Pathobiochemical signatures of cholestatic liver disease in bile duct ligated mice

Kerstin Abshagen[1¶*], Matthias König[2¶], Andreas Hoppe[2], Isabell Müller[1], Matthias Ebert[4], Honglei Weng[4], Herrmann-Georg Holzhütter[2], Ulrich M. Zanger[3], Johannes Bode[5], Brigitte Vollmar[1], Maria Thomas[3], Steven Dooley[4]

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

[2]Institute for Biochemistry, Computational Systems Biochemistry Group, Charité University Medicine Berlin, 10117 Berlin, Germany

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

[4]Department of Medicine II, Section Molecular Hepatology, Medical Faculty Mannheim, Heidelberg University, 68167 Mannheim, Germany

[5]Department for Gastroenterology, Hepatology and Infectiology, Heinrich-Heine University of Düsseldorf, 40225 Düsseldorf

[¶]These authors contributed equally to this work

*Correspondence should be addressed to:

Kerstin Abshagen, PhD

Institute for Experimental Surgery

Rostock University Medical Center

Schillingallee 69a

18057 Rostock

Germany

phone: +49 381 494 2503

fax: +49 381 494 2502

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

Email addresses:

matthias.koenig@charite.de

andreas.hoppe@charite.de

Maria.Thomas@ikp-stuttgart.de

isabell-mueller86@web.de

matthias.ebert@umm.de

honglei.weng@medma.uni-heidelberg.de

hergo@charite.de

Uli.Zanger@ikp-stuttgart.de

Johannes.Bode@med.uni-duesseldorf.de

brigitte.vollmar@med.uni-rostock.de

Steven.Dooley@medma.uni-heidelberg.de

Running title: Cellular and molecular signatures of cholestasis

Abstract

Background: Disrupted bile secretion leads to liver damage characterized by inflammation, fibrosis, eventually cirrhosis, and hepatocellular cancer. As obstructive cholestasis often progresses insidiously, markers for the diagnosis and staging of the disease are urgently needed. To this end we compiled a comprehensive data set of serum markers, histological parameters and transcript profiles at 8 time points after bile duct ligation (BDL) in mice.

Results: Statistical analysis of the more than 6,000 data points revealed distinct temporal phases of disease. Putting time-courses of biochemical, histochemical and mRNA transcripts (=factors) together, time-series correlation analysis defined 6 distinct clusters related to different phases of disease progression. The count of CTGF-positive cells provided the most reliable overall measure for disease progression at the histological level, bilirubin at biochemical level, and metalloproteinase inhibitor 1 (Timp1) on transcript level. Prominent molecular events exhibited by strong transcript peaks are found for the transcriptional regulator Nr0b2 (Shp) and 1,25-dihydroxyvitamin D(3) 24-hydroxylase (Cyp24a1) at 6 h. Based on these clusters we constructed a decision tree to identify factor combinations that can be used as markers for specific time intervals of disease progression. Best prediction of onset of disease is achieved by fibronectin (Fn1), the initial phase by (Cyp1a2), transition to the perpetuation phase by collagen α-1 I (Col1a1), and transition to the progression phase by interleukin 17-a (Il17a), with early and late progression separated by Col1a1. Notably, these predictions remained stable even for randomly chosen small sub-sets of factors from the clusters.

Conclusion: Our explorative study demonstrate the existence of a reproducible and well-coordinated temporal gene-expression program underlying disease progression in BLD cholestasis. To further elucidate the regulatory network behind this expression program additional studies are needed which have to include knowledge on transcription factor activation due to accumulation of bile salts and shared transcriptions factor binding motifs of genes belonging to the same transcript cluster.

Keywords: liver injury, mouse, systems biology, fibrosis, cell proliferation, bile duct ligation, cholestasis, morphological profiling, VirtualLiver Network

Background

Cholestatic liver diseases are caused by an impaired flow of the bile from the liver to the duodenum. The main fraction of organic molecules in the bile are bile salts which are strong detergents required for the extraction of lipids from the apical membrane of hepatocytes into the bile fluid and for the emulgation of lipids in the gut. Moreover, the bile fluid, contains numerous endogenous end products (e.g. bilirubin) and potentially toxic compounds cleared by the liver. Hence, accumulation of bile compounds in the cholestatic liver causes an unspecific cell damage that initiates a cascade of inflammatory events: Necrosis of hepatocytes and cholangiocytes, activation of macrophages releasing of pro-inflammatory cytokines and chemokines, neutrophil infiltration, cholangiocyte and hepatocyte proliferation, stellate cell activation with progressive fibrosis causing secondary biliary cirrhosis, ultimately leading to liver failure {Wang2005}.

Multiple pathologies may represent the primary trigger of impaired bile flow as, for example, defects in the export of the bile from hepatocytes to the extracellular space (hepatocellular cholestasis) or obstruction of bile ducts by gall stones or local tumor impingement (extrahepatic cholestasis) {Woolbright2013}. Among the most common causes of cholestatic liver disease in the adult population are primary biliary cirrhosis (PBC) and primary sclerosing cholangitis (PSC), while biliary atresia and Alagille syndrome are commonly found in the pediatric population {Lindblad1977,Trautner1998, Trauner2004, Wang2013, Nguyen2014}.

One of the main causes of cholestasis is biliary tract obstruction, which is a common mechanism of hepatic injury in a variety of clinical settings, including obstructing neoplasms, post-operative bile duct injury, biliary atresia, sclerosing cholangitis, and primary biliary cirrhosis. The corresponding experimental model to induce obstructive cholestatic injury in mice and rats is surgical bile duct ligation (BDL) {Tag2015, Georgiev2008}, which results in stereotypical histopathological changes of cholestasis. The BDL experimental model has been well described and evaluated in rats and mice and is widely used to study cholestatic liver injury and fibrogenesis {Huss2010}.

