1 **Pathobiochemical signatures of cholestatic liver disease in bile duct**

2 **ligated mice**

3

4 **Kerstin Abshagen[1][¶][*], Matthias König[2][¶], Andreas Hoppe[2], Isabell Müller[1], Matthias Ebert[4], Honglei**

5 **Weng[4], Herrmann-Georg Holzhütter[2], Ulrich M. Zanger[3], Johannes Bode[5], Brigitte Vollmar[1], Maria**

6 **Thomas[3], Steven Dooley[4]**

7 [1]Institute for Experimental Surgery, University Medicine Rostock, 18057 Rostock, Germany

8 [2]Institute for Biochemistry, Computational Systems Biochemistry Group, Charité University Medicine

9 Berlin, 10117 Berlin, Germany

10 [3]Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology, 70376 Stuttgart, and University of

11 Tuebingen, Tuebingen, Germany

12 [4]Department of Medicine II, Section Molecular Hepatology, Medical Faculty Mannheim, Heidelberg

13 University, 68167 Mannheim, Germany

14 [5]Department for Gastroenterology, Hepatology and Infectiology, Heinrich-Heine University of Düsseldorf,

15 40225 Düsseldorf

16

17 [¶]These authors contributed equally to this work

18

19 *Correspondence should be addressed to:

20 Kerstin Abshagen, PhD

21 Institute for Experimental Surgery

22 Rostock University Medical Center

23 Schillingallee 69a

24 18057 Rostock

25 Germany

26 phone: +49 381 494 2503

27 fax: +49 381 494 2502

28 E-mail: kerstin.abshagen@uni-rostock.de

29

1   **Email addresses:**

2   matthias.koenig@charite.de

3   andreas.hoppe@charite.de

4   Maria.Thomas@ikp-stuttgart.de

5   isabell-mueller86@web.de

6   matthias.ebert@umm.de

7   honglei.weng@medma.uni-heidelberg.de

8   hergo@charite.de

9   Uli.Zanger@ikp-stuttgart.de

10   Johannes.Bode@med.uni-duesseldorf.de

11   brigitte.vollmar@med.uni-rostock.de

12   Steven.Dooley@medma.uni-heidelberg.de

13

14   **Running title:** Cellular and molecular signatures of cholestasis

1 Abstract

2 **Background:** Disrupted bile secretion leads to liver damage characterized by inflammation, fibrosis,

3 eventually cirrhosis, and hepatocellular cancer. As obstructive cholestasis often progresses insidiously,

4 markers for the diagnosis and staging of the disease are urgently needed. To this end we compiled a

5 comprehensive data set of serum markers, histological parameters and transcript profiles at 8 time points

6 after bile duct ligation (BDL) in mice.

7 **Results:** Statistical analysis of the more than 6,000 data points revealed distinct temporal phases of

8 disease. Putting time-courses of biochemical, histochemical and mRNA transcripts (=factors) together,

9 time-series correlation analysis defined 6 distinct clusters related to different phases of disease

10 progression. The count of CTGF-positive cells provided the most reliable overall measure for disease

11 progression at the histological level, bilirubin at biochemical level, and metalloproteinase inhibitor 1

12 (Timp1) on transcript level. Prominent molecular events exhibited by strong transcript peaks are found for

13 the transcriptional regulator Nr0b2 (Shp) and 1,25-dihydroxyvitamin D(3) 24-hydroxylase (Cyp24a1) at 6

14 h. Based on these clusters we constructed a decision tree to identify factor combinations that can be used

15 as markers for specific time intervals of disease progression. Best prediction of onset of disease is

16 achieved by fibronectin (Fn1), the initial phase by (Cyp1a2), transition to the perpetuation phase by

17 collagen α-1 I (Col1a1), and transition to the progression phase by interleukin 17-a (Il17a), with early and

18 late progression separated by Col1a1. Notably, these predictions remained stable even for randomly

19 chosen small sub-sets of factors from the clusters.

20 **Conclusion:** Our explorative study demonstrate the existence of a reproducible and well-coordinated

21 temporal gene-expression program underlying disease progression in BLD cholestasis. To further

22 elucidate the regulatory network behind this expression program additional studies are needed which have

23 to include knowledge on transcription factor activation due to accumulation of bile salts and shared

24 transcriptions factor binding motifs of genes belonging to the same transcript cluster.

25

26 **Keywords:** liver injury, mouse, systems biology, fibrosis, cell proliferation, bile duct ligation, cholestasis,

27 morphological profiling, VirtualLiver Network

1 Background

2   Cholestatic liver diseases are caused by an impaired flow of the bile from the liver to the duodenum.

3   The main fraction of organic molecules in the bile are bile salts which are strong detergents required for

4   the extraction of lipids from the apical membrane of hepatocytes into the bile fluid and for the emulgation

5   of lipids in the gut. Moreover, the bile fluid, contains numerous endogenous end products (e.g. bilirubin)

6   and potentially toxic compounds cleared by the liver. Hence, accumulation of bile compounds in the

7   cholestatic liver causes an unspecific cell damage that initiates a cascade of inflammatory events:

8   Necrosis of hepatocytes and cholangiocytes, activation of macrophages releasing of pro-inflammatory

9   cytokines and chemokines, neutrophil infiltration, cholangiocyte and hepatocyte proliferation, stellate cell

10   activation with progressive fibrosis causing secondary biliary cirrhosis, ultimately leading to liver failure

11   {Wang2005}.

12   Multiple pathologies may represent the primary trigger of impaired bile flow as, for example, defects in

13   the export of the bile from hepatocytes to the extracellular space (hepatocellular cholestasis) or

14   obstruction of bile ducts by gall stones or local tumor impingement (extrahepatic cholestasis)

15   {Woolbright2013}. Among the most common causes of cholestatic liver disease in the adult population are

16   primary biliary cirrhosis (PBC) and primary sclerosing cholangitis (PSC), while biliary atresia and Alagille

17   syndrome are commonly found in the pediatric population {Lindblad1977,Trautner1998, Trauner2004,

18   Wang2013, Nguyen2014}.

19   One of the main causes of cholestasis is biliary tract obstruction, which is a common mechanism of

20   hepatic injury in a variety of clinical settings, including obstructing neoplasms, post-operative bile duct

21   injury, biliary atresia, sclerosing cholangitis, and primary biliary cirrhosis. The corresponding experimental

22   model to induce obstructive cholestatic injury in mice and rats is surgical bile duct ligation (BDL)

23   {Tag2015, Georgiev2008}, which results in stereotypical histopathological changes of cholestasis. The

24   BDL experimental model has been well described and evaluated in rats and mice and is widely used to

25   study cholestatic liver injury and fibrogenesis {Huss2010}.

26   Chronic liver diseases (CLD) like cholestasis represent with characteristic temporal morphologic,

27   biochemical and molecular changes in liver and serum. For instance, in the BDL model of cholestasis an

1   early phase of acute hepatocyte injury is followed by a proliferative response, up-regulation of pro-

2   inflammatory cytokines and pro-fibrotic metabolic enzymes, and subsequent development of fibrosis after

3   around 7 days {Georgiev2008, Huss2010}. Such alterations can be exploited to identify biomarkers

4   reporting on specific stages of disease progression and severity of the disease. Hitherto semi-quantitative

5   morphological scoring is the standard technique for grading in CLD. However, nowadays the availability of

6   high-throughput technologies enables to flank the histological assessment of injured tissue with a

7   comprehensive molecular profiling of gene transcripts, gene products (proteins) and metabolites. Such

8   analyses not only provide a more detailed characterization and thus more refined staging of disease

9   progression, they also lead to a deeper understanding of the molecular networks governing the

10 histological and physical alterations observed at higher scales of tissue organization. Extracting from high-

11 throughput data relevant information for clinical translation is still in its beginnings. The identification of key

12 processes triggering the transition between different phases of disease progression requires mathematical

13 analyses which take into account multiple parallel processes, process dynamics, and experimental data

14 from different levels of cellular organization [4].

15   In this study we explicitly wanted to systemize existing and newly acquired knowledge on the

16 molecular biomarkers of cholestasis and to analyze the time-dependent disease progression following

17 BDL in mice in a comprehensive manner. Our approach sets itself apart from existing studies which either

18 describe the time course of a limited number selected parameters after BDL {Georgiev2008, Huss2010} or

19 provide gene expression profiling for a limited number of time points, thereby missing the acute damage

20 after BDL in the first 24h and long-term effects after 7 days {Tanaka2006}. None of these preceding

21 studies provide predictive models based on the acquired time course data. Therefore, a central aim of our

22 study was to identify molecular markers for the temporal progression of BDL cholestasis by correlating

23 high-accuracy image data and transcriptional profiles of preselected targets with pathobiochemical

24 markers, whereby all data were obtained at different time points after onset of obstructive cholestasis.

25   We collected from 8 different time points after BDL more than 6,000 experimental data points,

26 comprising immunohistochemistry, biochemistry and molecular profiling measures. Statistical methods

27 were applied to unravel robust interrelations in this large-scale data set and find clusters of parameters

28 corresponding to characteristic time profiles in cholestatic injury induced by BDL. We correlated level and

1 timing of pathophysiological events with transcriptional changes in order to define molecular markers, and

2 developed predictive decision trees which allow assessment of the different disease phases occurring

3 during the development of cholestasis based on a subset of biomarkers.

4 # Materials and Methods

5 **Ethic Statement**

6 All experiments were approved by the local government Landesamt für Landwirtschaft,

7 Lebensmittelsicherheit und Fischerei Mecklenburg-Voprommern (LALLF M-V/TSD/7221.3-1.2-049/09) and

8 performed in accordance with the German legislation on protection of animals and the National Institutes

9 of ''Health Guide for the Care and Use of Laboratory Animals'' (Institute of Laboratory Animal Resources,

10 National Research Council; NIH publication 86-23 revised 1985).

11 **Mice**

12 Male C57BL/6J (Charles River Laboratories, Sulzfeld, Germany) at 8–10 weeks of age with a body

13 weight of 23–26 g were kept on water and standard laboratory chow ad libitum.

