, cell-matrix interactions and circulatory system development. The regulatory small GTPase-mediated signaling pathway was also enriched. A PPI interaction network with 308 unique genes for these processes and filtered further by ≥6 interactions yielded a final network with 149 genes enriched for tube morphogenesis and circulatory system development (adjusted *p* value <0.05, Table S12–S15, and Fig. 6). This restricted gene set was also enriched for regulation of several developmental and signaling pathways (unadjusted *p* value <0.05), thereby corroborating and further extending the significance of the gene networks generated from the RNA-sequencing data.

Discussion

We report the novel association of BA requiring LT with the top-ranked *AFAP1* and second-ranked *TUSC3* genes in this largest ever GWAS of BA cases. These associations are based on an intronic lead SNP that achieves genome-wide significance in *AFAP1*, and a lead SNP just below genome-wide significance in *TUSC3* (Table S2). Both lead SNPs are accompanied by credible peaks of other SNPs in LD in the corresponding genes. Gene-based GWAS confirms this ranking among genes associated with BA in our study (Table S5). Inducible pluripotent stem cell-derived liver organoids and fetal liver immunostaining confirm that the corresponding proteins are present in developing BECs. Persistence of

corresponding proteins in explanted human livers with BA, but not other cholestatic diseases, suggests that bile ducts develop abnormally in BA, in part via disruption of ciliogenesis and ciliary function, as shown by knockdown of these genes in mouse tracheal epithelial cells.

AFAP1 modulates actin filament integrity, and has a previously reported association with glaucoma, but not liver disease.[43] The TUSC3 protein participates in cellular magnesium uptake, N-linked protein glycosylation, and embryonic development, and has been associated previously with congenital disorder of glycosylation Type 1n, but not BA.[44–46] Glycosylation disorders can affect the central nervous system, liver, heart, bones, and the immune system. AFAP1 and TUSC3 were upregulated 1.67-fold and 2.3-fold, respectively, by real-time qPCR in BA livers compared with normal livers. Both genes are members of the recently described CPLANE network comprising 2,436 unique genes (Table S3).[38] Variants in CPLANE genes have been associated with birth defects of single and multiple organs, including the ciliopathies. Remarkably, the CPLANE network also includes five BA susceptibility genes identified in previous GWAS: *GPC1, ARF6, ADD3, EFEMP1*, and *MAN1A2*. Cilia mediate two-way signaling between cells and the microenvironment by sensing ion flow and chemical gradients from the cell surface, and cell movement.

Consistent with previously described associations of several CPLANE genes with BA, our analysis strongly implicates polygenic risk due to common variants represented by a SNP-gene set of 6,005 common variants in 102 CPLANE genes (p = 5.48E-15) in the pathogenesis of BA (Table S4). This association places BA on the spectrum of several other congenital anomalies. Examples of sporadic birth defects with polygenic susceptibility include transposition of the great vessels or tetralogy of Fallot, where most cases are non-familial, and rare variants are only present in a minority of cases.[47,48] Polygenic susceptibility also provides a common mechanistic link between the aforementioned previously reported BA susceptibility genes. These genes participate in pathways such as hedgehog and epidermal growth factor signaling involved in branching morphogenesis during morphogenesis of organs like the liver.[9–14] Thus, increasing reports of abnormal ciliogenesis in BA are to be expected, as are reports describing rare variants in ciliary genes in some BA cases.[49–51] Using published genome-wide statistics from another GWAS study which included 215 BA cases from our current study, we have also found that genome-wide PRSs are higher in 596 remaining BA cases (p = 2.2E-16) compared with 4,654 controls. This finding indicates additional contributions from non-CPLANE genes to BA pathogenesis. In this regard, the suggestive or nominal association of the non-CPLANE *SOX17* gene in SNP- and gene-based GWASs, respectively, is noteworthy. Haploinsufficiency of *SOX17* leads to BA in mice, attesting to its role in regulating the development of the extrahepatic biliary tree from the common pancreaticobiliary duct.[52,53] The toxin biliatresone, which also leads to BA in animal models, does so by inducing SOX17 deficiency, a finding consistent with reduced SOX17 immunostaining in BECs from affected children.[54,55] Other BA-associated *SOX17* SNPs rs1692037, rs10958412, rs112233799 and rs7839526 in our GWAS further support this association and validate the extensive previously reported functional relevance of *SOX17* in BA pathogenesis. Given this large number of BA-associated variants in CPLANE and non-CPLANE genes, polygenic

risk must be specified comprehensively with a search for novel common or rare variants ‘tagged’ by disease-associated SNPs using WGS.

Our rare variant analysis also shows that common and rare variants converge on common pathways and mechanisms via unique genes to create disease susceptibility. In exploratory WGS of a subset of BA cases, CPLANE genes with rare variants were more likely to harbor variants assigned to seven BA-associated HPO terms by variant prioritization algorithms, compared with non-CPLANE genes, 87% *vs.* 40%, $p < 0.0001$. However, these rare variants were less likely to be found in genes with disease-associated common variants, suggesting that rare variants confer disease risk independently from the common variants. Among 1,176 nominally significant BA-associated genes ($p \leq 0.05$) identified by gene-based association testing (Table S5), rare variants were only present in 5 of 125 CPLANE and 7 of 1,051 non-CPLANE genes (Fig. 3). Second, at least one rare variant was present in 97 (nearly all) of 100 BA cases. The majority, or 77 of 97 cases harbored multiple rare variants in one or more CPLANE, non-CPLANE or both types of genes. This high prevalence of rare variants in a subset of BA cases in our study suggests that rare variants also confer multivariate risk by themselves, similar to and additively with polygenic risk due to common variants. In a final illustration of convergent pathways and mechanisms, each candidate gene list, whether based on common or rare variants or differential expression in the BA liver transcriptome (Tables S5, S7 or S8, respectively), included genes associated with diagnoses of isolated BA (ORPHA:30391) and syndromic BA (ORPHA:244283) in Orphanet. The numbers of genes associated with these two diagnoses were seven and one unique genes with common variants, respectively, two and four unique genes with rare variants, respectively, and 29 and 18 unique DEGs, respectively (summarized in Table S16). Thus, several unique genes with common and rare variants are likely to contribute to BA pathogenesis via common pathways. This inference is also supported by a recent analysis of 394,783 exomes from the UK biobank, which showed that common and rare variants contribute additively to disease heritability.[40]