Chronic liver diseases (CLD) like cholestasis represent with characteristic temporal morphologic, biochemical and molecular changes in liver and serum. For instance, in the BDL model of cholestasis an early phase of acute hepatocyte injury is followed by a proliferative response, up-regulation of pro-inflammatory cytokines and pro-fibrotic metabolic enzymes, and subsequent development of fibrosis after

around 7 days {Georgiev2008, Huss2010}. Such alterations can be exploited to identify biomarkers reporting on specific stages of disease progression and severity of the disease. Hitherto semi-quantitative morphological scoring is the standard technique for grading in CLD. However, nowadays the availability of high-throughput technologies enables to flank the histological assessment of injured tissue with a comprehensive molecular profiling of gene transcripts, gene products (proteins) and metabolites. Such analyses not only provide a more detailed characterization and thus more refined staging of disease progression, they also lead to a deeper understanding of the molecular networks governing the histological and physical alterations observed at higher scales of tissue organization. Extracting from high-throughput data relevant information for clinical translation is still in its beginnings. The identification of key processes triggering the transition between different phases of disease progression requires mathematical analyses which take into account multiple parallel processes, process dynamics, and experimental data from different levels of cellular organization [4].

In this study we explicitly wanted to systemize existing and newly acquired knowledge on the molecular biomarkers of cholestasis and to analyze the time-dependent disease progression following BDL in mice in a comprehensive manner. Our approach sets itself apart from existing studies which either describe the time course of a limited number selected parameters after BDL {Georgiev2008, Huss2010} or provide gene expression profiling for a limited number of time points, thereby missing the acute damage after BDL in the first 24h and long-term effects after 7 days {Tanaka2006}. None of these preceding studies providing predictive models based on the acquired time course data. Therefore, a central aim of our study was to identify molecular markers for the temporal progression of BDL cholestasis by correlating high-accuracy image data and transcriptional profiles of preselected targets with pathobiochemical markers, whereby all data were obtained at different time points after onset of obstructive cholestasis. We collected from 8 different time points after BDL more than 6,000 experimental data points, comprising immunohistochemistry, biochemistry and molecular profiling measures. Statistical methods were applied to unravel robust interrelations in this large-scale data set and find clusters of parameters corresponding to characteristic time profiles in cholestatic injury induced by BDL. We correlated level and timing of pathophysiological events with transcriptional changes in order to define molecular markers, and developed predictive decision trees

that allow assessment of the different disease phases occurring during the development of cholestasis based on a subset of biomarkers.

Materials and Methods

Ethic Statement

All experiments were approved by the local government Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei Mecklenburg-Voprommern (LALLF M-V/TSD/7221.3-1.2-049/09) and performed in accordance with the German legislation on protection of animals and the National Institutes of ‘‘Health Guide for the Care and Use of Laboratory Animals’’ (Institute of Laboratory Animal Resources, National Research Council; NIH publication 86-23 revised 1985).

Mice

Male C57BL/6J (Charles River Laboratories, Sulzfeld, Germany) at 8–10 weeks of age with a body weight of 23–26 g were kept on water and standard laboratory chow ad libitum.

Surgical procedure and experimental groups

Mice were anesthetized by breathing isoflurane (1.5 vol%). BDL was performed after midline laparotomy. The common bile duct was ligated three times with 5-0 silk and transected between the two most distal ligations. Sham operation was performed similarly, except for ligation and transection of the bile duct (0 h, n=5). All surgical procedures were performed under aseptic conditions. Animals were allowed to recover from anesthesia and surgery under a red warming lamp and were held in single cages until subsequent experiments followed at postoperative hours 6, 12, 18 and 30 (n=5 animals per time point), and at 2, 5 and 14 days after BDL (n=5 animals per time point). Sham-operated animals without BDL served as controls (n=5). To analyze the regenerative response in regard to proliferation of different cell types, 5-bromo-2-deoxyuridine (BrdU; 50 mg/kg bw ip) was injected 1 h prior to harvest of liver tissue. BrdU incorporation into DNA was analyzed by immunohistochemistry. To obtain blood and liver samples, mice were killed at the indicated time points. Liver tissue was partially embedded in paraffin for morphology analysis and snap frozen for molecular biology and biochemistry analyses. In addition, liver tissue served for the parallel Taqman qRT-PCR using microfluidic Fluidigm Biomark™ platform (Fluidigm, CA, USA).

Hematological measurements and plasma enzyme levels

Animals were anesthetized and exsanguinated by puncture of the vena cava inferior. Red blood cell and blood platelets count, hemoglobin, and hematocrit were assessed with an automated cell counter (Sysmex KX-21, Sysmex). Plasma activities of alanine aminotransferase (ALT), aspartate aminotransferase (AST) and glutamate dehydrogenase (GLDH) were measured spectrophoto-metrically.

Assays

EDTA plasma served for the analysis of albumin as a parameter of liver function, which was determined with a commercially available enzyme-linked immunosorbent assay kit in accordance with the manufacturer's instructions (Assaypro, MO, USA).

Histopathology and image analysis

Liver tissue samples were fixed in formalin for 2 to 3 days and embedded in paraffin. 5 µm sections were stained with hematoxylin and eosin (H&E) for routine examination and quantification of bile infarcts. Sirius red staining served for quantification of collagen deposition. All samples from a series of experiments were stained simultaneously and evaluated in a blinded manner. For histomorphometric analysis, images of 20 random low power fields (x10 magnification, Olympus BX 51, Hamburg, Germany) were acquired with a Color View II FW camera (Color View, Munich, Germany) and evaluated using an image analysis system (Adobe Photoshop, Adobe Systems, Uxbridge, UK). Fibrosis deposition was quantified as a percentage of Sirius red stained area compared with the total section area. The surfaces of large centrilobular veins and large portal tracts were excluded from this calculation. Bile infarcts were quantified in H&E-stained sections in a similar manner and the percentage of the focal necrosis surface to the whole liver section area was assessed.

Immunohistochemistry and image analysis

For analyzing DNA-incorporated BrdU in liver cells, 4 µm sections collected on poly-L-lysine-coated glass slides were incubated with monoclonal mouse anti-BrdU antibody (1:50; Dako Cytomation, Hamburg, Germany) overnight at 4°C, followed by horseradish-peroxidase (HRP)-conjugated goat anti-mouse immunoglobin (LSAB kit plus; Dako). Sites of peroxidase-binding were detected by 3,3`-diaminobenzidine (Dako). Sections were counterstained with hemalaun. BrdU-positive hepatocellular nuclei were counted in a blinded manner within 30 consecutive high power fields (HPF) (x40 objective, numerical aperture 0.65)

and are given as cells/mm². In analogy, BrdU-expressing non-parenchymal cells were assessed and also given as cells/mm².

