List of Supplementary Materials

Supplementary Material and Methods

Figure S1. Receiver Operating Characteristic curves of subgroup analyses in subjects with serum GGT 100-300 U/L and greater 300 U/L

Figure S2. Relationship between MMP-7 and hepatic fibrosis at the time of diagnosis of BA

Figure S3. Hepatic mRNA expression of *GGT* in biliary atresia (BA), intrahepatic cholestasis (IHC), and normal controls and of *Ggt1* in the experimental BA.

Table S1. List of 76 serum proteins differentially expressed biliary atresia (BA) and intrahepatic cholestasis (IHC)

Table S2. Sensitivity of MMP-7 to predict biliary atresia from the validation cohort #2

Table S3. Primer sequences and PCR product sizes for mouse *Mmp-7, Ggt1, Th1, TNF-α and IL8* related genes

Supplementary file. Protein results from SOMAScan®

References

1. C. Liu, B. J. Aronow, A. G. Jegga, N. Wang, A. Miethke, R. Mourya, J. A. Bezerra, Novel resequencing chip customized to diagnose mutations in patients with inherited syndromes of intrahepatic cholestasis. *Gastroenterology* **132**, 119-126 (2007).
2. Y. Yang, D. M. Muzny, J. G. Reid, M. N. Bainbridge, A. Willis, P. A. Ward, A. Braxton, J. Beuten, F. Xia, Z. Niu, M. Hardison, R. Person, M. R. Bekheirnia, M. S. Leduc, A. Kirby, P. Pham, J. Scull, M. Wang, Y. Ding, S. E. Plon, J. R. Lupski, A. L. Beaudet, R.

SUPPLEMENTARY MATERIAL

Material and Methods

Tissue Histology

Histological analyses of tissues stained with hematoxylin/eosin were performed by an investigator unaware of the experimental or control origin of samples. Bile ducts were classified as obstructed or patent, while liver histology was scored according to portal inflammation and area of lobular necrosis as follows: 1) mild portal inflammation without area of necrosis = mild, 2) mild to moderate portal inflammation with patchy necrosis = moderate, and 3) moderate to severe portal inflammation with large area of necrosis = severe.

Singleplex Bead-based assay and PCR

MMP-7 concentrations in the sample supernatants were determined by enzyme-linked immunosorbent assay (ELISA) using Milliplex^{TM} Multiplex kits (Millipore, Billerica, MA, USA) according to manufacturer's protocol, and described in Supplementary Methods. Briefly, 25μL sample in duplicate was incubated with 25μL antibody coated beads in a 96 well multiscreen filter plate overnight at 4°C on a plate shaker. Plates were then washed twice on a vacuum apparatus, followed by the addition of 25μL of secondary antibody and incubation at room temperature for 1 hour with gentle shaking. To the plate, 25μL of streptavidin-RPE were added and incubated for 30 minutes, then washed and read using the Bio-Plex^{TM} (Bio-Rad, Hercules, CA, USA). Concentrations were calculated from standard curves using recombinant proteins and expressed in ng/mL.

Quantitative PCR

Total RNA was isolated from the livers and extrahepatic bile ducts of RRV- or saline-injected BALB/c mice with or without Batimastat- or vehicle- injection using the RNeasy Mini Kit, according to manufacturer's protocol (Qiagen Inc); RNA integrity was verified by agarose gel electrophoresis as described previously.(*29*) All samples were then reverse transcribed to cDNA using the High Capacity RNA-to-cDNA Kit (Applied Biosystems, Waltham, MA USA) and used for real-time PCR on a Stratagene Mx3000P thermocycler (Agilent Technologies Inc.) to detect relative concentrations of individual mRNAs. PCR amplifications were performed with specific primers of *Mmp-7, Ggt1, Cxcl1, Cxcl2, Cxcl5, Cxcl9, Cxcl10, Cxcl11, Tnfa, Ifng, and IL12p40* (table S3). mRNA expression of target genes was normalized to the endogenous reference *Gapdh* gene.