The association of BA with vascular development revealed by integrative analysis has clinical relevance in explaining the association of BA with cardiovascular defects and portal venous anomalies, and the high incidence of portal vein thrombosis after LT for BA. This association also highlights the role of early liver development processes in BA.[56,57] During early liver development, portal endothelial and biliary epithelial networks develop interdependently, after the liver bud grows into the septum transversum, and acquires sinusoidal endothelial lakes with cells derived in part from the sinus venosus of the developing heart.[56–59] This interdependence is regulated by the polarity of intervening hepatocytes.[59] Planar polarity of cells also regulates BEC orientation and is disrupted in BA.[59,60] Thus, BA likely originates during early gestation, precluding routine *in utero* surgical intervention. Regulatory processes identified in integrative analysis are suitable targets for preventive intervention. Modeled after prevention of neural tube defects with folate administration during pregnancy, interventions that could prevent BA include those that reduce oxidative stress and stabilize regulatory pathways like hypoxia signaling which are pivotal for normal organogenesis during fetal development.[54,61]

We acknowledge the need for an independent cohort to replicate the observed genetic associations. The rarity of BA will require a multi-year international prospective collection to accrue a similar-sized independent cohort for this purpose. Another limitation is the need for experimental evidence to determine the additive effects of common and rare variants. To this end, we are planning experiments with multi-locus CRISPR editing of candidate regions with common and rare variants in F0 mouse generations used by one of us previously (CL).

BA is associated with polygenic susceptibility due to common variants in CPLANE genes. These defective genes promote ciliary dysgenesis during development but are less likely to contain potentially causal rare variants. A high prevalence of genome-wide rare variants with a preferential distribution in unique CPLANE genes is demonstrable in most BA cases in an exploratory sub-cohort analysis. These findings strongly suggest that common and rare variants confer additive risk via unique genes in common pathways in the pathogenesis of BA. Integrative analysis with the BA liver transcriptome indicates downstream disruption of vascular and epithelial tube morphogenesis, providing an explanation for portal vein atresia, which occurs in up to half of all cases with BA.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Financial support

This study was supported by award 5RO1DK109365-05 to RS, DS, and SS from the NIDDK branch of the National Institutes of Health, Bethesda, MD; Hillman Foundation of Pittsburgh (RS), and the Raizman-Haney and Szalay, Family endowments (RS); the Herridge, Jane't, Giventer-Braff, Smirch families; Marjory K. Harmer Endowment for Research in Pediatric Pathology (MRM); NIH project R01HL153407 (MZ); NIH projects EB024562 (MRE). CHiLDReN sample collections were supported by U01 grants from the NIDDK (DK 62445 [Mount Sinai School of Medicine], DK 62497 [Cincinnati Children's Hospital Medical Center], DK 62470 [Children's Healthcare of Atlanta], DK 62481 [The Children's Hospital of Philadelphia], DK 62456 [The University of Michigan], DK 84536 [Riley Hospital for Children], DK 84575 [Seattle Children's Hospital], DK 62500 [UCSF Children's Hospital], DK 62466 [Children's Hospital of Pittsburgh of University of Pittsburgh Medical Center (UPMC)], DK 62453 [Children's Hospital Colorado], DK 84538 [Children's Hospital Los Angeles], DK 62436 [Ann & Robert H Lurie Children's Hospital of Chicago], DK 103149 [Texas Children's Hospital], DK 103135 [The Hospital for Sick Children], and DK 103140 [University of Utah]). In addition, the project described was supported by the following NIH, NCATS, Clinical and Translational Sciences Award grants: University of Colorado UL1 TR002535, UCSF Children's Hospital UL1 TR001872, Children's Hospital of Pittsburgh of UPMC UL1 TR001857, The Children's Hospital of Philadelphia UL1 TR001878, Seattle Children's Hospital UL1 TR000423 and UL1 RR025014, Children's Healthcare of Atlanta UL1 TR002378, and Children's Hospital of Los Angeles UL1 TR00130. The ChiLDReN Data Coordinating Center and the University of Michigan and Arbor Collaborative for Research are supported by National Institutes of Health U01 DK062456.

Data availability statement

Raw data will be accessible on dbGap under study title Mapping Disease Pathways for Biliary Atresia, under accession number(s) that can be requested from RS, when this information becomes available. GWAS summary data are available from https://www.ebi.ac.uk/biostudies/ (S-BSST182) and https://www.ebi.ac.uk/gwas/ (GCST90296484). DNA samples from participants were isolated and banked at the Rutgers University NIDDK biorepository. All other relevant data and results are within the paper and its Supporting Information files. Controls: All data authorized for dbGaP submission have been deposited to dbGaP (accessions: phs000490.v1.p1, phs000607.v3.p2,

phs000371.v1.p1, phs000490.v1.p1, phs001194.v2.p2, phs001194.v2.p2.c1, phs001661, and phs000233).