To specify the proliferative response of non-parenchymal cells upon BDL, we performed double immunohistochemistry of BrdU with specific markers for different liver cells: F4-80/BrdU for Kupffer cells and SM22α/BrdU for BEC. For each protocol, the immune-staining procedure for the specific marker was conducted after the BrdU staining protocol. Resident liver macrophages were analyzed using the F4-80 antigen (1:10; Serotec, Oxford, UK). Overnight incubation (4°C) with the first antibody (polyclonal rat anti-F4-80) was followed by alkaline-phosphatase (AP) conjugated mouse anti-rat immunoglobulin (1:200; Santa Cruz Biotechnology, Santa Cruz, CA, USA). The sites of AP-binding were detected using the chromogen fuchsin (Dako).

BECs and oval cells were detected by overnight incubation (4°C) with a polyclonal rabbit anti-SM22α antibody (1:100; Abcam, Cambridge, UK) followed by AP conjugated goat anti-rabbit immunoglobulin as secondary antibody (1:100; Dako). The sites of AP-binding were detected by Permanent Red (Dako).

Slides were viewed under a light microscope (Olympus BX 51) and the number of BrdU-positive cells co-expressing F4-80 or SM22α were counted in a blinded manner within 30 consecutive high power fields (HPF) (x40 objective, numerical aperture 0.65) and are given as cells/mm².

Antibodies for detection of α-SMA in HSCs and of S100a4-positive cells were from DAKO (M0851 and A5114, 1:500 and 1:200 dilution, respectively). CTGF antibody was from Santa Cruz (sc-1439, 1:200 dilution). Sections were de-paraffinized in serial ethanol dilutions. After a PBS wash, sections were transferred into 10 mM sodium citrate buffer (pH 6.0) and antigen unmasking was performed in a microwave. After cooling down, sections were incubated in peroxidase blocking reagent (Dako) for 1 h and with primary antibodies overnight at 4°C. EnVision peroxidase (Dako) was applied for 1 h at room temperature after a PBS wash next day. Sections were developed with diaminobenzidine for 5 minutes. The number of α-SMA-, CTGF- and S100a4-positive cells was quantified under a Leica light microscope (x20) by counting three fields.

High-throughput quantitative Taqman RT-PCR analysis

Total RNA was isolated from the liver tissue samples using RNeasy Mini Kit including on column genomic DNA digestion with RNase free DNase Set (Qiagen, Hilden, Germany). RNA was reverse

transcribed to cDNA with TaqMan Reverse Transcription Reagents (Applera GmbH, Darmstadt, Germany). For quantitative real-time PCR, we used the Fluidigm's Biomark high-throughput qPCR chip platform (Fluidigm Corporation, San Francisco, CA, USA) with pre-designed gene expression assays from Applied Biosystems, according to the manufacturer's instructions [6]. Data were analyzed using the ddCt method and expression values were normalized to the expression levels of the Gapdh gene.

Statistical data analysis

Dimension reduction: A one-way analysis of variance (ANOVA) was applied to reduce the data set to the parameter subset showing significant (p_{adj}<0.05) up- or down-regulation during the time course. Multiple testing correction was performed using the Holm procedure {Holm1979}.

Correlation analysis: Correlation between two parameter time courses was calculated via Y^{S3} a modified correlation coefficient based similarity measure for clustering time-course gene expression data {Son2008}. The correlation $Y_{i,j}^{S3}$ between two factors *i* and j is the linear combination of three terms: i) a classical correlation part based on Spearman correlation $S_{i,j}^{*}$, ii) $A_{i,j}^{**}$ accounting for the similarity in changes between the two time courses, and iii) $M_{i,j}^{*}$ comparing the location of the minimum and maximum values of the time course (see {Son2008} and supplement for definitions).

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

$S_{i,j}^{*}$ is calculated on individual data points, $A_{i,j}^{**}$ and $M_{i,j}^{*}$ on the mean time courses averaged over the repeats per time point. In the analysis the following weights were used:

$$\omega_1 = 0.5,\ \omega_2 = 0.3,\ \omega_3 = 0.2$$

All reported correlation values are $\bar{Y}_{i,j}^{S3}$ in the interval [-1,1] calculated by:

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

For the calculation of the similarity of changes within time courses, $A_{i,j}^{**}$, we used Spearman (S) correlation which is more robust against outliers as Pearson (R) correlation:

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

Simple Pearson and Spearman correlation were calculated for comparison.

Cluster analysis of the correlation matrix used complete-linkage hierarchical clustering with euclidian distance measurement. This combination of methods provided the best enrichments on gene-expression time-series in a recent comparisons of methods {Jaskowiak2014, Jaskowiak2013}.

Normalization of factors was performed separately for each factor f_k for all time points *i* and repeats *r* via

$$f_k^{norm}(t_{i,r}) = \frac{(f_k(t_{i,r}) - <f_k>)}{\max(f_k) - \min(f_k)}$$

Decision Trees: For the prediction of distinct time points of disease progression after BDL a regression tree model based on recursive partitioning {Breiman1984} was fitted using *rpart* with the mean normalized data of the 6 clusters as predictor variables and the log transformed time points as dependent variable

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

The logarithmic transformation was performed to obtain approximately equal intervals between time points. The regression tree was fitted using the full trainings, i.e. all 5 repeats per time point, with the minimum number of observations in a node for which a split is computed as 6, the minimum number of observations in a terminal node as 2, and a complexity parameter of $c_p = 0.01$. The splitting criterion used to decide 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. The predictive capacity of the model was evaluated using all single combinations of individual factors from the clusters, and a random subset of 10000 two factor combinations from each cluster. The predictive capacity of a combination of factors was scored using the root mean square distance of all samples on log scale

$$\mathrm{d} = \exp\left(\frac{\sqrt{\sum_{i=1}^{N}\left(\tilde{t}_i^{pre} - \tilde{t}_i^{exp}\right)^2}}{N}\right) - 1$$

The best combination of single factors minimizes d.

All computations were performed in R with source code, data and the full analysis available in supplement S2.