Immunohistochemistry and automated image analysis

Immunohistochemistry was performed on 5 µm paraffin-embedded tissue sections as described previously (*43*), using 1:50 polyclonal MMP-7 antibody (Cloud-Clone Corp, Houston, TX, USA) in blocking serum overnight at 4°C. Biotinylated anti-rabbit antibody (Vector Laboratories, Burlingame, CA, USA) was used as secondary antibody followed by Avidin/Biotinylated-detection (VECTASTAIN ABC reagent PK-4001, Vector Laboratories) and the 3,3’-diaminobenzidine (DAB) substrate (Vector Kit, SK-4100, Vector Laboratories). The same protocol was used in human and mouse tissue. An additional liver biopsy slide from each BA subject was subjected for Masson’s trichrome staining. For image analyses, entire sections were scanned and subdivided to depict the portal tracts using Olympus BX51 microscope (Olympus America Inc., PA, USA) and CellSens Dimension digital imaging software (Olympus Corporation, version 1.8.1). Automated quantification of MMP-7 expression was performed on scanned images with NIS Nikon Elements software using red-green-blue (RGB) general analysis. Positive expression was defined choosing DAB-positive brown area (DAB+). Portal tracts were selected and enumerated as regions of interest (ROI). RGB general analysis was used in hue, saturation, and intensity modes to identify DAB+ areas. All subdivided regions were averaged to report surface area of DAB+ per portal tract in each sample. For Masson’s trichrome staining, positive collagen was defined by blue areas with use of hue, saturation and intensity modes. The quantification was reported as surface area of collagen (blue stained)/portal tract.

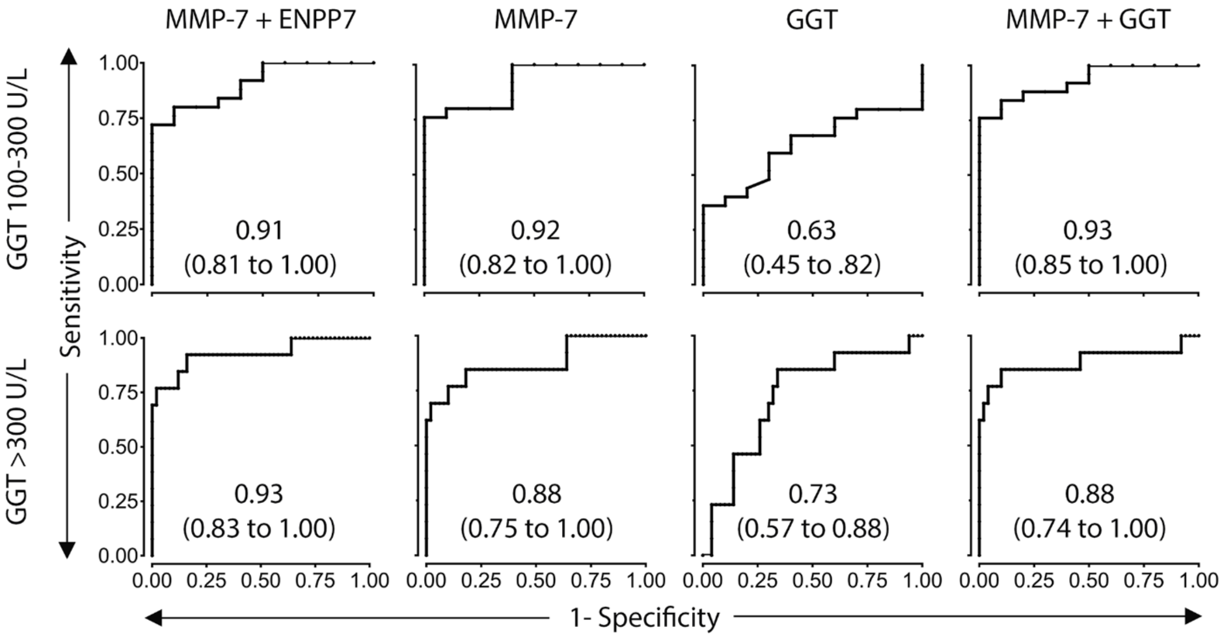

Fig. S1. Receiver operating characteristic curves and area under the curves (95% CI) for subjects with GGT 100-300 U/L (N=10 for BA, N=25 for IHC) and >300 U/L (N=50 for BA, N=13 for IHC). Groups were formed with a combination of subjects with BA and IHC in the discovery cohort and validation cohort #1 based on the availability of GGT.