14 **Surgical procedure and experimental groups**

15 Mice were anesthetized by breathing isoflurane (1.5 vol%). BDL was performed after midline

16 laparotomy. The common bile duct was ligated three times with 5-0 silk and transected between the two

17 most distal ligations. Sham operation was performed similarly, except for ligation and transection of the

18 bile duct (0 h, n=5). All surgical procedures were performed under aseptic conditions. Animals were

19 allowed to recover from anesthesia and surgery under a red warming lamp and were held in single cages

20 until subsequent experiments followed at postoperative hours 6, 12, 18 and 30 (n=5 animals per time

21 point), and at 2, 5 and 14 days after BDL (n=5 animals per time point). Sham-operated animals without

22 BDL served as controls (n=5). To analyze the regenerative response in regard to proliferation of different

23 cell types, 5-bromo-2-deoxyuridine (BrdU; 50 mg/kg bw ip) was injected 1 h prior to harvest of liver tissue.

24 BrdU incorporation into DNA was analyzed by immunohistochemistry. To obtain blood and liver samples,

25 mice were killed at the indicated time points. Liver tissue was partially embedded in paraffin for

26 morphology analysis and snap frozen for molecular biology and biochemistry analyses. In addition, liver

1   tissue served for the parallel Taqman qRT-PCR using microfluidic Fluidigm Biomark™ platform (Fluidigm,

2   CA, USA).

3   **Hematological measurements and plasma enzyme levels**

4   Animals were anesthetized and exsanguinated by puncture of the vena cava inferior. Red blood cell

5   and blood platelets count, hemoglobin, and hematocrit were assessed with an automated cell counter

6   (Sysmex KX-21, Sysmex). Plasma activities of alanine aminotransferase (ALT), aspartate

7   aminotransferase (AST) and glutamate dehydrogenase (GLDH) were measured spectrophoto-metrically.

8   **Assays**

9   EDTA plasma served for the analysis of albumin as a parameter of liver function, which was

10   determined with a commercially available enzyme-linked immunosorbent assay kit in accordance with the

11   manufacturer's instructions (Assaypro, MO, USA).

12   **Histopathology and image analysis**

13   Liver tissue samples were fixed in formalin for 2 to 3 days and embedded in paraffin. 5 µm sections

14   were stained with hematoxylin and eosin (H&E) for routine examination and quantification of bile infarcts.

15   Sirius red staining served for quantification of collagen deposition. All samples from a series of

16   experiments were stained simultaneously and evaluated in a blinded manner. For histomorphometric

17   analysis, images of 20 random low power fields (x10 magnification, Olympus BX 51, Hamburg, Germany)

18   were acquired with a Color View II FW camera (Color View, Munich, Germany) and evaluated using an

19   image analysis system (Adobe Photoshop, Adobe Systems, Uxbridge, UK). Fibrosis deposition was

20   quantified as a percentage of Sirius red stained area compared with the total section area. The surfaces of

21   large centrilobular veins and large portal tracts were excluded from this calculation. Bile infarcts were

22   quantified in H&E-stained sections in a similar manner and the percentage of the focal necrosis surface to

23   the whole liver section area was assessed.

24   **Immunohistochemistry and image analysis**

25   For analyzing DNA-incorporated BrdU in liver cells, 4 µm sections collected on poly-L-lysine-coated

26   glass slides were incubated with monoclonal mouse anti-BrdU antibody (1:50; Dako Cytomation,

27   Hamburg, Germany) overnight at 4°C, followed by horseradish-peroxidase (HRP)-conjugated goat anti-

28   mouse immunoglobin (LSAB kit plus; Dako). Sites of peroxidase-binding were detected by 3,3`-

1   diaminobenzidine (Dako). Sections were counterstained with hemalaun. BrdU-positive hepatocellular

2   nuclei were counted in a blinded manner within 30 consecutive high power fields (HPF) (x40 objective,

3   numerical aperture 0.65) and are given as cells/mm². In analogy, BrdU-expressing non-parenchymal cells

4   were assessed and also given as cells/mm².

5   To specify the proliferative response of non-parenchymal cells upon BDL, we performed double

6   immunohistochemistry of BrdU with specific markers for different liver cells: F4-80/BrdU for Kupffer cells

7   and SM22α/BrdU for BEC. For each protocol, the immune-staining procedure for the specific marker was

8   conducted after the BrdU staining protocol. Resident liver macrophages were analyzed using the F4-80

9   antigen (1:10; Serotec, Oxford, UK). Overnight incubation (4°C) with the first antibody (polyclonal rat anti-

10   F4-80) was followed by alkaline-phosphatase (AP) conjugated mouse anti-rat immunoglobulin (1:200;

11   Santa Cruz Biotechnology, Santa Cruz, CA, USA). The sites of AP-binding were detected using the

12   chromogen fuchsin (Dako).

13   BECs and oval cells were detected by overnight incubation (4°C) with a polyclonal rabbit anti-SM22α

14   antibody (1:100; Abcam, Cambridge, UK) followed by AP conjugated goat anti-rabbit immunoglobulin as

15   secondary antibody (1:100; Dako). The sites of AP-binding were detected by Permanent Red (Dako).

16   Slides were viewed under a light microscope (Olympus BX 51) and the number of BrdU-positive cells

17   co-expressing F4-80 or SM22α were counted in a blinded manner within 30 consecutive high power fields

18   (HPF) (x40 objective, numerical aperture 0.65) and are given as cells/mm².

19   Antibodies for detection of α-SMA in HSCs and of S100a4-positive cells were from DAKO (M0851

20   and A5114, 1:500 and 1:200 dilution, respectively). CTGF antibody was from Santa Cruz (sc-1439, 1:200

21   dilution). Sections were de-paraffinized in serial ethanol dilutions. After a PBS wash, sections were

22   transferred into 10 mM sodium citrate buffer (pH 6.0) and antigen unmasking was performed in a

23   microwave. After cooling down, sections were incubated in peroxidase blocking reagent (Dako) for 1 h and

24   with primary antibodies overnight at 4°C. EnVision peroxidase (Dako) was applied for 1 h at room

25   temperature after a PBS wash next day. Sections were developed with diaminobenzidine for 5 minutes.

26   The number of α-SMA-, CTGF- and S100a4-positive cells was quantified under a Leica light microscope

27   (x20) by counting three fields.

28   **High-throughput quantitative Taqman RT-PCR analysis**

1   Total RNA was isolated from the liver tissue samples using RNeasy Mini Kit including on column

2   genomic DNA digestion with RNase free DNase Set (Qiagen, Hilden, Germany). RNA was reverse

3   transcribed to cDNA with TaqMan Reverse Transcription Reagents (Applera GmbH, Darmstadt,

4   Germany). For quantitative real-time PCR, we used the Fluidigm's Biomark high-throughput qPCR chip

5   platform (Fluidigm Corporation, San Francisco, CA, USA) with pre-designed gene expression assays from

6   Applied Biosystems, according to the manufacturer's instructions [6]. Data were analyzed using the ddCt

7   method and expression values were normalized to the expression levels of the Gapdh gene.

8   **Statistical data analysis**

9   *Dimension reduction*: A one-way analysis of variance (ANOVA) was applied to reduce the data set to

10   the parameter subset showing significant ($p_{adj} < 0.05$) up- or down-regulation during the time course.

11   Multiple testing correction was performed using the Holm procedure {Holm1979}. To specifically test for

12   the initial changes a two-tailed unpaired t-test (Welch correction for nonhomogeneity of variance) was

13   performed for all factors between the classes 0h and 6h.

14   *Correlation analysis*: Correlation between two factor time courses was calculated via Y^{S3} a modified

15   correlation coefficient based similarity measure for clustering time-course gene expression data

16   {Son2008}. The correlation $Y_{i,j}^{S3}$ between two factors i and j is the linear combination of three terms: i) a

17   classical correlation part based on Spearman correlation $S_{i,j}^*$, ii) $A_{i,j}^{**}$ accounting for the similarity in changes

18   between the two time courses, and iii) $M_{i,j}^*$ comparing the location of the minimum and maximum values of

19   the time course (see {Son2008} and Supplement 2 for definitions). $S_{i,j}^*$ is calculated on individual data

20   points, $A_{i,j}^{**}$ and $M_{i,j}^*$ on the mean time courses averaged over the repeats per time point. For the calculation

21   of the similarity of changes within time courses, $A_{i,j}^{**}$, we used Spearman (S) correlation which is more

22   robust against outliers as Pearson (R) correlation:

23
$$A_{i,j}^{**} = (S(d_i, d_j) + 1)/2$$

24
$$Y_{i,j}^{S3} = \omega_1 S_{i,j}^* + \omega_2 A_{i,j}^{**} + \omega_3 M_{i,j}^*$$

25   $S_{i,j}^*$In the analysis the following weights were used: $\omega_1 = 0.5$, $\omega_2 = 0.3$, $\omega_3 = 0.2$. All reported correlations

26   are $\bar{Y}_{i,j}^{S3}$ values in the interval [-1,1]:

27
$$\bar{Y}_{i,j}^{S3} = 2(Y_{i,j}^{S3} - 0.5)$$

1

$$A_{i,j}^{**} = (S(d_i, d_j) + 1)/2$$

Cluster analysis of the correlation matrix used complete-linkage hierarchical clustering with Euclidian distance measurement. This combination of complete-linkage with Y^S provided the best enrichments on gene-expression time-series in a recent comparisons of methods {Jaskowiak2014, Jaskowiak2013}. The number of clusters $N_c = 6$, was selected as maximum number of clusters with all clusters containing more than one factor. Normalization of factors was performed separately for each factor f_k for all time points $i = 1, \ldots, N_t$ and repeats $r = 1, \ldots, N_r$ with $N_t = 8$ and $N_r = 5$ via

$$\bar{f}_k(t_{i,r}) = \frac{(f_k(t_{i,r}) - <f_k>)}{\max(f_k) - \min(f_k)}$$

Decision Trees: For the prediction of distinct time points of disease progression a regression tree with the mean normalized factor data of the 6 clusters as predictor variables and the log transformed time points $\tilde{t}_i$ as dependent variable was fitted based on recursive partitioning using *rpart* {Breiman1984}. Logarithmic transformation was performed to obtain approximately equal-distant time points.

$$\tilde{t}_i = \log(t_i + 1)$$

The regression tree was fitted using the complete trainings set $(N_s = N_t N_r = 40)$, with the minimum number of observations in a node for which a split was computed being 6, the minimum number of observations in a terminal node as 2, and the complexity parameter $c_p = 0.01$. The splitting criterion deciding which predictor variable gives the best split for nodes in the regression tree was $S_T - (S_L + S_R)$, with $S_T = \sum(\tilde{t}_i - <\tilde{t}>)^2$ the sum of squares for node T, and S_T and S_L the sums of squares for the left and right child. A leave-one-out approach was used to test the robustness of the predicted time classes and predictive performance: For each sample $(N_s = 40$ mice), the regression tree was generated under the exclusion of data from the sample, with subsequent prediction on the left out test data (see Supplement 2).