Results and Discussion

Temporal changes of biochemical, cellular and histochemical markers after BDL

In mice, BDL over 14 days induces time dependently progressing stages of a secondary biliary CLD. The first week after BDL begins with an acute cholestatic injury associated with necroinflammation, followed by a chronic injury stage, comprising hepatitis and liver fibrosis. Macroscopically, marked dilation of the gallbladder and formation of bilioma are found, associated with weight loss and a mortality rate of 10% in the first week due to bile leakage and rupture of the gallbladder [7]. Pathophysiologically, BDL interferes with glandular liver function and hepatobiliary transport, which comprises its detoxification and secretion function including bile formation. Obstruction of the bile duct leads to afflux of newly generated bile fluid. The main components of the bile, bile acids and phospholipids, induce toxicity and damage towards hepatocytes and cholangiocytes, therewith initiating the disease process. Rapidly after BDL, mice develop obstructive jaundice and cholestasis, as displayed by markedly elevated serum transaminases and bilirubin levels (Figure 1), macroscopically evident from yellow ears and urine. Within the first 30 h, there is a massive release of liver enzymes, like ALT and GLDH, reflecting the initial hepatocyte damage as initial pathophysiological event in the process of BDL-induced liver fibrosis (Figures 1A, B).

The plasma levels of diagnostic liver enzymes remained elevated over several days, but then dropped to reach after 14 days values that were slightly above those of sham operated mice. Concomitantly, liver detoxification capacity was deteriorated, as indicated by the rise of total bilirubin, a classical plasma marker of cholestasis (Figure 1C). Notably the plasma level of albumin, an important parameter for the evaluation of liver functionality remained relatively constant over the time course of 14 d (Figure 1D). The systemic blood cell count showed constant levels of erythrocytes and platelets up to day 5 (Table 1). In contrast, leukocytes decreased by 50% during the first two days, reflecting intrahepatic cell entrapment, and recovered to values of sham operated animals within the subsequent observation period (Table 1). During progression of fibrosis, red blood cells, hemoglobin and hematocrit slightly decreased.

In consequence of intrahepatic toxic bile accumulation, progressive development of confluent bile lakes is a hallmark of cholestasis. Histological quantification of bile infarcts, defined as clusters of injured hepatocytes, revealed a steady rise of infarct areas until day 14 after BDL (Figure 2A). The typical

appearance of liver tissue at representative time points after BDL using H&E staining is depicted in Figure 2B. Further histopathological changes of the livers after BDL included enlargement of portal tracts, accompanied by dilation of bile canaliculi and proliferation of BECs and oval cells (Figure 3A), resulting in formation of artificial bile ductules (Figure 2C), a cellular response termed 'ductular reaction' [5; 8]. However, recent data from lineage tracing experiments indicated that these cells, however, do not contribute to the population of ECM producing/fibrogenic cells, which in the BDL model is largely consisting of HSCs [9].

The inflammatory response resulting from chronic hepatocyte injury resulted in the accumulation of immune cells in the liver, among others, T cells, macrophages and dendritic cells, which were mainly found within and around bile infarct areas (Figure 2C, asterisk) [10]. The inflammatory and fibrogenic response is initiated by resident liver cells, primarily liver macrophages (Kupffer cells) and activated hepatic stellate cells (HSCs) secreting a wide range of cytokines and chemokines which determine the quality and quantity of inflammatory and consequently fibrotic responses [11; 12]. Upon parenchymal damage, quiescent HSCs undergo phenotypical changes to myofibroblasts (MFBs). The most prominent role of MFBs is extracellular matrix (ECM) production and reorganization, as reflected by, among others, increased synthesis of α-SMA, type I collagen and TIMPs. A marked increase of the number of α-SMA- and S100a4-positive cells, as measured by immunohistochemistry, was observed after BDL (Figures 3B, C). Migration of MFBs to the site of injury and their contractility contribute to liver scarring. This is accompanied by parenchymal cell proliferation, which began to rise at day 2 as regenerative response but decreased at day 14 (Figure 3D). With a slightly faster response as compared to HSCs, starting at 30 h upon BDL, Kupffer cells started to proliferate (Figure 3E).

The overall hepatic proliferative response as analyzed by immunohistochemistry was confirmed by elevated mRNA expression of Ki67 (Figure 4A).

Between days 5 and 14 after BDL, periportal alterations were associated with fibrotic changes. As demonstrated by Sirius red staining, extensive fibrosis, characterized by a several-fold increase of collagen deposition (Figure 4B), including bridging, was observed 5 days after BDL (Figure 4C). We further stained for connective tissue growth factor (CTGF), a prominent fibrogenic cytokine and enhancer of TGF-β effects [13]. Appearance of CTGF-positive cells started as early as 12 h upon BDL and their count increased continuously (Figure 3F).

Evidences for different time phases of diseases progression after BDL

For the identification of distinct phases in disease progression induced by BDL we included besides biochemical and histochemical markers time-resolved transcriptomics profiles of three different gene panels related to hepatocyte metabolism fibrogenesis and inflammation using the Fluidigm platform (Figure 5). The selection of representative genes for ADME- (absorption, distribution, metabolism, and excretion) (Figure 5A) fibrogenesis- (Figure 5B), and inflammation-related genes (Figure 5C) was made arbitrary based on state-of the art knowledge.

In a first step, a one-way analysis of variance (ANOVA) was applied to reduce the complete data set of biochemical, histochemical and transcript data consisting of 154 factors to a subset of those factors showing significant ($p_{adj} < 0.05$) changes during the time course of disease progression (see section ANOVA in Supplement 2). This resulted in a subset of 90 factors comprising two biochemical markers (B: bilirubin, GLDH), five histochemical markers (H: HSC, NHC, Kupffer, Sirius red), three immunostainings by antibodies (A: CTGF, a-SMA, S100a4) and 80 genes (14 ADME, 22 fibrosis, 44 inflammation panel). Within this subset, a bivariate time-dependent correlation analysis (see Methods and Supplement 2) was performed for all pairs of factors to identify those factors displaying similar temporal profiles (see Figure 7). Based on the obtained correlation matrix a hierarchical cluster analysis was applied whereby the minimum cluster size was put to 2. This yielded 6 different clusters comprising between 2 – 61 factors with distinct time courses attaining their maximum at different time points (see Fig. X). Accordingly, the clusters can be assigned to the early, middle and late phase of disease progression. For example, mRNA levels of genes involved in metabolism (ADME, Cluster ?) such as members of the cytochrome P450 system are immediately induced during the first 6 h upon damage and then steadily decrease with time after BDL. This time course suggests that early after the insult the detoxification activity of hepatocytes is increased but than decreases owing to ongoing necrosis of hepatocytes and feedback inhibition of CYP 450 expression by accumulating bile salts.