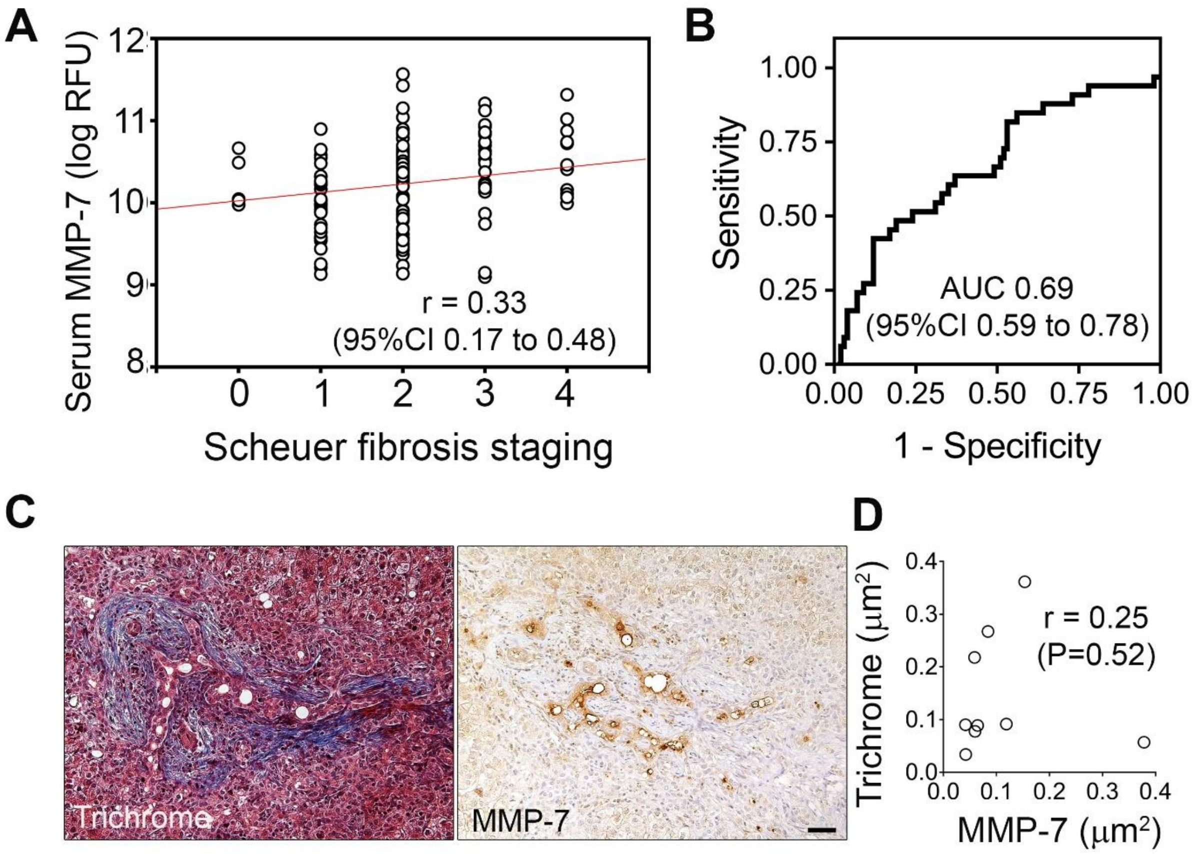

Fig. S2. Relationship between MMP-7 and hepatic fibrosis at the time of diagnosis of BA. (A) Low correlation between serum MMP-7 levels and histological fibrosis at the time of diagnosis (Spearman correlation). (B) The ROC curve and predictive value of serum MMP-7 for histological fibrosis staging (no-mild fibrosis: Scheuer fibrosis stage=0-1 vs. moderate-severe stage=2-4; N=135). (C) Representative immunostaining of MMP-7 and of trichrome staining performed in consecutive sections of the same liver biopsy from subjects with biliary atresia. Scale bar = 50 µm. (D) Surface area of tissue sections stained for MMP-7 by immunohistochemistry and trichrome stain in biliary atresia (N=9). Dots represent average surface area of DAB+ (or trichrome+) in portal tract areas of each sample. AUC, area under the curve; RFU, relative fluorescent unit.