Abbreviations

AFAP1	actin filament associated protein 1
BA	biliary atresia
BECs	biliary epithelial cells
ChiLDReN	Childhood Liver Disease Research Network
CHP	Children's Hospital of Pittsburgh
CPLANE	ciliogenesis and planar polarity effector
DEGs	differentially expressed genes
GOBP	Gene Ontology Biological Processes
GWAS	genome-wide association study
HPO	Human Phenotype Ontology
LD	linkage disequilibrium
LT	liver transplantation
MAF	minor allele frequency
MAN1A2	mannosidase alpha class 1A member 2
PC	principal component
PPI	protein-protein interaction
PRS	polygenic risk score
SNP	single nucleotide polymorphism
SOX17	SRY-box transcription factor 17
TUSC3	tumor suppressor candidate 3

References

Author names in bold designate shared co-first authorship

[1]. Otte JB, de Ville de Goyet J, Reding R, Hausleithner V, Sokal E, Chardot C, et al. Sequential treatment of biliary atresia with Kasai portoenterostomy and liver transplantation: a review. Hepatology 1994;20:41S–48S. [PubMed: 8005579]

[2]. Karrer FM, Price MR, Bensard DD, Sokol RJ, Narkewicz MR, Smith DJ, et al. Long-term results with the Kasai operation for biliary atresia. Arch Surg 1996;131:493–496. [PubMed: 8624194]

[3]. Schwarz KB, Haber BH, Rosenthal P, Mack CL, Moore J, Bove K, et al. Extrahepatic anomalies in infants with biliary atresia: results of a large prospective North American multicenter study. Hepatology 2013;58:1724–1731. [PubMed: 23703680]

[4]. Hernandez-Cano AM, Geis JR, Rumack CH, Stellin GP, Lilly JR. Portal vein dynamics in biliary atresia. J Pediatr Surg 1987;22:519–521. [PubMed: 3302199]

[5]. Wanek EA, Horgan JG, Karrer FM, Hall RJ, Rumack CM, Lilly JR. Portal venous velocity in biliary atresia. J Pediatr Surg 1990;25:146–148. [PubMed: 2405146]

[6]. Tanano H, Hasegawa T, Kawahara H, Sasaki T, Okada A. Biliary atresia associated with congenital structural anomalies. J Pediatr Surg 1999;34:1687–1690. [PubMed: 10591571]

[7]. Neto JS, Fonseca EA, Feier FH, Pugliese R, Candido HL, Benavides MR, et al. Analysis of factors associated with portal vein thrombosis in pediatric living donor liver transplant recipients. Liver Transpl 2014;20:1157–1167. [PubMed: 24954288]

[8]. Chardot C, Herrera JM, Debray D, Branchereau S, De Dreuzy O, Devictor D, et al. Portal vein complications after liver transplantation for biliary atresia. Liver Transpl Surg 1997;3:351–358. [PubMed: 9346763]

[9]. Cui S, Leyva-Vega M, Tsai EA, EauClaire SF, Glessner JT, Hakonarson H, et al. Evidence from human and zebrafish that GPC1 is a biliary atresia susceptibility gene. Gastroenterology 2013;144:1107–1115 e1103. [PubMed: 23336978]

[10]. Ningappa M, So J, Glessner J, Ashokkumar C, Ranganathan S, Min J, et al. The role of ARF6 in biliary atresia. PLoS One 2015;10:e0138381. [PubMed: 26379158]

[11]. So J, Ningappa M, Glessner J, Min J, Ashokkumar C, Ranganathan S, et al. Biliary-atresia-associated mannosidase-1-alpha-2 gene regulates biliary and ciliary morphogenesis and laterality. Front Physiol 2020;11:538701. [PubMed: 33192543]

[12]. Cheng G, Tang CS, Wong EH, Cheng WW, So MT, Miao X, et al. Common genetic variants regulating ADD3 gene expression alter biliary atresia risk. J Hepatol 2013;59:1285–1291. [PubMed: 23872602]

[13]. Chen Y, Gilbert MA, Grochowski CM, McEldrew D, Llewellyn J, Waisbourd-Zinman O, et al. A genome-wide association study identifies a susceptibility locus for biliary atresia on 2p16.1 within the gene EFEMP1. PLoS Genet 2018;14:e1007532. [PubMed: 30102696]

[14]. Tsai EA, Grochowski CM, Loomes KM, Bessho K, Hakonarson H, Bezerra JA, et al. Replication of a GWAS signal in a Caucasian population implicates ADD3 in susceptibility to biliary atresia. Hum Genet 2014;133:235–243. [PubMed: 24104524]

[15]. Zhao R, Song Z, Dong R, Li H, Shen C, Zheng S. PolymorphismofITGB2gene3'-UTR+145C/A is associated with biliary atresia. Digestion 2013;88:65–71. [PubMed: 23921155]

[16]. Garcia-Barcelo MM, Yeung MY, Miao XP, Tang CS, Cheng G, So MT, et al. Genome-wide association study identifies a susceptibility locus for biliary atresia on 10q24.2. Hum Mol Genet 2010;19:2917–2925. [PubMed: 20460270]

[17]. Zeng S, Sun P, Chen Z, Mao J, Wang J, Wang B, et al. Association between single nucleotide polymorphisms in the ADD3 gene and susceptibility to biliary atresia. PLoS One 2014;9:e107977. [PubMed: 25285724]

[18]. Tang V, Cofer ZC, Cui S, Sapp V, Loomes KM, Matthews RP. Loss of a candidate biliary atresia susceptibility gene, add3a, causes biliary developmental defects in zebrafish. J Pediatr Gastroenterol Nutr 2016;63:524–530. [PubMed: 27526058]

[19]. Udomsinprasert W, Tencomnao T, Honsawek S, Anomasiri W, Vejchapipat P, Chongsrisawat V, et al. +276 G/T single nucleotide polymorphism of the adiponectin gene is associated with the susceptibility to biliary atresia. World J Pediatr 2012;8:328–334.

[20]. Camaj P, Seeliger H, Ischenko I, Krebs S, Blum H, De Toni EN, et al. EFEMP1 binds the EGF receptor and activates MAPK and Akt pathways in pancreatic carcinoma cells. Biol Chem 2009;390:1293–1302. [PubMed: 19804359]

[21]. Lee HC, Chang TY, Yeung CY, Chan WT, Jiang CB, Chen WF, et al. Association of interferon-gamma gene polymorphisms in Taiwanese children with biliary atresia. J Clin Immunol 2010;30:68–73. [PubMed: 19756986]

[22]. Arikan C, Berdeli A, Ozgenc F, Tumgor G, Yagci RV, Aydogdu S. Positive association of macrophage migration inhibitory factor gene-173G/C polymorphism with biliary atresia. J Pediatr Gastroenterol Nutr 2006;42:77–82. [PubMed: 16385258]