The found clusters comprise both ‘classical’ biochemical and histochemical factors and genes characteristic for a specific phase of disease progression. The correlation between these two cardinally different types of factors is shown in Fig. 8. The time course of each of the ‘classical’ factors contained in the ANOVA-subset can be correlated with the expression time course of at least one gene. Only for GLDH and BrdU Sirius Red this correlation is weak (p=?).

In what follows we discuss in more detail the biological significance of the 6 clusters of timely correlated factors and the correlation between different types of factors shown in Fig. 8 thereby including already exiting knowledge.

mRNA levels of genes involved in metabolism (ADME, Figure 5A), such as the classical representatives of the detoxifying cytochrome P450 system are immediately induced during the first 6 h upon damage and are then steadily decreasing with time after BDL (Figure 5A). This indicates that early after the insult, detoxification activity is increased to interfere with damage that however is too strong. Consequently, ongoing hepatocyte depletion leads to significant decreases in total liver enzyme expression and activity. Several exceptions for genes that play a role in detoxification and in antagonizing oxidative stress, such as Cyp3a11, Gsta2 or Sult1a1, exist. BDL-induced significant changes were identified for genes playing a role in the regulation of oxidative stress, e.g. Nos2 and Nfkbia.

For example, analysis of both fibrillar collagen1α1 and 3α1, which predominantly exist in fibrotic livers, showed a significant upregulation rapidly after BDL and a continuous increase with severity of liver fibrosis up to 14 days after BDL (Figures 6E, F). Among the peptide mediators, not very surprisingly, Tgf-β (isoforms 1 and 2) mRNA expression was steadily increasing, confirming its postulated role as fibrogenic master cytokine [14]. Associated with induction of fibrogenesis-related genes (Figure 5B), which particularly are representative for HSC activation, the dynamics of the inflammation gene signature (Figure 5C) nicely matches with the increase in the number of proliferating Kupffer cells (F4-80/BrdU staining values) observed from day 2 onwards (Figure 3E). Very low expression levels were present immediately after BDL, except for the chemokines Cxcl1 und 2. Starting at time points between 2 and 5 d after BDL, most cytokines and chemokines in the list were strongly upregulated for a longer period of time, until day 14.

Taken together, the detailed time-resolved transcriptional profiling of liver homogenates following BDL revealed a coordinated induction of detoxification processes immediately after surgery and an upregulation of an inflammatory response along with activation of metabolically active genes, which can be explained by physiological recovery and adaptation of the mice to the bile acid exposure.

Correlations

To assess interrelations between level and time of pathophysiological events and predictive markers within the complex scenario of cholestatic liver disease, correlations were analyzed. The correlation of the time point averages (column Av in Figs. 7 and 8) is high for most parameter pairs, not unexpectedly, as most histologic parameters and cell count observations, as well as most genes related to inflammation and fibrogenesis increase with disease progression (about 2/3 of all factors studied). However, there is large variance in values for many analyzed factors, when comparing individual mice of the same time points. For example, at 5 days, the infarct area varies from 0.9% to 12% and the collagen deposition area varies from 0.8%-5.9%. Thus, we decided to study correlations of determined parameters from individual mice (All) with regard to time points (T0h...T14d) and time frames (6-18h and so on) to yield more specific results. With that strategy, we were able to identify, for example, parameter pairs displaying a correlation only at specific phases. As all the identified correlations, both time course and single mouse correlations, represent complementary aspects, a consensus measure is calculated, which is used for a final assessment of the relation of two factors, as shown in the hierarchical clustering in Figure 7 and to determine the ranking of correlated factors in Figs. 8 and 9.

To identify global connections between factors, the consensus correlation coefficients have been subjected to a cluster analysis (Figure 7). RNA levels of commonly regulated genes form highly correlated clusters, e.g., (i) cytokines and growth factors, such as Il6 and Tgfb (lower left, blue), (ii) fibrosis-related genes, such as Coll1a1 and Timp1 (bottom right, cyan), (iii) ADME-related genes, such as isoforms of Cyp24a1 and Nr0b2 (red, top right). There are also RNAs, which poorly correlate with any other parameter, such as Rarres1 and Igf1. GLDH as well as albumin are not well-connected with any other parameter, while serum ALT clusters with Cyp7a1 by a negative correlation. Sirius red-positive area (collagen) is quite isolated in the graph map, and most closely relates to the cytokines' cluster (left, grey). BrdU-positive Kupffer cells, BECs and S100a4-positive cell numbers are clustered with the closely related RNAs of Gstm1 and Gsta2 (top left, purple, brown, light blue). BrdU-positive hepatocytes are clustered together with Notch1, Birc5, and Mki67 (bottom, cyan). CTGF and α-SMA positive cells are clustered with the RNA of Pparg and Gstp1 (top left, grey blue).

Initial response

Immediately after BDL, there is a massive release of liver enzymes up to day 5, followed by a drop down to almost values of sham-operated livers (Figures 1A, B). It can be interpreted as the beginning breakdown of hepatocellular protein synthesis or the cease of cell necrosis.

The factor with the strongest negative correlation to ALT in serum is Cyp7a1 (Figure 8A). As Cyp7a1 mRNA encodes cholesterol-7-α-hydroxylase, down-regulation of bile acid synthesis can be seen as a fast and straightforward response to cholestasis. As underlying mechanism, activation of the JNK/c-Jun pathway has been reported [15]. In later time phases Cyp7a1 increases again, most likely SREBP-mediated, since plasma cholesterol concentration increases simultaneously [16]. The second highest correlated factor to ALT is Hmox1, encoding heme oxygenase, which was reported as increased upon BDL [17]. It therefore also can be defined as “early response” parameter, which subsequently remains at increased levels as compared to healthy liver. GLDH is another serum marker (Figure 8B) for hepatocellular injury that positively correlates to ALT, but less significant than Cyp7a1 and Hmox1, indicating that ALT and GLDH are suitable as independent measures. Additionally, a high correlation is found for ALT with Cebpb, which encodes CCAAT/enhancer-binding protein β, a regulator of the inflammatory response, e.g. via up-regulating Il6 [18]. The two chemokines Cxcl1, encoding neutrophil-activating protein 3, and Cxcl2, encoding macrophage inflammatory protein 2-α, are positively correlated to ALT. As both proteins are excreted, it is likely that they can be detected in the plasma and thus may be further investigated as potential diagnostic marker.