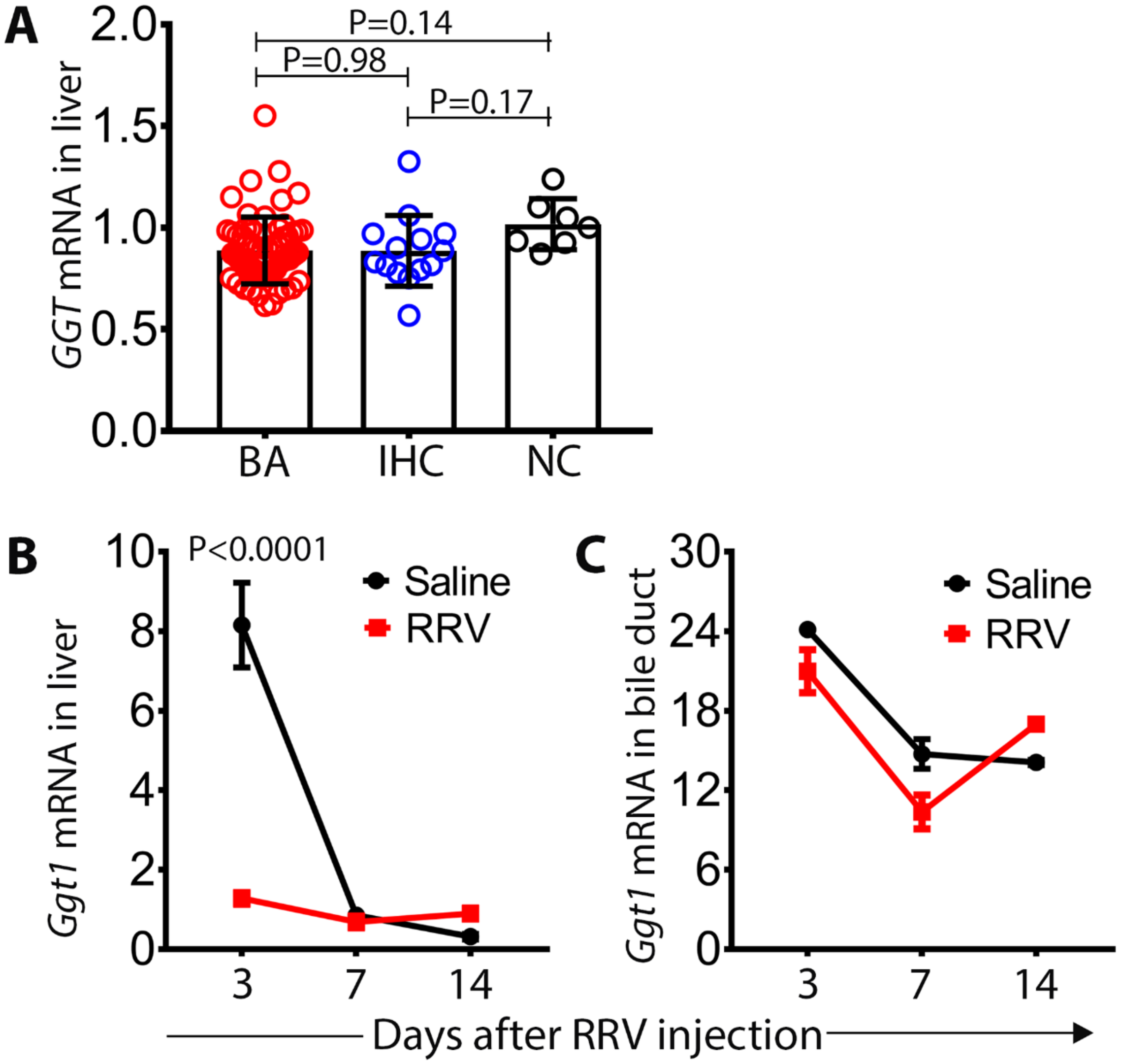

Fig. S3. *GGT* mRNA expression (A) hepatic *GGT* mRNA in biliary atresia (BA; N=64), intrahepatic cholestasis (IHC; N=14) and normal controls (NC; N=7). (B) *Ggt1* mRNA in the experimental BA; liver and bile duct (C). N= 3-4/group.