[23]. Lee HC, Chang TY, Yeung CY, Chan WT, Jiang CB, Chen WF, et al. Genetic variation in the vascular endothelial growth factor gene is associated with biliary atresia. J Clin Gastroenterol 2010;44:135–139. [PubMed: 19713864]

[24]. Liu B, Wei J, Li M, Jiang J, Zhang H, Yang L, et al. Association of common genetic variants in VEGFA with biliary atresia susceptibility in Northwestern Han Chinese. Gene 2017;628:87–92. [PubMed: 28710035]

[25]. Kaewkiattiyot S, Honsawek S, Vejchapipat P, Chongsrisawat V, Poovorawan Y. Association of X-prolyl aminopeptidase 1 rs17095355 polymorphism with biliary atresia in Thai children. Hepatol Res 2011;41:1249–1252. [PubMed: 22118303]

[26]. Sindhi R, Higgs BW, Weeks DE, Ashokkumar C, Jaffe R, Kim C, et al. Genetic variants in major histocompatibility complex-linked genes associate with pediatric liver transplant rejection. Gastroenterology 2008;135:830–839. 839 e831–810.

[27]. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284–1287. [PubMed: 27571263]

[28]. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336–2337. [PubMed: 20634204]

[29]. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82. [PubMed: 21167468]

[30]. Bakshi A, Zhu Z, Vinkhuyzen AA, Hill WD, McRae AF, Visscher PM, et al. Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. Sci Rep 2016;6:32894. [PubMed: 27604177]

[31]. Mishra A, Macgregor S. VEGAS2: software for more flexible gene-based testing. Twin Res Hum Genet 2015;18:86–91. [PubMed: 25518859]

[32]. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun 2019;10:1776. [PubMed: 30992449]

[33]. Strong A, Simone L, Krentz A, Vaccaro C, Watson D, Ron H, et al. Expanding the genetic landscape of oral-facial-digital syndrome with two novel genes. Am J Med Genet A 2021;185:2409–2416. [PubMed: 34132027]

[34]. Ebiki M, Okazaki T, Kai M, Adachi K, Nanba E. Comparison of causative variant prioritization tools using next-generation sequencing data in Japanese patients with mendelian disorders. Yonago Acta Med 2019;62:244–252. [PubMed: 31582890]

[35]. Tosco-Herrera E, Munoz-Barrera A, Jaspez D, Rubio-Rodriguez LA, Mendoza-Alvarez A, Rodriguez-Perez H, et al. Evaluation of a whole-exome sequencing pipeline and benchmarking of causal germline variant prioritizers. Hum Mutat 2022;43:2010–2020. [PubMed: 36054330]

[36]. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019;47:D607–D613. [PubMed: 30476243]

[37]. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–2504. [PubMed: 14597658]

[38]. Toriyama M, Lee C, Taylor SP, Duran I, Cohn DH, Bruel AL, et al. The ciliopathy-associated CPLANE proteins direct basal body recruitment of intraflagellar transport machinery. Nat Genet 2016;48:648–656. [PubMed: 27158779]

[39]. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol 2010;8:e1000294. [PubMed: 20126254]

[40]. Weiner DJ, Nadig A, Jagadeesh KA, Dey KK, Neale BM, Robinson EB, et al. Polygenic architecture of rare coding variation across 394,783 exomes. Nature 2023 Feb;614(7948):492–499. 10.1038/s41586-022-05684-z. Epub 2023 Feb 8. [PubMed: 36755099]

[41]. Velazquez JJ, LeGraw R, Moghadam F, Tan Y, Kilbourne J, Maggiore JC, et al. Gene regulatoy network analysis and engineering directs development and vascularization of multilineage human liver organoids. Cell Syst 2021;12:41–55 e11. [PubMed: 33290741]

[42]. 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;60:211–223. [PubMed: 24493287]

[43]. Shiga Y, Akiyama M, Nishiguchi KM, Sato K, Shimozawa N, Takahashi A, et al. Genome-wide association study identifies seven novel susceptibility loci for primary open-angle glaucoma. Hum Mol Genet 2018;27:1486–1496. [PubMed: 29452408]

[44]. Zhou H, Clapham DE. Mammalian MagT1 and TUSC3 are required for cellular magnesium uptake and vertebrate embryonic development. Proc Natl Acad Sci U S A 2009;106:15750–15755. [PubMed: 19717468]

[45]. Mohorko E, Owen RL, Malojcic G, Brozzo MS, Aebi M, Glockshuber R. Structural basis of substrate specificity of human oligosaccharyl transferase subunit N33/Tusc3 and its role in regulating protein N-glycosylation. Structure 2014;22:590–601. [PubMed: 24685145]

[46]. Chang IJ, He M, Lam CT. Congenital disorders of glycosylation. Ann Transl Med 2018;6:477. [PubMed: 30740408]

[47]. Skoric-Milosavljevic D, Tadros R, Bosada FM, Tessadori F, van Weerd JH, Woudstra OI, et al. Common genetic variants contribute to risk of transposition of the great arteries. Circ Res 2022;130:166–180. [PubMed: 34886679]

[48]. Grunert M, Dorn C, Schueler M, Dunkel I, Schlesinger J, Mebus S, et al. Rare and private variations in neural crest, apoptosis and sarcomere genes define the polygenic background of isolated Tetralogy of Fallot. Hum Mol Genet 2014;23:3115–3128. [PubMed: 24459294]

[49]. Chu AS, Russo PA, Wells RG. Cholangiocyte cilia are abnormal in syndromic and non-syndromic biliary atresia. Mod Pathol 2012;25:751–757. [PubMed: 22301700]

[50]. 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. [PubMed: 34392560]

[51]. Lam WY, Tang CS, So MT, Yue H, Hsu JS, Chung PH, et al. Identification of a wide spectrum of ciliary gene mutations in nonsyndromic biliary atresia patients implicates ciliary dysfunction as a novel disease mechanism. EBioMedicine 2021;71:103530. [PubMed: 34455394]

[52]. Uemura M, Ozawa A, Nagata T, Kurasawa K, Tsunekawa N, Nobuhisa I, et al. Sox17 haploinsufficiency results in perinatal biliary atresia and hepatitis in C57BL/6 background mice. Development 2013;140:639–648. [PubMed: 23293295]