The predictive capacity of the regression tree was evaluated using all single combinations of individual factors from the clusters (88572), and a random subset of 10000 two factor combinations from each cluster. Predictions for a given combination of factors $\langle f_1 \rangle, \ldots, \langle f_6 \rangle$ from the 6 clusters were scored using the root mean square distance on log scale *d*, with the best combination of factors minimizing d

1
$$d(\langle f_1 \rangle, \ldots, \langle f_6 \rangle) = \frac{1}{N_s} \sqrt{\sum_{i=1}^{N_s} (t_i^{pre} - t_i^{exp})^2}$$

2   All computations were performed in R with source code, data and the full analysis available in Supplement

3   S2 and from https://github.com/matthiaskoenig/bdl-analysis.

4

5   Results and Discussion

6   Temporal changes of biochemical, cellular and histochemical markers after BDL

7   In mice, BDL over 14 days induces time dependently progressing stages of a secondary biliary CLD.

8   The first week after BDL begins with an acute cholestatic injury associated with necroinflammation,

9   followed by a chronic injury stage, comprising hepatitis and liver fibrosis. Macroscopically, marked dilation

10   of the gallbladder and formation of bilioma are found, associated with weight loss and a mortality rate of

11   10% in the first week due to bile leakage and rupture of the gallbladder [7]. Pathophysiologically, BDL

12   interferes with glandular liver function and hepatobiliary transport, which comprises its detoxification and

13   secretion function including bile formation. Obstruction of the bile duct leads to afflux of newly generated

14   bile fluid. The main components of the bile, bile acids and phospholipids, induce toxicity and damage

15   towards hepatocytes and cholangiocytes, therewith initiating the disease process. Rapidly after BDL, mice

16   develop obstructive jaundice and cholestasis, as displayed by markedly elevated serum transaminases

17   and bilirubin levels (Figure 1), macroscopically evident from yellow ears and urine. Within the first 30 h,

18   there is a massive release of liver enzymes, like ALT and GLDH, reflecting the initial hepatocyte damage

19   as initial pathophysiological event in the process of BDL-induced liver fibrosis (Figures 1A, B).

20   The plasma levels of diagnostic liver enzymes remained elevated over several days, but then

21   dropped to reach after 14 days values that were slightly above those of sham operated mice.

22   Concomitantly, liver detoxification capacity was deteriorated, as indicated by the rise of total bilirubin, a

23   classical plasma marker of cholestasis (Figure 1C). Notably the plasma level of albumin, an important

24   parameter for the evaluation of liver functionality remained relatively constant over the time course of 14 d

25   (Figure 1D). The systemic blood cell count showed constant levels of erythrocytes and platelets up to day

1   5 (Table 1). In contrast, leukocytes decreased by 50% during the first two days, reflecting intrahepatic cell

2   entrapment, and recovered to values of sham operated animals within the subsequent observation period

3   (Table 1). During progression of fibrosis, red blood cells, hemoglobin and hematocrit slightly decreased.

4   In consequence of intrahepatic toxic bile accumulation, progressive development of confluent bile

5   lakes is a hallmark of cholestasis. Histological quantification of bile infarcts, defined as clusters of injured

6   hepatocytes, revealed a steady rise of infarct areas until day 14 after BDL (Figure 2A). The typical

7   appearance of liver tissue at representative time points after BDL using H&E staining is depicted in Figure

8   2B. Further histopathological changes of the livers after BDL included enlargement of portal tracts,

9   accompanied by dilation of bile canaliculi and proliferation of BECs and oval cells (Figure 3A), resulting in

10   formation of artificial bile ductules (Figure 2C), a cellular response termed 'ductular reaction' [5; 8].

11   However, recent data from lineage tracing experiments indicated that these cells, however, do not

12   contribute to the population of ECM producing/fibrogenic cells, which in the BDL model is largely

13   consisting of HSCs [9].

14   The inflammatory response resulting from chronic hepatocyte injury resulted in the accumulation of

15   immune cells in the liver, among others, T cells, macrophages and dendritic cells, which were mainly

16   found within and around bile infarct areas (Figure 2C, asterisk) [10]. The inflammatory and fibrogenic

17   response is initiated by resident liver cells, primarily liver macrophages (Kupffer cells) and activated

18   hepatic stellate cells (HSCs) secreting a wide range of cytokines and chemokines which determine the

19   quality and quantity of inflammatory and consequently fibrotic responses [11; 12]. Upon parenchymal

20   damage, quiescent HSCs undergo phenotypical changes to myofibroblasts (MFBs). The most prominent

21   role of MFBs is extracellular matrix (ECM) production and reorganization, as reflected by, among others,

22   increased synthesis of α-SMA, type I collagen and TIMPs. A marked increase of the number of α-SMA-

23   and S100a4-positive cells, as measured by immunohistochemistry, was observed after BDL (Figures 3B,

24   C). Migration of MFBs to the site of injury and their contractility contribute to liver scarring. This is

25   accompanied by parenchymal cell proliferation, which began to rise at day 2 as regenerative response but

26   decreased at day 14 (Figure 3D). With a slightly faster response as compared to HSCs, starting at 30 h

27   upon BDL, Kupffer cells started to proliferate (Figure 3E). The overall hepatic proliferative response as

28   analyzed by immunohistochemistry was confirmed by elevated mRNA expression of Ki67 (Figure 4A).

1   Between days 5 and 14 after BDL, periportal alterations were associated with fibrotic changes. As

2   demonstrated by Sirius red staining, extensive fibrosis, characterized by a several-fold increase of

3   collagen deposition (Figure 4B), including bridging, was observed 5 days after BDL (Figure 4C). We

4   further stained for connective tissue growth factor (CTGF), a prominent fibrogenic cytokine and enhancer

5   of TGF-β effects [13]. Appearance of CTGF-positive cells started as early as 12 h upon BDL and their

6   count increased continuously (Figure 3F).

7   **Evidences for different time phases of diseases progression after BDL**

8   For the identification of distinct phases in disease progression induced by BDL we included besides

9   biochemical and histochemical markers time-resolved transcriptomics profiles of three different gene

10   panels related to hepatocyte metabolism, fibrogenesis, and inflammation using the Fluidigm platform

11   (Figure 5). The selection of representative genes for ADME- (absorption, distribution, metabolism, and

12   excretion) (Figure 5A) fibrogenesis- (Figure 5B), and inflammation-related genes (Figure 5C) was made

13   arbitrary based on state-of the art knowledge.

14   In a first step, a one-way analysis of variance (ANOVA) was applied to reduce the complete data set

15   of biochemical, histochemical and transcript data consisting of 153 factors to a subset of those factors

16   showing significant ($p_{adj} < 0.05$) changes during the time course of disease progression (see section

17   ANOVA in Supplement 2). This resulted in a subset of 90 factors comprising two biochemical markers

18   (bilirubin, GLDH), five histochemical markers (BEC, NHC, Kupffer cells, Sirius red), three immunostainings

19   by antibodies (CTGF, α-SMA, S100a4) and 80 genes (14/47 ADME, 22/46 fibrosis, 44/47 inflammation

20   panel). Many of the ADME genes and fibrosis genes were filtered out, whereas almost all of the

21   inflammation panel was retained. The top significant factors were: Cyp1a2, serum bilirubin, Il10rb, Tgfb1,

22   Ccl2, Cd86, Ccr2, and Mrc1. Within this subset, a bivariate time-dependent correlation analysis (see

23   Methods and Supplement 2) was performed for all pairs of factors to identify those factors displaying

24   similar temporal profiles (see Figure 7). The top correlations for the biochemical, histological and

25   immunostainings are depicted in Figure 8. Based on the obtained correlation matrix a hierarchical cluster

26   analysis was applied resulting in 6 different clusters comprising between 2 - 61 factors with distinct time

27   courses attaining their maximum at different time points (see Figure 9). To identify global connections

13

1   between factors, the consensus correlation coefficients have been subjected to a cluster analysis (Figure

2   7). RNA levels of commonly regulated genes form highly correlated clusters, e.g., (i) cytokines and growth

3   factors, such as Il6 and Tgfb (lower left, blue), (ii) fibrosis-related genes, such as Coll1a1 and Timp1

4   (bottom right, cyan), (iii) ADME-related genes, such as isoforms of Cyp24a1 and Nr0b2 (red, top right).

5   Accordingly, the clusters can be assigned to the early, middle and late phase of disease progression.

6   The found clusters comprise both 'classical' biochemical and histochemical factors and genes

7   characteristic for a specific phase of disease progression.

Correlations between transcripts and non-RNA factors

9   In what follows we discuss in more detail the biological significance of the correlations between

10   different types of factors shown in Figure 8 thereby including already existing knowledge. The time course

11   of each of the 'classical' factors contained in the ANOVA-subset can be correlated with the expression

12   time course of at least one gene (Figure 8A). Only for GLDH and Sirius Red this the correlations is weak.