Table S1: List of 76 serum proteins differentially expressed biliary atresia (BA) and intrahepatic cholestasis (IHC) (FDR adjusted P-value ≤ 0.05)

Protein	Protein Full Name	EntrezGene	Mean expression* in BA	Mean expression* in IHC	Mean expression* in NC	Adjusted P BA vs IHC	Fold change BA : IHC
MMP-7	Matrilysin	MMP7	10.148	8.724	8.099	1.05E-14	4.153
IL-8	Interleukin-8	CXCL8	9.735	8.669	7.971	4.16E-08	2.902
SARP-2	Secreted frizzled-related protein 1	SFRP1	8.038	7.062	6.708	0.000111	2.655
IGFBP-1	Insulin-like growth factor-binding protein 1	IGFBP1	8.800	7.991	8.894	0.000396	2.244
Midkine	Midkine	MDK	9.743	8.950	8.250	0.006104	2.209
PIGR	Polymeric immunoglobulin receptor	PIGR	10.594	9.807	8.613	0.000535	2.195
Collectin Kidney 1	Collectin-11	COLEC11	10.273	9.517	8.857	1.3E-07	2.130
ADAM 9	Disintegrin and metalloproteinase domain-containing protein 9	ADAM9	8.882	8.145	7.889	1.28E-05	2.090
GCP-2	C-X-C motif chemokine 6	CXCL6	9.152	8.450	7.943	0.003651	2.019
ENPP7	Ectonucleotide pyrophosphatase/phosphodiesterase family member 7	ENPP7	10.828	10.154	8.993	0.027054	1.961
IL-18 Ra	Interleukin-18 receptor 1	IL18R1	7.614	7.062	6.581	2.16E-07	1.737
BSSP4	Brain-specific serine protease 4	PRSS22	8.932	8.454	8.205	0.000205	1.613
HGF	Hepatocyte growth factor	HGF	7.895	7.421	7.253	0.000163	1.608
TSP2	Thrombospondin-2	THBS2	8.536	8.089	7.078	0.013479	1.563
LIMP II	Lysosome membrane protein 2	SCARB2	7.134	6.692	6.590	0.000202	1.557
PRL	Prolactin	PRL	9.061	8.624	8.560	0.002421	1.547
ASAH2	Neutral ceramidase	ASAH2	7.875	7.439	7.884	0.006097	1.547
Fractalkine/ CX3CL-1	Fractalkine	CX3CL1	8.336	7.940	7.834	0.000357	1.486
FSTL3	Follistatin-related protein 3	FSTL3	9.990	9.595	9.083	0.00035	1.485
annexin II	Annexin A2	ANXA2	9.587	9.229	9.074	0.012437	1.430
TIMP-1	Metalloproteinase inhibitor 1	TIMP1	6.952	6.616	6.242	8.71E-05	1.399
Activin A	Inhibin beta A chain	INHBA	9.330	9.009	8.699	0.047014	1.378