[53]. Spence JR, Lange AW, Lin SC, Kaestner KH, Lowy AM, Kim I, et al. Sox17 regulates organ lineage segregation of ventral foregut progenitor cells. Dev Cell 2009;17:62–74. [PubMed: 19619492]

[54]. Fried S, Gilboa D, Har-Zahav A, Lavrut PM, Du Y, Karjoo S, et al. Extrahepatic cholangiocyte obstruction is mediated by decreased glutathione, Wnt and Notch signaling pathways in a toxic model of biliary atresia. Sci Rep 2020;10:7599. [PubMed: 32371929]

[55]. El-Goday SF, El-Azab DS, Kandil MA, Aiad HA, Sira MM, Tantawy MS. High hairy and enhancer of split 1 and low SOX17 immunohistochemical hepatic expression is associated with human biliary atresia. Egypt J Pathol 2019;39:18–25. 2019.

[56]. El-Goday SF, El-Azab DS, Kandil MA, Aiad HA, Sira MM, Tantawy MS. High hairy and enhancer of split 1 and low SOX17 immunohistochemical hepatic expression is associated with human biliary atresia. Egypt J Pathol 2019;39:18–25. 2019.

[57]. Collardeau-Frachon S, Scoazec JY. Vascular development and differentiation during human liver organogenesis. Anat Rec (Hoboken) 2008;291:614–627. [PubMed: 18484606]

[58]. Ober EA, Lemaigre FP. Development of the liver: insights into organ and tissue morphogenesis. J Hepatol 2018;68:1049–1062. [PubMed: 29339113]

[59]. Sakaguchi TF, Sadler KC, Crosnier C, Stainier DY. Endothelial signals modulate hepatocyte apicobasal polarization in zebrafish. Curr Biol 2008;18:1565–1571. [PubMed: 18951027]

[60]. Cui S, Capecci LM, Matthews RP. Disruption of planar cell polarity activity leads to developmental biliary defects. Dev Biol 2011;351:229–241. [PubMed: 21215262]

[61]. Min J, Ningappa M, So J, Shin D, Sindhi R, Subramaniam S. Systems analysis of biliary atresia through integration of high-throughput biological data. Front Physiol 2020;11:966. [PubMed: 32848883]

Impact and implications

Liver transplantation is needed to cure most children born with biliary atresia, a poorly understood rare disease. Transplant immunosuppression increases the likelihood of life-threatening infections and cancers. To improve care by preventing this disease and its progression to transplantation, we examined its genetic basis. We find that this disease is associated with both common and rare mutations in highly specialized genes which maintain normal communication and movement of cells, and their organization into bile ducts and blood vessels during early development of the human embryo. Because defects in these genes also cause other birth defects, our findings could lead to preventive strategies to lower the incidence of biliary atresia and potentially other birth defects.

Highlights

- GWAS of common SNPs in 811 BA cases that required liver transplantation and 4,654 controls.
- BA is strongly associated with polygenic susceptibility in 102 ciliogenesis and planar polarity effector genes.
- Functional data implicate ciliary dysgenesis, abnormal biliary epithelial cell development and vasculogenesis.
- Rare variants enriched for relevant Human Phenotype Ontology terms occur in several unique genes in most BA cases.

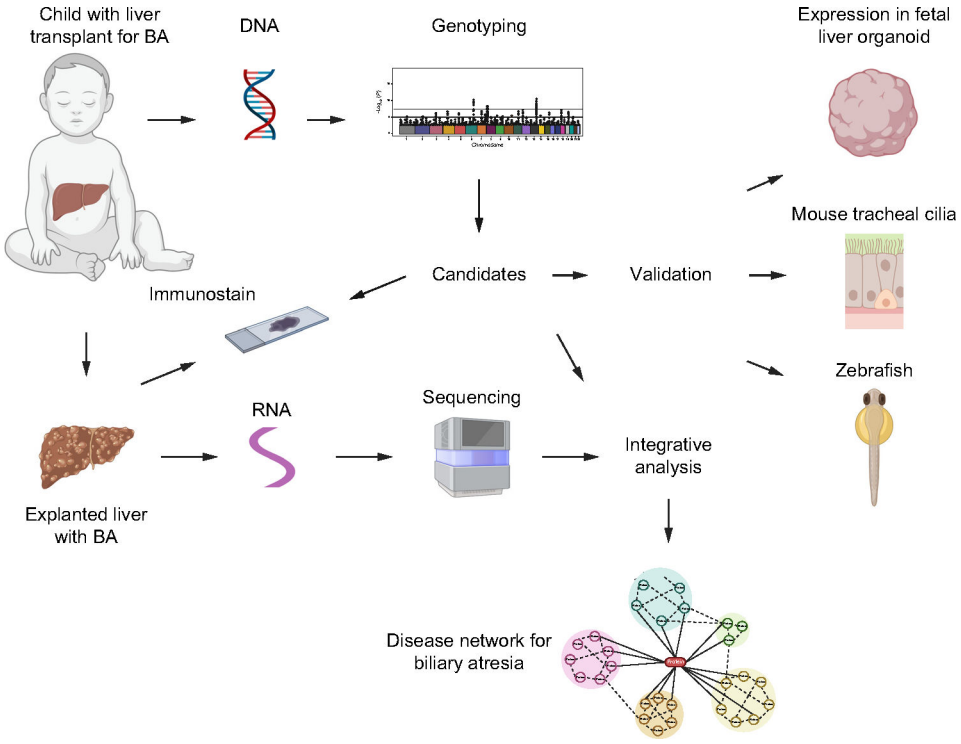


Fig. 1. Study design.

BA-associated common SNPs were identified using GWAS in 811 BA cases and 4,654 controls. Candidate genes underwent functional validation with immunostaining, fetal liver organoids, mouse tracheal cilia, and zebrafish. Whole-transcriptome sequencing was performed on 64 BA livers and eight normal donor liver samples. An integrative analysis was performed using genes identified by GWAS and differentially expressed genes in BA livers. BA, biliary atresia, GWAS, genome-wide association study; SNPs, single nucleotide polymorphisms.