GLDH like ALT shows a strong initial response but in contrast to the latter as increases further up to 18h, before it decreases gradually (Figure 8B). It is highly correlated to ALT, but shows a low correlation to all other factors (Figure 8B). In particular, the anti-correlation of GLDH to Cyp7a1, the 2nd highest correlation, is considerably lower than the ALT-Cyp7a1 correlation (see Supporting Information Dataset S3, section 4.2). Further, it is correlated to RNA levels of Fibronectin (Fn1, see Figure 6G), but only in time frames including the 18h time point (T18h, 6-18h, Perp).

Macroscopic organ damage

As documented with the H&E staining in Figure 2B, necroinflammation is caused by the BDL-induced intrahepatic toxic bile accumulation with individual liver cell death and progressive development of confluent

bile infarct areas. The total area of infarcts increases steadily with relatively high variance (Figure 2A). The most consistently positively correlated factor with bile infarcts is plasma bilirubin, which occurs in most time frames and even at single time points (Figure 8C). The correlation is not very strong and the significance level is <5%, valid only for 3/11 time frames and 2/7 time points (see Supporting Information Dataset S3, section 4.10). The two parameters are biologically related as the infarct area inversely accounts for the number of functional hepatocytes.

Next most correlated is the RNA level of Il17a encoding interleukin-17A (Figure 8C), which plays a pivotal role in cholestatic liver fibrosis in mice by activation of both the KCs and HSCs [19]. The correlation is high only for later time points. Il17a is switched on between 2d and 5d to very high RNA levels, and as it is a secreted protein, it is likely to be detectable in the blood, thus representing a candidate diagnostic marker. Expression of Notch3 is initially negatively and later positively correlated with bile infarcts. This corresponds to the pattern of Notch3 expression. It drops below the level of untreated mice and is increased between days 2 and 5. Notch3 is reported to be significantly up-regulated in fibrotic liver tissues, most supposedly by regulating the activation of HSCs [20]. Next, Prom1, encoding prominin 1 (CD133), is reported to be increased in alcoholic hepatitis [21] and chronic liver injury [22], and was dedicated to be regulated by the DNA methylation in HSCs [23]. Many of the selected genes were positively correlated to a similar extent, with the highest values occurring for the later time frames. Among those, Col3a1 (see Figure 6F), Sparc, Col8a1, Wisp1, and Edn1 show a similar pattern as Notch3 for the initial anti-correlation. Remarkably, all genes whose expression values have a high consensus correlation to the infarct area show a high correlation for the 5d time point and the time frames covering day 5, while there is low correlation for the earlier time points and frames (Figure 8C). In addition, 5 days is the time point with the highest variance of the infarct area (Figure 2A). This indicates that different routes of the disease process exist, defined by the infarct area at day 5 and marked by expression of genes such as Il17a, Notch3, Prom1, Col3a1 (see Figure 6F), Actb, Sparc, Tnc, and Tgfbr2.

Loss of liver function

There is a large variability of bilirubin levels among different mice from the same time point. As bile cannot leave the liver via the biliary system due to cholestasis, differences in the measured serum bilirubin

show the capacity of the hepatocytes and the bile lakes to store conjugated bilirubin. The highest correlation with serum bilirubin levels can be observed for Tnfrsf1b, encoding tumor necrosis factor receptor 2 (Figure 8D). Next, a high correlation is found for Ccl2, encoding chemokine (C-C motif) ligand 2, a soluble biomarker for hepatic fibrosis in NAFLD [24]. The 3rd highest correlation is found for Il10rb, encoding the interleukin 10 receptor β subunit, which was found up-regulated in NASH with fibrosis [25]. Intriguingly, the 12 factors most correlated to bilirubin levels show a negative correlation for the final time frame 14d (see Supporting Information Dataset S3*,* section 4.4). For instance, the mouse with the lowest bilirubin levels at day 14 (IM2, 9.12, see Supporting Information Dataset S1) has the highest Tnfrsf1b expression, leading to the hypothesis of alternate progression courses in the final disease stage.

Surprisingly, albumin synthesis is maintained relatively constant over the observation period of 14 days (Figure 1D). Apparently, hepatic degradation of albumin is affected at similar levels as its synthesis. Due to this finding, it is not surprising that albumin is poorly correlated to the dynamics of other factors investigated (Figure 8E). Cyp2b10 shows high correlations to Albumin only for separate time frames and time points.

Hepatic cell proliferative response

In later phase of the disease process, different hepatic cells proliferate, underlined by co-immunostaining with BrdU and specific markers for hepatocytes, KC, and BECs (Figure 3), and indirectly reflected by the marked up-regulation of Ki67 mRNA (Figure 4A).

As compensatory activity of the liver to restore the damaged parenchyma, hepatocytes proliferate, monitored by the parameter BrdU-positive hepatocytes. It is markedly triggered between the 30h and 2d time point (Figure 3D). It is highly correlated with Mki67 expression (Figure 9B), encoding the antigen Ki-67, a known proliferation marker (Figure 9B). The 2nd highest correlated gene is Birc5, encoding survivin. Yes-associated protein has been reported to regulate the hepatic response after bile duct ligation via modulation of survivin [28], underlining the role of survivin in hepatic tissue restructure. The 3rd highest correlation is found for Notch1, a transmembrane receptor involved in developmental processes, and its increase can as well be seen as a sign of cell plasticity and tissue restructuring. Next in line is Cdh2, encoding cadherin-2, which is normally associated to cancerous cells. The correlation is high at the time points 30h to 5d. Then Lama1, encoding laminin subunit α-1, the next most correlated factor, follows an

opposite pattern, as the correlation is high at early (6h) and late (2d-14d) time points. Lama1 was found to be increased in nonalcoholic fatty liver disease [29]. Further, Timp2, encoding tissue inhibitor of metalloproteinases 2, an antagonist for degradation of extracellular matrix (ECM), also correlates with the hepatocyte proliferative response and reflects increased ECM deposition and buildup of fibrotic tissue.