Cathepsin B	Cathepsin B	CTSB	8.728	8.421	7.493	0.041262	1.359
6Ckine	C-C motif chemokine 21	CCL21	9.583	9.278	9.218	0.006979	1.357
BAFF	Tumor necrosis factor ligand superfamily member 13B	TNFSF13B	8.308	8.023	7.872	0.044416	1.330
C7	Complement component C7	C7	7.285	7.026	6.950	0.000745	1.296
LG3BP	Galectin-3-binding protein	LGALS3BP	7.128	6.871	7.237	0.046014	1.293
IGFBP-2	Insulin-like growth factor-binding protein 2	IGFBP2	11.543	11.287	11.090	0.027364	1.292
PTHrP	Parathyroid hormone-related protein	PTHLH	7.188	6.958	6.952	0.036308	1.258
a1-Antichymotrypsin	Alpha-1-antichymotrypsin	SERPINA3	12.183	11.978	12.285	0.001733	1.227
NID2	Nidogen-2	NID2	9.298	9.168	9.079	0.047006	1.138
CTACK	C-C motif chemokine 27	CCL27	7.544	7.655	7.484	0.004261	0.894
CRDL1	Chordin-like protein 1	CHRDL1	7.339	7.470	7.481	0.007267	0.877
DLL1	Delta-like protein 1	DLL1	7.910	8.047	8.057	0.025671	0.872
IGFBP-6	Insulin-like growth factor-binding protein 6	IGFBP6	6.074	6.217	6.156	0.012437	0.867
IGFBP-5	Insulin-like growth factor-binding protein 5	IGFBP5	6.490	6.635	6.715	0.044088	0.866
CFC1	Cryptic protein	CFC1	8.288	8.438	8.496	0.008787	0.861
GX	Group 10 secretory phospholipase A2	PLA2G10	6.619	6.780	6.661	0.036161	0.851
CHST6	Carbohydrate sulfotransferase 6	CHST6	6.313	6.481	6.390	0.044378	0.846
RGMB	RGM domain family member B	RGMB	7.843	8.014	7.969	0.005148	0.843
Gelsolin	Gelsolin	GSN	6.493	6.666	6.713	0.008173	0.841
TNR4	Tumor necrosis factor receptor superfamily member 4	TNFRSF4	7.164	7.341	7.292	0.012211	0.838
WIF-1	Wnt inhibitory factor 1	WIF1	8.058	8.242	8.457	0.036308	0.832
LSAMP	Limbic system-associated membrane protein	LSAMP	8.129	8.323	8.293	0.027364	0.823
SET	Protein SET	SET	9.406	9.602	9.894	0.035958	0.822
EPHB2	Ephrin type-B receptor 2	EPHB2	7.650	7.855	7.698	0.035958	0.814
LRIG3	Leucine-rich repeats and immunoglobulin-like domains protein 3	LRIG3	9.602	9.839	9.731	0.044256	0.789
Kallikrein 11	Kallikrein-11	KLK11	8.107	8.346	8.074	0.012437	0.787

Notch-3	Neurogenic locus notch homolog protein 3	NOTCH3	6.481	6.721	6.709	0.027102	0.786
CSF-1	Macrophage colony-stimulating factor 1	CSF1	8.055	8.296	7.764	0.006979	0.785
TSP4	Thrombospondin-4	THBS4	11.330	11.573	11.368	0.047167	0.784
RET	Proto-oncogene tyrosine-protein kinase receptor Ret	RET	7.273	7.526	7.559	0.001617	0.776
MATN2	Matrilin-2	MATN2	10.218	10.479	10.308	0.006784	0.770
Aggrecan	Aggrecan core protein	ACAN	8.641	8.904	9.103	0.036308	0.768
Periostin	Periostin	POSTN	9.290	9.558	9.279	0.037223	0.764
HPLN1	Hyaluronan and proteoglycan link protein 1	HAPLN1	9.404	9.676	9.863	0.043921	0.762
Angiopoietin-2	Angiopoietin-2	ANGPT2	7.587	7.863	7.512	0.016486	0.759
EMR2	EGF-like module-containing mucin-like hormone receptor-like 2	EMR2	6.545	6.825	6.631	0.047006	0.756
ROR1	Tyrosine-protein kinase transmembrane receptor ROR1	ROR1	8.386	8.666	8.670	0.006665	0.756
bFGF-R	Fibroblast growth factor receptor 1	FGFR1	9.603	9.900	9.832	0.000878	0.743
PCI	Plasma serine protease inhibitor	SERPINA5	9.821	10.119	10.238	0.012211	0.742
DAF	Complement decay-accelerating factor	CD55	8.248	8.546	8.429	0.000059	0.742
IL-11 RA	Interleukin-11 receptor subunit alpha	IL11RA	9.429	9.729	9.788	0.007267	0.741
SHBG	Sex hormone-binding globulin	SHBG	8.616	8.920	8.740	0.016486	0.738
MRC2	C-type mannose receptor 2	MRC2	9.294	9.603	9.425	0.012211	0.735
MMP-13	Collagenase 3	MMP13	9.299	9.625	9.569	0.041262	0.722
BOC	Brother of CDO	BOC	8.855	9.187	9.024	0.034184	0.717
Sonic Hedgehog	Sonic hedgehog protein	SHH	7.056	7.412	7.593	0.022793	0.700
CDON	Cell adhesion molecule-related/down-regulated by oncogenes	CDON	10.169	10.555	10.610	0.009042	0.680
URB	Coiled-coil domain-containing protein 80	CCDC80	8.319	8.708	8.528	0.000202	0.678
kallikrein 8	Kallikrein-8	KLK8	9.438	9.843	9.645	0.008444	0.667
CK-MB	Creatine kinase M-type:Creatine kinase B-type heterodimer	CKB CKM	6.502	6.914	7.458	0.02934	0.662
SLIK5	SLIT and NTRK-like protein 5	SLITRK5	9.511	9.961	9.939	0.002152	0.638