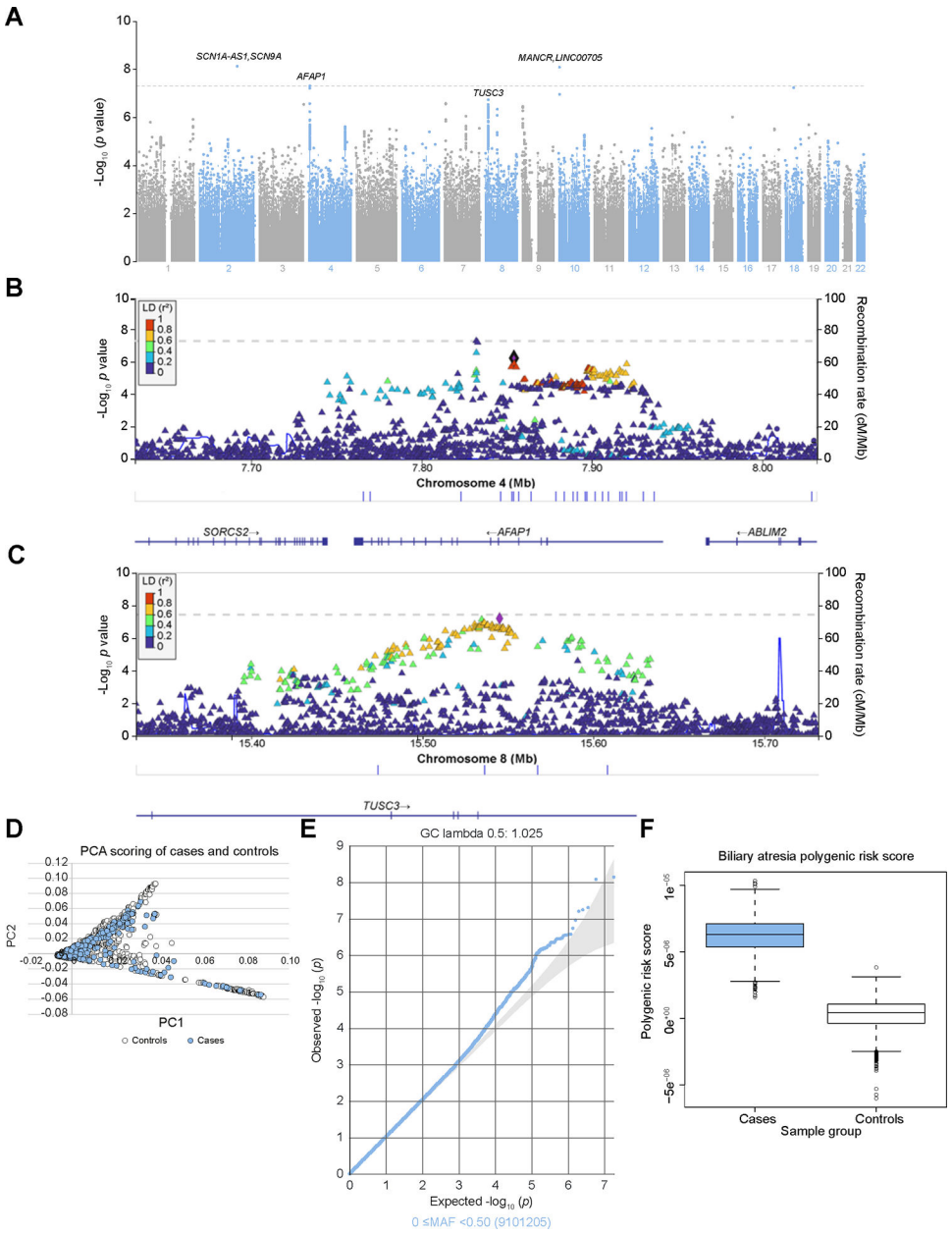


Fig. 2. Genome-wide association study.

(A-C) Genome-wide (A) and regional Manhattan plots (B and C) indicate robust genome-wide association signals for the *AFAP1* (B) and *TUSC3* (C) loci. Each locus shows a lead SNP (purple diamond) coupled with elevated significance of SNPs in linkage disequilibrium and flanking SNPs with low significance (dark blue dots). (D) PCA plot for cases (blue) and controls (orange). (E) Q-Q plot for association results. (F) Polygenic risk score distribution in cases and controls. PCA, principal component analysis; SNPs, single nucleotide polymorphisms.