13   All top correlations to genes come either from cluster 4 or cluster 1. Bilirubin, bile infarcts and the

14   immunostainings α-SMA, CTGF and S100a4 all have high correlations among each other, so do the BrdU

15   cell measurements for BEC, Kupffer and NHC (Figure 8B). GLDH and Sirius red do not have high

16   correlation with any classical factor. In the following the top correlated factors with the classical

17   biochemical, histological and immunostaining factors are discussed in the context of different aspects of

18   the disease process (Figure 8C). Main points are summarized in Figure 11 bringing the different aspects,

19   phases and markers together.

20   *Initial response*

21   Immediately after BDL, there is a massive release of liver enzymes up to day 5, followed by a drop

22   down to almost values of sham-operated livers (Figures 1A, B). It can be interpreted as the beginning

23   breakdown of hepatocellular protein synthesis or the cease of cell necrosis. GLDH like ALT shows a

24   strong initial increase, but in contrast to the latter increases further up to 18h, before it decreases

25   gradually. Prominent molecular events exhibited by strong transcript peaks are found for the members of

26   cluster 2, the transcriptional regulator Nr0b2 (small heterodimer partner, SHP, Figure 6D) and Cyp24a1

27   (mitochondrial 1,25-dihydroxyvitamin D3 24-hydroxylase, see Figure 6B) at 6 h. Nr0b2 is a repressor of

14

1   retinoid and thyroid hormone, as well as estrogen receptors and by being diminished in human cirrhotic

2   livers, and is suggested to be associated to cirrhosis and hepatic tumors [53]. This is also supported by

3   the fact that cholestatic liver fibrosis induced by BDL was increased in SHP-/- mice [54]. A second set of

4   initially up-regulated transcripts is found in cluster 3, Fn1 (fibronectin, Figure 6G) and Sult1a1

5   (Sulfotransferase 1A1). Additionally up-regulated transcripts (see Supplement 2, t-test for initial phase) are

6   Tnfrsf1a (Figure 6L), Il6st (Interleukin-6 receptor subunit beta), Osmr, Cd14, Cxcl1/2, Timp1 and Hmox1

7   (heme oxygenase), in line with reported marked increase in heme oxygenase activity following BDL in rats

8   [17]. A marked down-regulation was observed among others for Cdh2 and Pde4a and the main enzyme of

9   bile acid synthesis Cyp7a1 (cholesterol-7-α-hydroxylase), which can be seen as a fast and straightforward

10 response to cholestasis. As underlying mechanism, activation of the JNK/c-Jun pathway has been

11 reported [15]. With increasing progression Cyp7a1 increased again, most likely SREBP-mediated, since

12 plasma cholesterol concentration increases simultaneously [16].

13 *Macroscopic organ damage*

14   As documented with the H&E staining in Figure 2B, necroinflammation was caused by the BDL-

15 induced intrahepatic toxic bile accumulation with individual liver cell death and progressive development of

16 confluent bile infarct areas. The total area of infarcts increases steadily with relatively high variance

17 (Figure 2A). Bilirubin (Figure 1C) showed the highest correlation with bile infarcts followed by the

18 immunostainings by CTGF (Figure 3F) and α-SMA (Figure 3B). Gsta2 (Glutathione S-transferase A2),

19 Gstm1 (glutathione-S-transferase mu 1, Figure 6C) and Timp1 (Metalloproteinase inhibitor 1) showed the

20 highest positive correlation to bile infarcts. Timp1 is a metalloproteinase inhibitor that functions by forming

21 one to one complexes with target metalloproteinases, such as collagenases. In contrast, Cyp1a2

22 (Cytochrome P450 1A2, Figure 6A) and Cyp2e1 showed (Cytochrome P450 2E1) high negative

23 correlations to the bile infarct area, with Cyp1a2 decreased continuously after BDL in line with

24 observations in rats [60].

25 *Loss of liver function*

26   Liver function after BDL was measured using albumin (Figure 1D) and bilirubin (Figure 1C) levels.

27 Surprisingly, albumin synthesis is maintained relatively constant over the observation period of 14 days,

1   and was already filtered out based on ANOVA. Apparently, hepatic degradation of albumin is affected at

2   similar levels as its synthesis. Bilirubin on the other hand increased continuously after BDL. The highest

3   positive correlation with serum bilirubin levels are observed for the transcripts Timp1, Cd14. Ccl2 and

4   Ccl3. Ccl2 (chemokine C-C motif ligand 2) is a soluble biomarker for hepatic fibrosis in NAFLD [24].

5   Notably, a very high negative correlation was found between bilirubin and Slc10a1 (Sodium/bile acid

6   cotransporter), the hepatic sodium bile acid uptake system.

7   *Hepatic cell proliferative response*

8   During disease progression various hepatic cells start proliferating, underlined by co-immunostaining

9   with BrdU and specific markers for hepatocytes (Figure 3D), Kupffer cells (KC, Figure 3E), and BECs

10   (Figure 3A), and indirectly reflected by the marked up-regulation of Ki67 mRNA (Figure 4A). As

11   compensatory activity of the liver to restore the damaged parenchyma, hepatocytes proliferate, monitored

12   by the parameter BrdU-positive hepatocytes, which is markedly triggered between the 30h and 2d (Figure

13   3D), with very similar time course to the increase in KC, BECs and NHCs. Consequently, BEC, NHC and

14   KC all have high time course correlation with each other (Figure 8B). Additionally, S100a4 positive cells

15   are highly correlated with the proliferation time course of these three cell types.

16   Kupffer cells (liver macrophages) infiltrate the liver tissue, are activated, and proliferate starting at 30

17   h in our time course to reach a maximum at day 2 and to decrease again thereafter. They have the

18   highest correlation to the transcript Mki67 (antigen Ki-67), a known proliferation marker, followed by the

19   transcripts Birc5 (Baculoviral IAP repeat-containing protein 5, survivin) and Notch1, a transmembrane

20   receptor involved in developmental processes (Figure 8C). Yes-associated protein has been reported to

21   regulate the hepatic response after bile duct ligation via modulation of survivin [28], underlining the role of

22   survivin in hepatic tissue restructuring.

23   Biliary epithelial cells (BEC) also start proliferating after BDL, with the main increase between days 2

24   and 5, with a minor increase already observed after 30 h. Interestingly, the highest correlations to

25   transcripts are all negative, namely Cyp2c37 (Cytochrome P450 2C37), Slc10a1, Cyp2e1, and Cyp2c29.

26   Among the top correlated transcripts to the classical factors (Figure 8A) Cyp2c37 and Slc10a1 are two

27   very interesting candidates being, with Cyp1a2 and Ppara, the only ones from time course cluster 1 and

1   having high negative correlations. The decrease in Slc10a1, encoding the Na+-taurocholate co-

2   transporting polypeptide, which transports bile acids, has been shown to protect hepatocytes from

3   cholestasis-induced injury [32].

4   *Increase in fibrogenic cells*

5   In accordance to the proliferative activity of HSCs, immunohistochemical analyses demonstrate the

6   gradual rise in CTGF- (Figure 3F), α-SMA- (Figure 3B) and S100a4-expressing cells (Figure 3C),

7   reflecting activated HSC, activated KC and trans-differentiated hepatocytes. CTGF and α-SMA are highly

8   correlated (respective top correlation Figure 8C), with S100a4 having strong correlation to both of them

9   (Figure 8B).

10   CTGF is a highly pro-fibrogenic protein expressed by HSCs, BECs and hepatocytes [33; 34] and

11   mediates extracellular matrix inducing properties. Levels of have been reported significantly up-regulated

12   in experimental liver fibrogenesis and human chronic liver disease patients of various etiologies [35; 36].

13   CTGF-positive cell number is the best candidate to monitor the disease progress among the selected

14   biochemical, histological and immunostaining parameters, showing a steady increase with comparatively

15   little variance (among top ANOVA results, p_{adj}=7.9E-10). This is consistent with data from other studies,

16   which observed a correlation of increased CTGF levels with histological fibrosis stages [62; 63],

17   suggesting CTGF as a valuable diagnostic target, since it can be measured in patients' blood and have a

18   potential application in the follow-up of patients suffering from chronic liver diseases [64]. The highest

19   transcript correlation with CTGF shows Tgfb2 (cytokine TGF-β2, Figures 6I) followed by Pdgfb (platelet-

20   derived growth factor subunit B). TGF-β is the major stimulus for CTGF expression in hepatocytes [37],

21   and elevated levels of Tgfb2 were reported for bile duct ligated rats [38]. Pdgfb has been reported up-

22   regulated in liver inflammation and fibrosis [40]. Additionally, there are considerable correlations to several

23   other genes such as Tgfbr2, encoding the transforming growth factor β receptor 2 [41], Cd14, Cxcl5, Ccr2,

24   and Timp1.

25   α-SMA-positive cells, a marker of HSC activation, increase steadily during disease progression

26   (Figure 3B), highly correlated to CTGF. Consequently, the top transcript correlations are the very similar:

27   Tgfb2, Cxcl4, Timp1, Tnc, and Pdgfb. Notably, α-SMA staining, but also CTGT and S100a, are highly

1   negatively correlated to Cyp1a2 (cytochrome P450 1A2, Figure 6A), known to be lowered in liver cirrhosis

2   mediated by AhR [27], and to Ppara (Figure 8A).

3   S100a4-positive cells show a steady rise until day 2, after which it stays elevated (Figure 3C).

4   S100a4-positive cells are also a good marker for disease progression with a similar time course than

5   CTGF (Figure 3B), but larger variation than CTGF from 18h on. Many transcripts are highly correlated with

6   S100a4, e.g. Pdgfb, Birc5, Tgfb2 or Notch1.

7   *Fibrosis*

8   The progression of fibrogenesis was histomorphologically characterized by excessive deposition of

9   extracellular matrix visible by Sirius red staining of liver slices (Figure 4BC). Surprisingly, Sirius red did not

10   have high correlations to other factors (Figure 8ABC), mainly due to the high variability in measurements

11   from 0h to 2d and was in this study not a very reliable predictor for fibrogenesis. After day 2 a strong

12   increase in Sirius red was oberserved. Both Col1a1 (fibrillar collagen 1α1, Figure 6E) and Col3a1 (fibrillar

13   collagen 3α1, Figure 6F) transcripts, which predominantly exist in fibrotic livers, showed up-regulation

14   beginning 30h after BDL and a continuous increase with severity of liver fibrosis up to 14 days. Among the

15   peptide mediators, Tgfb1 and Tgfb2 (Tgf-β isoforms 1 and 2) expression was increasing after 2d,

16   confirming its postulated role as fibrogenic master cytokine [14]. Tgfb, encoding the cytokine TGF-β is well

17   known to correspond with the fibrotic process in a positive feedback loop [30]. Associated with induction of

18   fibrogenesis-related genes (Figure 5B), which particularly are representative for HSC activation, the

19   dynamics of the inflammation gene signature (Figure 5C) nicely matches with the increase in the number

20   of proliferating Kupffer cells (F4-80/BrdU staining values) observed from day 2 onwards (Figure 3E). Very

21   low expression levels were present immediately after BDL, except for the chemokines Cxcl1 und 2.