IL-17B	Interleukin-17B	IL17B	6.220	6.672	6.401	0.003651	0.636
DAN	Neuroblastoma suppressor of tumorigenicity 1	NBL1	8.763	9.304	9.067	0.000435	0.582
Carbonic anhydrase 6	Carbonic anhydrase 6	CA6	6.961	7.569	8.504	0.029891	0.544

*Serum expression expressed in log-transformed relative fluorescence unit.

BA: Biliary atresia

IHC: Intrahepatic cholestasis

NC: Normal controls

Table S2. Sensitivity of MMP-7 to predict biliary atresia from the validation cohort #2 (BA subjects, N = 105)

Cutoff (probability)	True positive	False negative	Sensitivity
0	105	0	1
0.1	104	1	0.99
0.2	103	2	0.98
0.3	103	2	0.98
0.4	102	3	0.97
0.5	101	4	0.96
0.6	100	5	0.95
0.7	98	7	0.93
0.8	97	8	0.92
0.9	86	19	0.82
1.0	0	105	0

Table S3. Primer sequences and PCR product sizes for mouse *Mmp-7, Ggt1, Th1, TNF-α* and *IL8* related genes

Gene	Mouse Primer Sequences	Annealing Temp (˚C)	Product size (bp)
Mmp-7	For: 5’-GCACATCAGTGGGAACAGGC-3’ Rev: 5’-GCATTTCCTTGAGGTTGTCCAC-3’	51.9	103
Ggt1	For: 5’-AGCCCGCATCACTGATGAAA-3’ Rev: 5’-CTCTGAAACCGCAGACAGGT-3’	76.4	103
Cxcl1	For: 5’-CAAACCGAAGTCATAGCCACAC-3’ Rev: 5’-TTGGGGACACCTTTTAGCATC-3’	54.0	107
Cxcl2	For: 5’-TGAACAAAGGCAAGGCTAACTG-3’ Rev: 5’-AAGTGAACTCTCAGACAGCGAGG-3’	56.4	146
Cxcl5	For: 5’-TGGGCAGTGACAAAAAGAAAGC-3’ Rev: 5’-AAATCCGTGGGTGGAGAGAATC-3’	55.2	138
Cxcl9	For: 5’-GAGCTAGATAGACCTCACCAAG-3’ Rev: 5-‘CCATTAGCACCATCTCTGA-3’	54.7	101
Cxcl10	For: 5’-TCGCTCAAGTGGCTGGGATG-3’ Rev: 5’-TAGGGAGGACAAGGAGGGTGTG-3’	58.9	117
Cxcl11	For: 5'- CAAGGCTTCCTTATGTTCAAACAG-3’ Rev: 5'- TTACAGAAGCTTTCTCGATCTCTG-3’	85	94
Tnfa	For: 5'-AAGGGAGAGTGGTCAGTTGCC-3’ Rev: 5'-CCTCAGGGAAGAGTCTGGAAAGG-3’	55	95
Ifng	For: 5′-GGCTGTCCCTGAAAGAAAGC -3’ Rev: 5′-GAGCGAGTTATTTGTCATTCGG-3’	53.5	102
IL12p40	For: 5′-AAAGGCTGGGTATCGGTGG-3’ Rev: 5′-ACTGGCTGTGCTGGAACTCC-3’	55.7	118
Gapdh	For: 5′-TGGTTTGACAATGAATACGGCTAC-3’ Rev: 5′-GGTGGGTGGTCCAAGGTTTC-3’	55.1	92