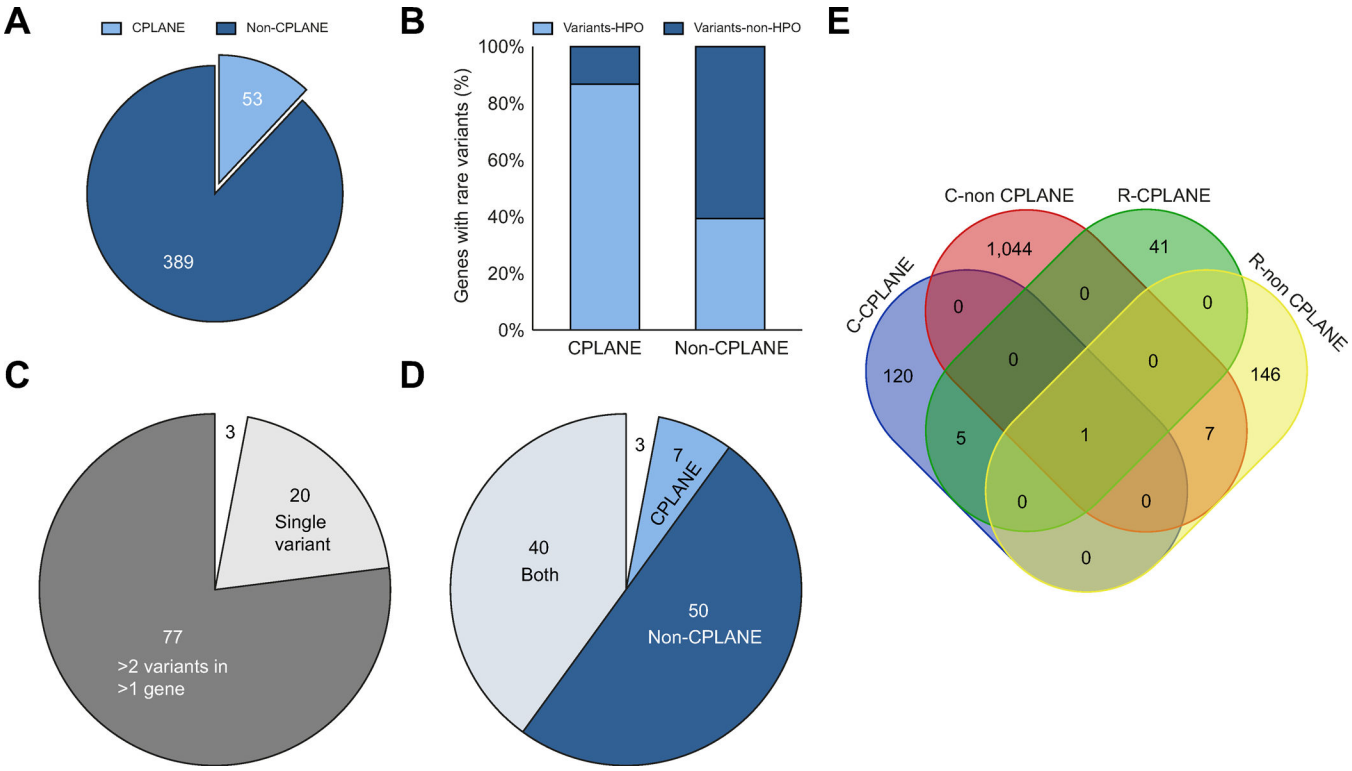


Fig. 3. Rare variants in BA.

(A) CPLANE and non-CPLANE genes with rare variants. (B) Proportions of CPLANE and non-CPLANE genes in Fig 2A, with rare variants which are associated with HPO terms. (C) BA cases with no rare variants (n = 3), single rare variants (light grey) and multiple rare variants (dark grey). (D) Distribution of CPLANE and non-CPLANE genes, and genes in both categories with rare variants among BA cases. (E) Venn diagram shows overlap between CPLANE and non-CPLANE genes with rare and common variants. BA, biliary atresia; CPLANE, ciliogenesis and planar polarity effector; HPO, Human Phenotype Ontology.

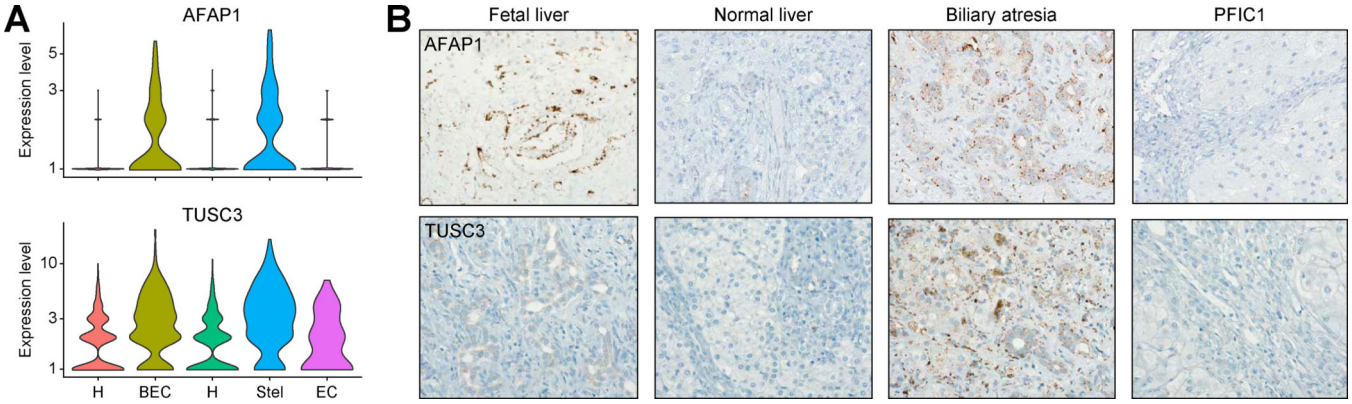


Fig. 4. Single-cell RNA-sequencing and immunostaining.

(A) Single-cell RNA-sequencing data from inducible pluripotent stem cell-derived human fetal liver organoids. Spindle plots show expression of *AFAP1* (upper panel) and *TUSC3* (lower panel) in progenitors of hepatocytes, biliary epithelial cells, stellate cells and endothelial cells (H, BEC, Stel, EC, respectively). Expression of both genes is enriched in BEC and Stel populations. (B) Immunoperoxidase staining shows expression of AFAP1 (upper panels) and TUSC3 (lower panels) in 18-week-old human fetal liver, normal human liver, and liver samples from individuals with biliary atresia and disease controls (primary familial intrahepatic cholestasis type 1, PFIC1). AFAP1 staining is predominantly seen in cholangiocytes. TUSC3 staining is more diffuse, in cholangiocytes, endothelial cells and other cells.

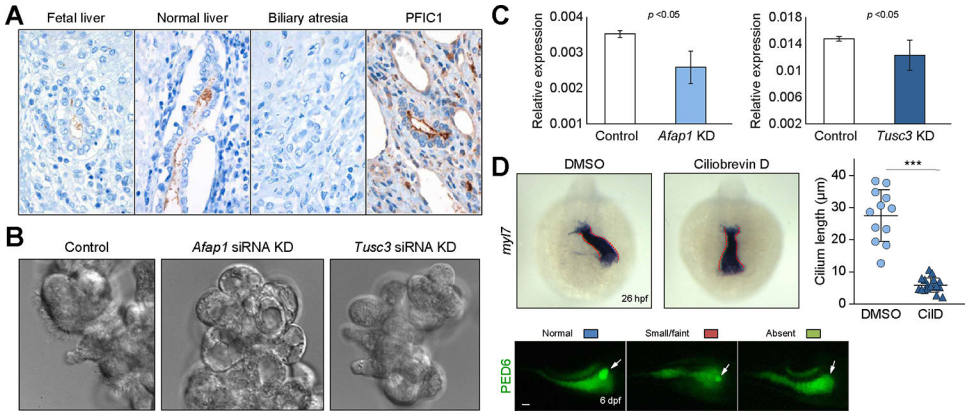


Fig. 5. Functional studies.