22   Starting at time points between 2 and 5 d after BDL, most cytokines and chemokines in the list were

23   strongly upregulated for a longer period of time, until day 14. During the perpetuation phase (18h-2d),

24   paracrine and autocrine cytokines amplify hepatic inflammation and HSC activation resulting in continued

25   ECM remodeling, being characterized by the enhanced mRNA expression of both fibrillar collagen1α1 and

26   3α1 (Figures 6E and F).

1 **Markers of disease progression**

2   The main focus of this study was to detect factors and factor combinations which characterize

3 particular stages of the disease process best. In what follows we discuss in more detail the biological

4 significance of the 6 time course clusters.

5   Cluster 1 (Figure 9A) decreases continuously over time, with no classical factor included in the c1.

6 Most of the members (8/11) come from the ADME chip, with exception of Cyp2e1, Rarres1 and Egfr

7 coming from the fibrosis chip. All members have very high significance in the ANOVA, with exception of

8 Rarres1 ($p_{adj} = 0.036$). The ADME genes of c1 decrease during the late initial, perpetuation and

9 progression phase. Top correlations with the cluster mean time course are in order of decreasing

10 correlation Cyp2c37, Cyp2e1, Cyp2e29, Ugt1a1, Cyp1a2 (Figure 6A), Rarres1 and Slc10a1 containing

11 many enzymes of the cytochrome P450 system. The down-regulation of Ugt1a1 (UDP-glucuronosyl-

12 transferase 1A), the main enzyme for conjugation of bilirubin, and Slc10a1, encoding the Na+-

13 taurocholate co-transporting polypeptide, which transports bile acids, are protective against the increased

14 concentration of conjugated bilirubin in hepatocytes after BDL.

15   Cluster 2 (Figure 9B) consists of strong transcript peaks at 6h by the transcriptional regulator Nr0b2

16 (SHP, Figure 6D) and Cyp24a1 (Figure 6B), both probes on the ADME chip. Nr0b2 ($p_{adj} = 1.30E^{-7}$) and

17 Cyp24a1 ($p_{adj} = 9.88E^{-3}$) both have very high ANOVA significance. Other transcripts also show an

18 increase in the initial phase at 6h like for instance the members of cluster 3, but none of them to baseline

19 during disease progression at 18h up to 14d. This makes Nr0b2 and Cyp24a1 interesting targets for the

20 detection of very acute cholestasis.

21   Cluster 3 (Figure 9C) is characterized by an increase in the initial phase up to 18h with subsequent

22 decrease during disease progression up to 14d. The cluster consists of the biochemical factor GLDH, the

23 fibrosis transcript Fn1 (Fibronectin, Figure 6G), and the ADME gene Sulf1a1 (Sulfotransferase 1A1),

24 making this an interesting cluster combining various factor types.

25   Cluster 4 (Figure 4) shows continuous increase starting in the initial phase, lasting throughout disease

26 progression up to 14d. Consequently, members of the cluster are good candidates for continuous disease

27 progression (fibrosis), among the top candidates the bilirubin, CTGF and α-SMA. Cluster 4 is the largest

1   cluster containing 61/90 significant factors of the ANOVA. Most of the classical markers are contained in

2   cluster 4: bilirubin, CTGF α-SMA, S100a4, bile infarcts and BEC. The transcripts contained in c4 come

3   hereby all either the inflammation chip or the fibrosis chip, with exception of Gsta2 and Gstm1 (ADME

4   chip).

5   The top transcript correlation with the cluster mean shows Timp1, followed by Ccr2, and Tgfbr2. It has to

6   be noted that a large number of transcripts has a high correlation.

7   Cluster 5 (Figure E) has a highly similar time course than c4, but only increasing from baseline after

8   around 30h compared to the continuous increase of c4 starting already in the initial phase. Cluster 5

9   contains the classical markers NHC, Kupffer cells and Sirius red. Top correlating transcripts are Gdl2 and

10   Cyp7a1, which shows a strong increase starting around 30h, despite the strong down-regulation in the

11   initial phase (see above). Notably, the interleukins Il2, Il17a (interleukin-17A, Figure 6H) and Il28b

12   (interleukin 28β, Figure 6K) are members of c5, which are secreted proteins, likely to be detectable in

13   blood, thus representing a candidate diagnostic marker. Il17a (interleukin-17A, Figure 6H), plays a pivotal

14   role in cholestatic liver fibrosis in mice by activation of both the KCs and HSCs [19].

15   Cluster 6 (Figure 9F) is characterized by an initial decrease followed by an increase in the late initial

16   phase at 12h up to 2d with subsequent decrease during disease progression at 5d and 14d. Both cluster

17   members Cdh2 (cadherin 2) and Bad1 are part of the fibrosis chip. Cluster 6 shows a similar up and down

18   than c3, but the increase starts later, the decrease starts later and the maximum transcript values are

19   consequently shifted to a later time point.

20   The large majority of histologic parameters, cell count observations, as well as most genes related to

21   inflammation and fibrogenesis increased with disease progression either in a continuous manner starting

22   in the perpetuation phase after ~1 day or at the latest beginning at day 5 (cluster 4 and 5). Based on the

23   strong increase between day 2 and 5 in cluster 5 (but also in cluster 4), we conclude that transition from

24   day 2 to day 5 can be easily monitored, also in clinical practice, as serum bilirubin is among the top

25   correlations of cluster 4, and the interleukins among cluster 5 (Il2, Il17a, Il28b, Figure 6H and K) as well as

26   the growth factors in cluster 4 (Pdgfb, Tgfb1, Tgfb2, Figure 6I, Hgf) and contributors to the extracellular

27   matrix among cluster 4 (Sparc, Col3a1, Col1a1, Figures 6F and E).

1   Interestingly no histological, biochemical or immunostaining based factors were found in the clusters

2   c1, c2 and c6. The transcripts in these clusters provide unique time course information which cannot be

3   captured with the subset of analyzed histological markers, providing crucial information for the initial and

4   perpetuation phase, not attainable via the c4 and c5.

5   **Decision trees for disease progression**

6   We next asked the question, which of the analyzed factors are best suited as markers for particular

7   stages of the disease process. To this end we constructed a decision tree based on the time courses of

8   the normalized clusters (Figure 9). This regression tree allows the prediction of a specific time interval of

9   the disease process based on the combination of time courses of factors from the clusters (Figure 10).

10   The algorithm used for the construction of the decision tree aims at avoiding over-fitting of the data by

11   balancing the number of knots against robustness of the classification tested by cross-validation.

12   Consequently, the decision tree assigns a pattern of cluster to time intervals of disease progression rather

13   than the discrete experimental time points, resulting in 6 time classes. Interestingly, mainly time points in

14   the late initial phase and perpetuation phase (12h, 18h, 30h, 48h) were merged into time classes, whereas

15   the control (0h), early initial phase (6h), and progression phase (5d, 14d) remained almost unchanged.

16   The robustness of the predicted time classes and prediction performance were tested with a leave-one-out

17   approach, resulting in reproducible time classes and good prediction performance on the left out test data

18   (see Supplement 2).

19   The resulting regression tree exploits time course information from cluster c1, c3, c4, and c5, whereas

20   clusters c2 and c6 are not used. Cluster 4 was found to possess the strongest discriminating power

21   performing the important split between early phase after BLD (classes 0h, 6h and 14h with range 0h-

22   21.8h) and the later perpetuation and progression phase (classes 34h, 6d,14d, with range 21.8h-14d): If

23   the mean value of the factors within this cluster is smaller than -0.12 at a measured time point this time

24   point of BLD induced disease is classified as being not larger than 21.8 hours. The more detailed

25   assignment of the respective time interval requires to interrogate additional clusters: The value of c3

26   decides between control and initial phase with subsequent splitting based on c1 in early and late initial

27   phase. Analogue, the value of c5 decides between perpetuation and progression phase with subsequent

1 splitting in early and late progression phase based on c4. Note that the values of cluster 4 appear twice in

2 the decision tree, owing to the monotonous increase after BDL.

3   The predictive performance of the regression tree on the mean cluster data is depicted in Figure 10B

4 (blue), providing the information which experimental time points were classified in which classes. All

5 samples of control, 6h, and 14d are assigned to their respective time classes 0h, 6h and 14d, whereas

6 neighboring time points are combined in classes 14h, 34h and 6d. In addition, we evaluated the decision

7 tree based on a subsets of factors from the time course clusters using either a single factor or two factors

8 randomly chosen from each cluster and their values were used for prediction (Figure 10B, single and

9 double factors). Even with only one factor selected from c1, c3, c4, and 5 surprisingly consistent

10 classifications were achieved. As expected, the quality of predictions was improved by increasing the

11 number of factors. The control, early initial phase and late progression phase can be predicted with high

12 accuracy, whereas the intermediate phase shows more misclassifications in case of the single and double

13 factors.