Liver macrophages infiltrate the liver tissue, are activated, and proliferate, monitored by the parameter "BrdU-positive Kupffer cells". This starts at 30 h in our time course to reach a maximum at day 2 and to decrease again thereafter (Figure 3E). There is no strong correlation with other parameters, yet the highest with Gsta2, encoding glutathione S-transferase A2 (Figure 9A), which facilitates bilirubin import [26]. The correlation is confined to the middle time points (30h-2d), the time frame of the strongest increase of KC proliferation, while in the beginning (6h-12h) and the end (5d-14d) a slight anti-correlation is found. The next highest correlation is to BrdU-positive BECs, again in the middle time frames (30h-2d), and with S100a4-positive cells, which itself has a higher correlation to BrdU-positive BECs. The same is true for the RNA level of the cell cycle gene Mki67. The factor Por, encoding cytochrome P450 reductase, is negatively correlated in the time frames and time point from 30h on, thus, resembles the up-and-down pattern inversely. Por is reported to be down-regulated in liver cirrhosis via the aryl hydrocarbon receptor AhR [27].

Also biliary epithelial cells (BEC) proliferate, monitored by the parameter "BrdU-positive BECs". The main increase occurs between days 2 and 5, where a minor increase can also be observed after 30 h (Figure 3A). Further, a remarkable variability of BEC proliferation in the different mice can be seen at day 14. The highest correlation of this event is observed for the number of S100a4-positive cells (Figure 9C), which is representative for the number of KC and is in line with the recent finding about feed-forward and feedback regulation of the two cell types BEC and KC in liver fibrogenesis. Only the time frames near the perpetuation phase show a high correlation. Not surprisingly, the gene expression most correlated to BrdU-positive BECs was found for Col1a1 (Figures 9C and 6E), encoding the collagen deposited in ECM in large quantities. The 3rd highest correlation is found for Tgfb, encoding the cytokine TGFβ, which is well known to correspond with the fibrotic process in a positive feedback loop [30]. Interestingly, the negative correlation at the 6h time point shows high significance suggesting a negative feedback regulation. The 4th highest correlation is found for Il10rb, encoding the β-subunit for the Il10 receptor, which was described to be increased in rat liver fibrosis [31]. Slc10a1, encoding the Na+-taurocholate co-transporting polypeptide, which transports bile

acids through the basolateral membranes of hepatocytes is negatively correlated with BEC proliferation, mainly in the later time frames (30h-14d). A decrease of Slc10a1 has been shown to protect hepatocytes from cholestasis-induced injury [32]. Another negatively correlated factor is Ugt1a1, encoding UDP-glucuronosyl-transferase 1A, the main enzyme for conjugation of bilirubin, whose down-regulation is considered protective against the increased concentration of conjugated bilirubin in hepatocytes.

Increase in fibrogenic cells

In accordance to the proliferative activity of HSCs, immunohistochemical analyses demonstrate the gradual rise in CTGF-, α-SMA- and S100a4-expressing cells (Figure 3), reflecting activated HSC, activated KC and trans-differentiated hepatocytes (such that they are producing hepatokines). CTGF is a highly pro-fibrogenic protein expressed by HSCs, BECs and hepatocytes [33; 34] and mediates extracellular matrix inducing properties. It is reported that levels of this protein are significantly up-regulated in experimental liver fibrogenesis and human chronic liver disease patients of various etiologies [35; 36]. The number of CTGF-positive show a steady increase with comparatively little variance (Figure 3F), as shown by the high significances in ANOVA analysis (see Supplementary File 4). The highest correlation with CTGF positive cell number shows Tgfb2, encoding the cytokine TGF-β2 (Figures 9D and 6I). TGF-β is the major stimulus for CTGF expression in hepatocytes [37] and elevated levels of Tgfb2 were reported for bile duct ligated rats [38]. This is the result of a direct interaction [39], and the correlation is close to 1 for several time frames (30h-14d) and all time points, except for 30h. The next highest correlation is found for Pdgfb, encoding platelet-derived growth factor subunit B, which is reported to be up-regulated in liver inflammation and fibrosis [40]. The number of cells containing CTGF and α-SMA is also highly correlated. There is also a strong correlation to Tgfbr2, encoding the transforming growth factor β receptor 2 [41]. Additionally, there are considerable correlations to several other genes such as Timp1, Cxcl5, Sparc, and Col8a1.

The parameter "α-SMA-positive cells", a marker of HSC activation, increases steadily with the course of disease progression (Figure 3B). Not surprisingly, it is positively correlated to the number of CTGF-positive and S100a4-positive cells (Figure 9E). Interestingly, CTGF staining is highly anti-correlated to Cyp1a2 (Figure 6A), encoding cytochrome P450 1A2, known to be lowered in liver cirrhosis mediated by AhR [27]. The next most correlated gene is Pparg, encoding peroxisome proliferator-activated receptor

gamma, but the most relevant correlation occurs in the early time frames, that is 18h-30h. PPAR-γ inhibits HSC activation [42]. As Pparg is increasing only in the early stages, we conclude that TGF-β and Pparg form a threshold system, where HSC activation is controlled in the first stage and overshooting in later time frames. Cyp2c37 shows a negative correlation in the early time frames (6h-18h), but the correlation of α-SMA-positive cells to S100a4 is higher, as outlined below.

The S100a4 protein has been dedicated to myofibroblasts upon epithelial mesenchymal transition of hepatocytes in liver [43; 44]. This was however disproved in a careful analysis, indicating that the protein marks Kupffer cells [45]. The parameter “S100a4-positive cells” shows a steady rise until day 2, after which it stays constant (Figure 3C). Not surprisingly, it is positively correlated to the BrdU-positive BECs and α-SMA-positive cells, but significant correlations do not occur after day 2 (Figure 9F). Expression of the gene Gstm1 (Figure 6C), encoding glutathione S-transferase Mu 1 is highly correlated. Genetic variation increases the susceptibility to alcoholic liver disease [46], but no connection between its regulation and other types of liver diseases has been reported. Cyp2c37 is negatively correlated, especially in the early time frames. The time course of Cyp2c37 expression shows a strong initial decline until 12 h, a plateau until day 2, and a final decline. Egfr, encoding epidermal growth factor receptor is also negatively correlated. Interestingly, both, but especially Egfr, show a higher anti-correlation for single time points than for time frames. Thus, although Egfr does not change much in the time course, it is a marker for increase of S100a4 positive cell number.