(A) Immunoperoxidase staining of Arl13b, which is expressed in primary or motile cilia shows luminal staining in 18-week-old human fetal liver, normal human liver, and liver disease-control (primary familial intrahepatic cholestasis type 1, PFIC1). Arl13b staining is absent in livers with BA. (B) Panels from left to right show mouse tracheal cell cultures before (left) and after treatment with *Afap1* siRNA and *Tusc3* siRNA. Treated cultures do not demonstrate the normal distribution of motile cilia seen in untreated cultures (Video S1). (C) The relative expression of *Afap1* and *Tusc3* in mouse tracheal epithelia cell cultures shows the effectiveness of KD after siRNA transfection (Video S2 and Video S3). (D) CilD treatment impairs left-right asymmetry and hepatic biliary function in zebrafish. Upper left. Whole-mount *in situ* hybridization images showing *myl7* expression in cardiomyocytes at 26 hpf. The embryos were treated with 5 μM CilD from 8 to 26 hpf. Dotted lines outline the heart. Two patterns of heart looping (normal and midline) were observed. Lower left panels. Epifluorescence images showing PED-6 accumulation in the gallbladder (arrowheads). The larvae were treated with 2.5 μM CilD from 3 to 6 dpf. Based on PED6 levels in the gallbladder, larvae were divided into three groups: normal, small/faint, and absent. Right panel. The length of cilium in cholangiocytes was significantly reduced in CilD-treated larvae (n = 6) compared to controls (n = 4) at 6 dpf. CilD, ciliobrevin D; Dpf, days post fertilization; Hpf, hours post fertilization; KD, knockdown; siRNA, small-interfering RNA.

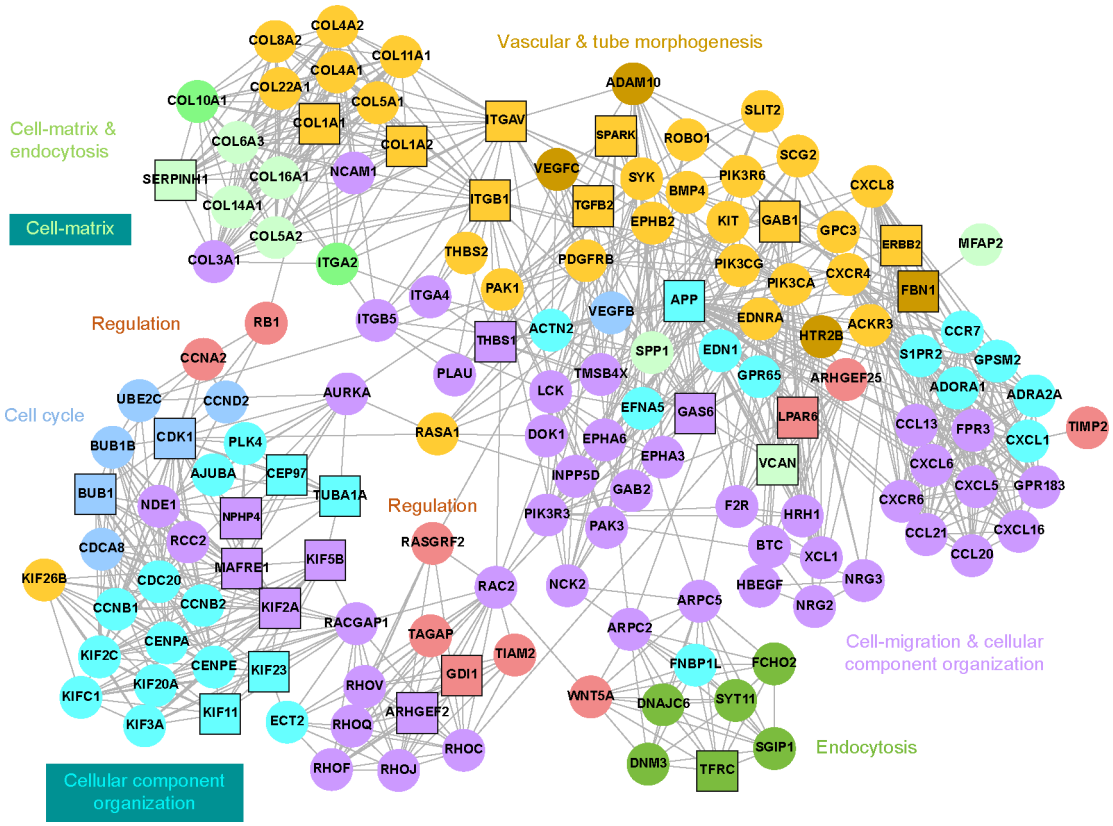


Fig. 6. Protein-protein interaction network of 149 highly connected nodes with ≥6 interactions identified in integrative analysis of RNA-sequencing and GWAS genes.

Target genes were upregulated in the BA transcriptome and selected from among 1,389 genes including 102 BA-associated CPLANE genes and genes in proximity to SNPs with significant differences between 811 BA cases and 4,654 controls (p <0.0001) and the first neighbors of these gene sets in the STRING database. These genes were present in significant gene ontology biological processes (adjusted p value <0.01). Node colors indicate the biological processes associated with the genes. Rectangular outlines identify CPLANE genes. Details for 149 network genes and associated pathways are provided in Tables S12–S15. BA, biliary atresia; GWAS, genome-wide association study. BA, biliary atresia, CPLANE, ciliogenesis and planar polarity effector; GWAS, genome-wide association study; SNPs, single nucleotide polymorphisms.