14   The best performing decision tree based on a single transcript from each cluster (Figure 10A) splits

15 on Col1a1 (Collagen alpha-1(I) chain, c4), Fn1 (Fibronectin, c3), Cyp1a2 (Cytochrome P450 1A2, c1), and

16 Il17a (Interleukin-17A, c5), all important factors involved in disease progression (discussed above). The

17 best performing decision tree based on all factors, i.e. histological, biochemical, immunostaining factors

18 and transcripts, resulted in a highly similar tree, with the single change of using S100a4 instead of Col1a1

19 for c4 splitting, and providing GLDH as alternative factor to Fn1 to perform the c3 splitting. By selecting a

20 subset of good performing transcripts very good prediction performance already on a small subset of

21 factors can be achieved (Figure10B red). Our approach allows the prediction of progression after BDL

22 from an arbitrary subset of measured parameters.

23 **Individual variability**

24   We observed a large variability of many analyzed factors, when comparing individual mice of the

25 same time points. For example at day 5, the infarct area varies between 0.9% and 12% (Figure 2A) and

26 the collagen deposition area measured by Sirius red between 0.8%-5.9% (Figure 4B). Similar high

27 variability can be observed in the proliferative response (BEC Figure 3A, Kupffer Figure 3E) or expression

1   of collagen (Col1a1 Figure 6E, Col3a1 Figure 6F), to name a few. These large variation in parameters

2   during the perpetuation and especially progression phase are an intriguing finding, considering the

3   homogeneity of the experimental system (see also heatmap of time courses in Supplement 2). A possible

4   explanation could be a different individual pace of disease progression due to variations in susceptibility to

5   the damage caused by BDL, for instance due to different genetic background. As a consequence

6   heterogenous time courses develop, with highly affected mice showing strong signs of fibrosis earlier.

7   Another hypothesis are variable routes of disease progression: One route is characterized by a strong

8   increase in necrotic tissue and a weaker activation of HSC and lower expression change of inflammation

9   factors. The other route is characterized by strong activation of fibrogenesis factors, and finally

10 macroscopic collagen tissue. Both routes are similarly connected to loss of liver function, however, the

11 histopathological display is quite different: the former contains a large amount of necrotic tissue while the

12 latter contains large fibrotic tissue. Such alternate disease routes can be of far-reaching importance for an

13 individualized therapy, as obviously the medical interventions avoiding necrosis differ from interventions

14 reducing overshooting fibrosis. The design of the study, which included the sacrifice of mice after a

15 specific time, did not allow to answer whether the alternate developments, shown macroscopically in later

16 time points, can also be observed in earlier time points.

Conclusion

18 The time-resolved analysis of a wide range of parameters in bile duct ligated mice has shown that

19 many factors share the pattern of an increase throughout disease progression (Figure 9). Particularly,

20 pronounced changes were observed during the transition from perpetuation to progression phase 2 to

21 5 days after BDL, characterized by strong increase of parameters, such as Il17A, Il2, Il28b or Il13. This

22 information has strong clinical relevance, as it indicates a switching point, and human homologs of the

23 respective interleukins are top candidates to be used as clinical markers.

24 Our experiments also found previously unknown molecular events which are probably elements of a

25 transcriptional program. For instance, there are strong transcript peaks for SHP (Nr0b2) at 6h, which now

26 need a focused analysis to delineate the chain of molecular interactions causing it and the functional

27 consequences for disease onset and progression.

1   Based on time course correlation analysis we found a distinct number of time course patterns

2   following BDL. Based on these results, we propose a decision tree from a subset of parameters, such as

3   in Figure 10, as a promising tool for assessment of disease progression. Therefore, as next step,

4   suitability of the parameters selected from the mouse model need to be confirmed for human patients as

5   in [69].

6   Many of the measured parameters display a large variability, which maybe the major reason why

7   translation of a set of such parameters into diagnostic approaches have not proven sufficient robustness

8   for valid predictions in human patients with chronic liver diseases. The here suggested approach of

9   pooling information from factors falling in the same time course classes could be a possible solution for

10 more robust predictors.

11 Taken together, the detailed time-resolved transcriptional profiling of liver homogenates following BDL

12 revealed a coordinated response after surgery, resulting in an upregulation of an inflammatory response

13 along with activation of metabolically active genes, which can be explained by physiological recovery and

14 adaptation of the mice to the bile acid exposure.

15 Availability of supporting data

16 The data sets supporting the results of this article are included within the article and its additional files.

17 Dataset S1. Measurements in bile duct ligated mice.

18 Dataset S2. Statistical analysis.

19 List of Abbreviations

20 CTGF: connective tissue growth factor; BDL: bile duct ligation; SHP: small heterodimer partner; CLD:

21 chronic liver disease; HSC: hepatic stellate cells; KC: Kupffer cells; ADME: absorption, distribution,

22 metabolism, and excretion; ALT: alanine aminotransferase; GLDH: glutamate dehydrogenase; EDTA:

23 ethylenediaminetetraacetic acid; H&E: hematoxylin and eosin;

24

25 Competing Interests

1   The authors declare that they have no conflicts of interest.

2

3   **Author Contributions**

4   Conceived and designed the experiments: KA AH MT IM BV SD. Performed the experiments: KA MT IM

5   HW. Analyzed the data: KA MK AH MT IM ME HW HH UMZ BV SD. Wrote the paper: KA MK AH MT ME

6   HW HH UMZ BV SD. All authors read and approved the final manuscript.

7   **Acknowledgments**

15

References

1. Rockey DC. Translating an understanding of the pathogenesis of hepatic fibrosis to novel therapies. Clin Gastroenterol Hepatol. 2013;11:224-31.

2. Diehl AM, Chute J. Underlying potential: cellular and molecular determinants of adult liver repair. J Clin Invest. 2013;123:1858-60.

3. Schuppan D, Kim YO. Evolving therapies for liver fibrosis. J Clin Invest. 2013;123:1887-1901.

4. An G, Mi Q, Dutta-Moscatom J, Vodovotz Y. Agent-based models in translational systems biology. Wiley Interdiscip Rev Syst Biol Med. 2009; 1:159-71.

5. Georgiev P, Jochum W, Heinrich S, Jang JH, Nocito A, Dahm F, et al. Characterization of time-related changes after experimental bile duct ligation. Br J Surg. 2008;95:646-56.

6. Spurgeon SL, Jones RC, Ramakrishnan R. High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. PLoS One. 2008;3:e1662.

7. Geerts AM, Vanheule E, Praet M, Van Vlierberghe H, De Vos M, Colle I. Comparison of three research models of portal hypertension in mice: macroscopic, histological and portal pressure evaluation. Int J Exp Pathol. 2008;89:251-63.

8. Kisseleva T, Brenner DA. Anti-fibrogenic strategies and the regression of fibrosis. Best Pract Res Clin Gastroenterol. 2011;25:305-17.

9. Mederacke I, Hsu CC, Troeger JS, Huebener P, Mu X, Dapito DH, et al. Fate tracing reveals hepatic stellate cells as dominant contributors to liver fibrosis independent of its aetiology. Nat Commun. 2013;4:2823.

10. Holt AP, Salmon M, Buckley CD, Adams DH. Immune interactions in hepatic fibrosis. Clin Liver Dis. 2008;12:861-82.

11. Wasmuth HE, Lammert F, Zaldivar MM, Weiskirchen R, Hellerbrand C, Scholten D, et al. Antifibrotic effects of CXCL9 and its receptor CXCR3 in livers of mice and humans. Gastroenterology. 2009;137:309-19.

12. Heinrichs D, Berres ML, Nellen A, Fischer P, Scholten D, Trautwein C, et al. The chemokine CCL3 promotes experimental liver fibrosis in mice. PLoS One. 2013;8:e66106.

1   13. Leask A, Abraham DJ. All in the CCN family: essential matricellular signaling modulators emerge
2   from the bunker. J Cell Sci. 2006;119:4803-10.

3   14. Gressner AM, Weiskirchen R, Breitkopf K, Dooley S. Roles of TGF-beta in hepatic fibrosis. Front
4   Biosci. 2002;7:d793-807.

5   15. Gupta S, Stravitz RT, Dent P, Hylemon PB. Down-regulation of cholesterol 7alpha-hydroxylase
6   (CYP7A1) gene expression by bile acids in primary rat hepatocytes is mediated by the c-Jun N-
7   terminal kinase pathway. J Biol Chem. 2001;276:15816-22.

8   16. Russell DW. Nuclear orphan receptors control cholesterol catabolism. Cell. 1999;97:539-42.

9   17. Schacter BA, Joseph E, Firneisz G. Effect of cholestasis produced by bile duct ligation on hepatic
10   heme and hemoprotein metabolism in rats. Gastroenterology. 1983;84:227-35.

11 18. Hattori T, Ohoka N, Hayashi H, Onozaki K. C/EBP homologous protein (CHOP) up-regulates IL-6
12   transcription by trapping negative regulating NF-IL6 isoform. FEBS Lett. 2003;541:33-9.

13 19. Hara M, Kono H, Furuya S, Hirayama K, Tsuchiya M, Fujii H. Interleukin-17A plays a pivotal role in
14   cholestatic liver fibrosis in mice. J Surg Res. 2013;183:574-82.

15 20. Chen YX, Weng ZH, Zhang SL. Notch3 regulates the activation of hepatic stellate cells. World J
16   Gastroenterol. 2012;18:1397-403.

17 21. Sancho-Bru P, Altamirano J, Rodrigo-Torres D, Coll M, Millán C, José Lozano J, et al. Liver
18   progenitor cell markers correlate with liver damage and predict short-term mortality in patients with
19   alcoholic hepatitis. Hepatology. 2012;55:1931-41.

20 22. Fujii T, Fuchs BC, Yamada S, Lauwers GY, Kulu Y, Goodwin JM, et al. Mouse model of carbon
21   tetrachloride induced liver fibrosis: Histopathological changes and expression of CD133 and
22   epidermal growth factor. BMC Gastroenterol. 2010.10:79.

23 23. Reister S, Kordes C, Sawitza I, Häussinger D. The epigenetic regulation of stem cell factors in
24   hepatic stellate cells. Stem Cells Dev. 2011;20:1687-99.

25 24. Page S, Birerdinc A, Estep M, Stepanova M, Afendy A, Petricoin E, et al. Knowledge-based
26   identification of soluble biomarkers: hepatic fibrosis in NAFLD as an example. PLoS One.
27   2013;8:e56009.

25. Estep JM, Baranova A, Hossain N, Elariny H, Ankrah K, Afendy A, et al. Expression of cytokine signaling genes in morbidly obese patients with non-alcoholic steatohepatitis and hepatic fibrosis. Obes Surg. 2009;19:617-24.

26. Huang W, Zhang J, Chua SS, Qatanani M, Han Y, Granata R, et al. Induction of bilirubin clearance by the constitutive androstane receptor (CAR). Proc Natl Acad Sci U S A. 2003;100:4156-61.

27. Floreani M, De Martin S, Gabbia D, Barbierato M, Nassi A, Mescoli C, et al. Severe liver cirrhosis markedly reduces AhR-mediated induction of cytochrome P450 in rats by decreasing the transcription of target genes. PLoS One, 2013;8:e61983.

28. Bai H, Zhang N, Xu Y, Chen Q, Khan M, Potter JJ, et al. Yes-associated protein regulates the hepatic response after bile duct ligation. Hepatology. 2012;56:1097-107.

29. Sookoian S, Gianotti TF, Rosselli MS, Burgueño AL, Castaño GO, Pirola CJ. Liver transcriptional profile of atherosclerosis-related genes in human nonalcoholic fatty liver disease. Atherosclerosis. 2011;218:378-85.

30. Inagaki Y, Okazaki I. Emerging insights into Transforming growth factor beta Smad signal in hepatic fibrogenesis. Gut. 2007;56:284-92.

31. Rachmawati H, Beljaars L, Reker-Smit C, Van Loenen-Weemaes AM, Hagens WI, Meijer DK, et al. Pharmacokinetic and biodistribution profile of recombinant human interleukin-10 following intravenous administration in rats with extensive liver fibrosis. Pharm Res. 2004;21:2072-8.

32. Eipel C, Menschikow E, Sigal M, Kuhla A, Abshagen K, Vollmar B. Hepatoprotection in bile duct ligated mice mediated by darbepoetin-α is not caused by changes in hepatobiliary transporter expression. Int J Clin Exp Pathol. 2013;6:80-90.

33. Paradis V, Dargere D, Vidaud M, De Gouville AC, Huet S, Martinez V, et al. Expression of connective tissue growth factor in experimental rat and human liver fibrosis. Hepatology. 1999;30:968-76.

34. Sedlaczek N, Jia JD, Bauer M, Herbst H, Ruehl M, Hahn EG, et al. Proliferating bile duct epithelial cells are a major source of connective tissue growth factor in rat biliary fibrosis. Am J Pathol. 2001;158:1239-44.

35. Rachfal AW, Brigstock DR. Connective tissue growth factor (CTGF/CCN2) in hepatic fibrosis. Hepatol Res. 2003;26:1-9.

36. Dendooven A, Gerritsen KG, Nguyen TQ, Kok RJ, Goldschmeding R. Connective tissue growth factor (CTGF/CCN2) ELISA: a novel tool for monitoring fibrosis. Biomarkers. 2011;16:289-301.

37. Liu Y, Liu H, Meyer C, Li J, Nadalin S, Königsrainer A, et al. Transforming growth factor-β (TGF-β)-mediated connective tissue growth factor (CTGF) expression in hepatic stellate cells requires Stat3 signaling activation. J Biol Chem. 2013;288:30708-19.

38. Schierwagen R, Leeming DJ, Klein S, Granzow M, Nielsen MJ, Sauerbruch T, et al. Serum markers of the extracellular matrix remodeling reflect antifibrotic therapy in bile-duct ligated rats. Front Physiol. 2013;4:195.

39. Wang Q, Usinger W, Nichols B, Gray J, Xu L, Seeley TW, et al. Cooperative interaction of CTGF and TGF-β in animal models of fibrotic disease. Fibrogenesis Tissue Repair. 2011;4:4.

40. Malizia G, Brunt EM, Peters MG, Rizzo A, Broekelmann TJ, McDonald JA. Growth factor and procollagen type I gene expression in human liver disease. Gastroenterology. 1995;108:145-56.

41. Yang L, Inokuchi S, Roh YS, Song J, Loomba R, Park EJ, et al. Transforming growth factor-β signaling in hepatocytes promotes hepatic fibrosis and carcinogenesis in mice with hepatocyte-specific deletion of TAK1. Gastroenterology. 2013;144:1042-54.

42. Qian J, Niu M, Zhai X, Zhou Q, Zhou Y. β-Catenin pathway is required for TGF-β1 inhibition of PPARγ expression in cultured hepatic stellate cells. Pharmacol Res. 2012;66:219-25.

43. Zeisberg M, Yang C, Martino M, Duncan MB, Rieder F, Tanjore H, et al. Fibroblasts derive from hepatocytes in liver fibrosis via epithelial to mesenchymal transition. J Biol Chem. 2007;282:23337-47.

44. Robertson H, Kirby JA, Yip WW, Jones DE, Burt AD. Biliary epithelial-mesenchymal transition in posttransplantation recurrence of primary biliary cirrhosis. Hepatology. 2007;45:977-81.

45. Österreicher CH, Penz-Österreicher M, Grivennikov SI, Guma M, Koltsova EK, Datz C, et al. Fibroblast-specific protein 1 identifies an inflammatory subpopulation of macrophages in the liver. Proc Natl Acad Sci U S A. 2011;108:308-13.

46. Baclig MO, Alvarez MR, Lozada XM, Mapua CA, Lozano-Kühne JP, Dimamay MP, et al. Association of glutathione S-transferase T1 and M1 genotypes with chronic liver diseases among Filipinos. Int J Mol Epidemiol Genet. 2012;3:153-9.

47. Znoyko I, Sohara N, Spicer SS, Trojanowska M, Reuben A. Expression of oncostatin M and its receptors in normal and cirrhotic human liver. J Hepatol. 2005;43:893-900.

48. Liang H, Block TM, Wang M, Nefsky B, Long R, Hafner J, et al. Interleukin-6 and oncostatin M are elevated in liver disease in conjunction with candidate hepatocellular carcinoma biomarker GP73. Cancer Biomark. 2012;11:161-71.

49. Barashi N, Weiss ID, Wald O, Wald H, Beider K, Abraham M, et al. Inflammation-induced hepatocellular carcinoma is dependent on CCR5 in mice. Hepatology. 2013;58:1021-30.

50. Seki E, De Minicis S, Gwak GY, Kluwe J, Inokuchi S, Bursill CA, et al. CCR1 and CCR5 promote hepatic fibrosis in mice. J Clin Invest. 2009;119:1858-70.

51. Sudo K, Yamada Y, Moriwaki H, Saito K, Seishima M. Lack of tumor necrosis factor receptor type 1 inhibits liver fibrosis induced by carbon tetrachloride in mice. Cytokine. 2005;29:236-44.

52. Zimmermann HW, Seidler S, Gassler N, Nattermann J, Luedde T, Trautwein C, et al. Interleukin-8 is activated in patients with chronic liver diseases and associated with hepatic macrophage accumulation in human liver fibrosis. PLoS One. 2011;6:e21381.

53. Smalling RL, Delker DA, Zhang Y, Nieto N, McGuiness MS, Liu S, et al. Genome-wide transcriptome analysis identifies novel gene signatures implicated in human chronic liver disease. Am J Physiol Gastrointest Liver Physiol. 2013;305:G364-74.

54. Zhang Y, Xu N, Xu J, Kong B, Copple B, Guo GL, et al. E2F1 is a novel fibrogenic gene that regulates cholestatic liver fibrosis through the Egr-1/SHP/EID1 network. Hepatology. 2014;60:919-30.

55. Horvath E, Lakatos P, Balla B, Kósa JP, Tóbiás B, Jozilan H, et al. Marked increase of CYP24A1 mRNA level in hepatocellular carcinoma cell lines following vitamin D administration. Anticancer Res. 2012;32:4791-6.

56. Fosang AJ, Neame PJ, Hardingham TE, Murphy G, Hamilton JA. Cleavage of cartilage proteoglycan between G1 and G2 domains by stromelysins. J Biol Chem. 1991;266:15579-82.

57. Garcia-Irigoyen O, Carotti S, Latasa MU, Uriarte I, Fernández-Barrena MG, Elizalde M, et al. Matrix metalloproteinase-10 expression is induced during hepatic injury and plays a fundamental role in liver tissue repair. Liver Int. 2014;34:e257-70.

58. Krampert M, Bloch W, Sasaki T, Bugnon P, Rülicke T, Wolf E, et al. Activities of the matrix metalloproteinase stromelysin-2 (MMP-10) in matrix degradation and keratinocyte organization in wounded skin. Mol Biol Cell. 2004;15:5242-54.

59. LaFramboise WA, Bombach KL, Pogozelski AR, Cullen RF, Muha N, Lyons-Weiler J, et al. Hepatic gene expression response to acute indomethacin exposure. Mol Diagn Ther. 2006;10:187-96.

60. Fukushima S, Okuno H, Shibatani N, Nakahashi Y, Seki T, Okazaki K. Effect of biliary obstruction and internal biliary drainage on hepatic cytochrome P450 isozymes in rats. World J Gastroenterol. 2008;14:2556-60.

61. Trombetta-Esilva J, Bradshaw AD. The function of SPARC as a mediator of fibrosis. Open Rheumatol J. 2012;6:146-155.

62. Hayashi N, Kakimuma T, Soma Y, Grotendorst GR, Tamaki K, Harada M, et al. Connective tissue growth factor is directly related to liver fibrosis. Hepatogastroenterology. 2002;49:133-5.

63. Colak Y, Senates E, Coskunpinar E, Oltulu YM, Zemheri E, Ozturk O, et al. Concentrations of connective tissue growth factor in patients with nonalcoholic fatty liver disease: association with liver fibrosis. Dis Markers. 2012;33:77-83.

64. Gressner OA, Fang M, Li H, Lu LG, Gressner AM, Gao CF. Connective tissue growth factor (CTGF/CCN2) in serum is an indicator of fibrogenic progression and malignant transformation in patients with chronic hepatitis B infection. Clin Chim Acta. 2013;421:126-31.

65. Aparicio-Vergara M, Hommelberg PP, Schreurs M, Gruben N, Stienstra R, Shiri-Sverdlov R, et al. Tumor necrosis factor receptor 1 gain-of-function mutation aggravates nonalcoholic fatty liver disease but does not cause insulin resistance in a murine model. Hepatology. 2013;57:566-76.

66. Tarrats N, Moles A, Morales A, García-Ruiz C, Fernández-Checa JC, Marí M. Critical role of tumor necrosis factor receptor 1, but not 2, in hepatic stellate cell proliferation, extracellular matrix remodeling, and liver fibrogenesis. Hepatology. 2011;54:319-27.

1   67. Khan AJ, Choudhuri G, Husain Q, Parmar D. Polymorphism in glutathione-S-transferases: a risk

2   factor in alcoholic liver cirrhosis. Drug Alcohol Depend. 2009;101:183-90.

3   68. Noureddin M, Wright EC, Alter HJ, Clark S, Thomas E, Chen R, et al. Association of IL28B genotype

4   with fibrosis progression and clinical outcomes in patients with chronic hepatitis C: a longitudinal

5   analysis. Hepatology. 2013;58:1548-57.

6   69. Gadd VL, Skoien R, Powell EE, Fagan KJ, Winterford C, Horsfall L, et al. The portal inflammatory

7   infiltrate and ductular reaction in human nonalcoholic fatty liver disease. Hepatology. 2014;59:1393-

8   1405.

9

1 Tables

2 **Table 1.** Analysis of systemic blood cell count of sham-operated mice (S) and mice underwent BDL.

3 Values are given as means ± SEM.

	erythrocytes	platelets	leukocytes	hemoglobin	hematocrit
S	8.4±0.1	1177±60	7.5±0.3	7.9±0.1	44.8±0.7
6h	8.1±0.1	1061±39	4.2±0.5	7.7±0.1	42.8±0.7
12h	8.2±0.1	1036±47	4.5±0.4	7.7±0.1	43.0±0.6
18h	8.7±0.3	856±110	4.1±0.2	8.1±0.3	45.6±1.9
30h	8.5±0.5	1071±100	5.8±0.9	7.9±0.5	44.9±2.8
2d	8.7±0.2	1117±65	4.7±1.3	6.5±1.7	45.9±0.9
5d	8.7±0.3	1295±107	7.6±1.2	7.8±0.3	46.5±1.6
14d	7.6±1.4	1362±58	7.4±1.1	6.6±0.2	38.4±1.3

4

1   **Figure legends**

2   **Figure 1: Analysis of liver injury and-function.** Plasma activities of alanine aminotransferase (ALT) (**A**)

3   and glutamate dehydrogenase (GLDH) (**B**) and concentrations of plasma bilirubin (**C**) and albumin (**D**) at

4   multiple time points after BDL. Values are given in means ± SEM of five independent experiments per time

5   point.

6

7   **Figure 2: Quantification of bile infarcts <u>in H&E-stained liver sections</u> at multiple time points after**

8   **BDL (A).** Values are given in means ± SEM of five independent experiments per time point.

9   Representative H&E stainings of paraffin-embedded liver sections for each time point after BDL (**B**; arrows

10 indicate bile lakes; magnification x10) with higher magnifications (x40) in (**C**), displaying cellular infiltrates

11 (asterisk) and formation of artificial bile ductules (arrowhead).

12

13 **Figure 3: Analysis of the proliferative and cellular response at multiple time points after BDL.**

14 Quantitative immunohistochemical analysis of BrdU-positive biliary epithelial cells (**A**), liver cells positive

15 for α-SMA (**B**) and S100a4 (**C**), BrdU-positive hepatocytes (**D**) and Kupffer cells (**E**) and CTGF-positive

16 cells (**F**). Values are given in means ± SEM of five independent experiments per time point.

17 Corresponding representative immunohistochemical stainings are shown in the right panel (magnifications

18 x40).

19

20 **Figure 4: Analysis of proliferation and extracellular matrix accumulation.** mRNA quantification of the

21 proliferation marker Ki67 (**A**) by Fluidigm real-time PCR. Values are given in means ± SEM of five

22 independent experiments per time point. Quantitative analysis <u>of extracellular matrix deposition</u> (**B**) and

23 representative histological images (**C**; magnification x10) of Sirius red-positive areas at multiple time

24 points after BDL. Values are given in means ± SEM of five independent experiments per time point.

25

26 **Figure 5: Heat maps displaying gene expression pattern at multiple time points after BDL.** Gene

27 expression relative to Gapdh gene, obtained from Fluidigm qPCR, are shown as fold changes to sham

34

1   mice (0 h) and are displayed in log2 scale. Red colour indicates up-regulation (2), blue down-regulation (-

2   2) and white transcription levels about 1. (**A**) selected ADME genes, (**B**) selected fibrogenesis genes, and

3   (**C**) selected cytokine genes.

4

5   **Figure 6: mRNA quantification of different <u>selected </u>genes by Fluidigm real-time PCR <u>displayed in</u>**

6   **<u>log2 scale</u>.** (**A**) Cyp1a2, (**B**) Cyp24a1, (**C**) Gstm1, (**D**) Nr0b2, (**E**) Col1α1, (**F**) Col3α1, (**G**), Fn1, (**H**)

7   Mmp10, (**I**) Tgfb2, (**J**) Il2, (**K**) Il28b, (**L**) Tnfrsf1a. Values are given in means ± SEM of five independent

8   experiments per time point.

9

10   **Figure 7: Correlation matrix of factors.** Correlation matrix (YS3) of the subset of significantly changed

11   factors determined by ANOVA. Positive correlation depicted in blue, negative correlation in red according

12   to color key. Side dendrogram depicts the hierarchical clustering results with the 6 time course clusters

13   marked in the color sidebar (see Figure 9 for the time courses corresponding to clusters c1-c6).

14   Histological factors marked with H, immunostainings with A, and biochemical factors with B.

15

16   **Figure 8: Histological (H), biochemical (B), and immunostaining (A) correlations.** All correlations are

17   YS3 correlation between the respective factors with numerical values provided in Supplement 2. (**A**)

18   Correlation between histological, biochemical and antibody factors and gene transcripts. Only genes with

19   at least one YS3 correlation of abs(YS3)>=0.6 are shown. Positive correlation in blue, negative correlation

20   in red analog to Figure 7, with the area of circles corresponding to the correlation coefficients. (**B**)

21   Correlation between histological, biochemical and antibody factors. (**C**) Highest absolute correlations

22   between histological, biochemical and antibody factors and all other ANOVA filtered factors. Data sorted

23   from left to right by absolute value of correlation. Color and size of the filled pie corresponds to the

24   respective correlation value.

25

26   **Figure 9: Time course clusters in BDL.** Time course clusters based on YS3 correlation with hierarchical

27   clustering based on complete linkage corresponding to clusters in Figure 8. Mean cluster time course

28   (averaged over all factors and repeats) depicted in blue, normalized factor representatives of cluster in

1   grey. Shaded grey area corresponding to SD of all mean factor time courses. Top correlations based on

2   YS3 between the factors in the cluster and the mean cluster time course are shown above the time

3   course. Histological factors marked with H, immunostainings with A, and biochemical factors with B.

4   Cluster members are fully enumerated in the figure for all classes with exception of cluster 4 with the

5   members and respective correlation to mean cluster time course being: Timp1 (0.94), bilirubin (B 0.92),

6   Ccr2 (0.92), CTGF (A 0.91), Tgfbr2 (0.89), α-SMA (A 0.89), Ccl5 (0.88), Tgfb1 (0.88), Ccl3 (0.87), Tnc

7   (0.87), Cd14 (0.87), Ccl2 (0.86), Cd86 (0.86), Pdgfb (0.86), Col1a1 (0.86), Cxcl3 (0.86), Ccl4 (0.85), Cxcl5

8   (0.85), Il10ra (0.85), Col3a1 (0.85), Il10rb (0.84), Ccl7 (0.82), Cd69 (0.82), Ifnar1 (0.82), Tnf (0.82), Osm

9   (0.81), Sparc (0.8), Il6 (0.8), Tnfrsf1b (0.8), Cxcr2 (0.78), Il1b (0.78), Timp2 (0.77), Ifnar2 (0.77), Ccr5

10   (0.77), Il10 (0.76), Osmr (0.75), Gsta2 (0.74), Il4 (0.71), Ifng (0.71), Ccl8 (0.71), Hgf (0.7), Bak1 (0.7),

11   Mrc1 (0.69), Tgfb2 (0.69), Ccr3 (0.68), Actb (0.68), S100a4 (A 0.66), Il13 (0.66), Met (0.66), bile infarcts

12   (H 0.65), Il6st (0.63), Tnfrsf1a (0.63), Mki67 (0.62), Birc5 (0.6), Ctgf (0.58), BEC (H 0.56), Bax (0.56),

13   Notch1 (0.54), Cxcr1 (0.51), Gstm1 (0.45), Cdh1 (0.42)

14

15   **Figure 10: Decision tree for disease progression.** (**A**) Regression tree for the prediction of time phases

16   after BDL based on mean cluster time courses. The best gene representatives from the clusters are

17   depicted above the decision rules. The regression tree classifies the data in 6 time classes 0h, 6h, 14h,

18   24h, 6d, 14d with information on mean time value, range, and number and percentage of samples falling

19   given in the respective class. The best tree based on genes, histological, biochemical and antibody factors

20   is highly similar to the best gene tree, with the single change of using S100a4 instead of Col1a1 for the

21   decision on cluster c4 and allowing GLDH as equally good alternative to Fn1 in c3. (**B**) Predictive

22   performance of decision tree. The predictive performance of the regression tree was evaluated using all

23   single factor combinations from the individual clusters (white), a random sample (N=10000) of two factors

24   from each cluster (gray), the best gene representative tree (red), and the mean cluster data (blue,

25   trainings data).

26

1   **Figure 11: Outline of the disease process.** Each box is dedicated to a specific disease aspect (first line)

2   which is represented by a commonly known marker (second line) or several markers. Below (in small font)

3   the genes are shown whose expression is correlated to the factor above.