Fibrosis

During the perpetuation phase (18h-2d), paracrine and autocrine cytokines amplify hepatic inflammation and HSC activation resulting in continued ECM remodeling, being characterized by the enhanced mRNA expression of both fibrillar collagen1α1 and 3α1 (Figures 6E and F). Additionally, the progression stage of fibrogenesis is histomorphologically characterized by excessive deposition of extracellular matrix (Figure 4B), as analyzed by Sirius red staining of liver slices (Figure 4C). Significant correlations of “Sirius red positive area” are found after day 2 (Figure 8F). The most correlated parameter is Osmr, encoding the oncostatin M receptor, which transduces, among others, Il31 and oncostatin signaling events. Interestingly, it has been reported that in cirrhotic liver, Osmr is not expressed [47]. Thus, expression

of this gene may be a candidate distinguishing factor for liver cirrhosis and cholestasis-induced liver injury, which however has to be confirmed with further investigations. Also Osm, encoding oncostatin M itself, is correlated, showing a combined induction of the ligand and its receptor [48]. Next most correlated with Sirius red staining is Ccr5, encoding C-C chemokine receptor type 5, which is a regulator of inflammation as well as macrophage recruitment and trafficking [49], thus representing a general promoter of hepatic fibrosis [50]. Additionally, blood bilirubin levels are correlated, as described above. Further correlated factors are Tnfrsf1b, encoding tumor necrosis factor receptor 2, and Tnfrsf1a, encoding isoform 1 of the receptor (see Figure 6L). The latter was previously reported as necessary for liver fibrosis in mice [51]. Then Cxcr1, encoding interleukin 8 receptor α, is reported to be highly up-regulated in chronic liver disease [52] and is found correlated with Sirius red (Figure 8F).

Markers of disease progression

We next asked the question, which of the analyzed factors characterize particular stages of the disease process best. One-way ANOVA identifies the parameters whose values at different time points are different in a general perspective. The calculated p-values allow a ranking how good the parameters deviate in the time course, the top significant are: Cyp1a2, serum bilirubin, Il10rb, Tgfb1, Ccl2, Cd86, Ccr2, and Mrc1.

In a more specific analysis, parameters were identified which have the largest difference of values in a particular time frame, a consecutive series of time points or a single time point. A separator for a particular time frame is a parameter, whose value range for mice belonging to the respective time frame is disjoint from the value range of mice outside this time frame. The quality of a separator is measured by the distance between the ranges in relation to the full range, the separation gap (see Supporting Information Dataset S3 for details).

We identified 41 separators for specific time frames. For 35 of these separators, the values of the factor show a single transition, while for 6 separators, an internal time range is marked. 12 separators select a single time point, 9 of which separate the 0h time point (the control) from ligated mice. The transcript abundance for Nr0b2 (small heterodimer partner, SHP, see Figure 6D) is strongly decreased only for mice at the 6h time point. SHP is a repressor of retinoid and thyroid hormone, as well as estrogen receptors and by being diminished in human cirrhotic livers, it is suggested to be associated to cirrhosis and hepatic tumors

[53]. This is also supported by the fact that cholestatic liver fibrosis induced by BDL was increased in SHP-/- mice [54]. Cyp24a1 (mitochondrial 1,25-dihydroxyvitamin D3 24-hydroxylase, see Figure 6B) is also a separator for the 6h time point, with a lower separation strength. It had been found increased in hepatoma cells [55]. Mmp10 (Stromelysin-2/transin-2, see Figure 6H) is a separator for the 18h time point, the only other internal time point with such a separator in the investigated parameter set. By degrading proteoglycans [56] and fibronectin [57], the metalloproteinase contributes to ECM break down [58] and is found increased after liver injury [57] and at intoxication [59]. The transcript of Cdh2 (cadherin 2) has increased levels only for 6h and 12h. Il28b and Il2 (Figures 6J and K) display a similar pattern, but the separation gap is small.

Among the 9 separators for untreated mice (0h), serum ALT is the strongest separator. Serum bilirubin, Hmox1 (decycling heme oxygenase 1), and serum GLDH have a separation >1%, while α-SMA-positive cells, Cxcl2 (chemokine (C-X-C motif) ligand 2), Cd14 (CD14 antigen), Timp1 (tissue inhibitor of metalloproteinase 1), and Mmp10 (Figure 6H) have smaller separation values.

CTGF-positive cell number is the separator with a large separation gaps for several time frames of transition (6h/12h, 12h/18h, 30h/2d, 5d/14d). Thus, it is clearly the best candidate from those selected to monitor the disease progress. It is also among the parameters with the highest ANOVA significance (which is particularly remarkable as only 3 could be analyzed compared to 5 for other parameters, See Supplementary File 4, section 1). S100a4-positive cells is a similarly good separator for two transitions (12h/18h, 30h/2d). The parameter α-SMA-positive cells is a good separator for the transitions 0h/6h, 6h/12h, 12h/18h, thus it can be considered as the best marker for the initial disease process.

Cyp1a2 (cytochrome P450 1A2, Figure 6A) was identified as best separator for the 6h/12h transition, whereby transcript levels are initially high and steeply decrease between 6h and 12h. Cyp1a2 shows the highest significance in the ANOVA test, and parameters pairs with Cyp1a2 as one partner yield the highest significances in MANOVA tests which show its robustness as a marker, see Supplementary File 4. The gene product has already been observed to decrease upon bile obstruction [60]. The two other RNA separators for this transition (Cd14, Ccl2) are poor.

There are no suitable RNA markers for 12h/18h, 18h/30h and 30h/2d transitions. For the 18h/30h transition, there are no separators at all, which supports the definition of a perpetuation phase represented by the 18h and 30h time points. Strong separators are abundantly available for the 2d/5d transition, which

indicates that large qualitative changes occur between day 2 and day 5 of the experimental disease process. The best separator is Il28b (interleukin 28B; Figures 11 and 6K). Based on this investigation, we conclude that transition from day 2 to day 5 can be easily monitored, also in clinical practice, as serum bilirubin is among the best separators, and the interleukins (Il28b, see Figure 6K, Il13, Il17a) as well as the growth factors (Pdgfb, Tgfb2, see Figure 6I, Hgf) and contributors to the extracellular matrix (Sparc, Col3a1, Col1a1, Figures 6F and E), all together encode excreted gene products, that should be detectable in blood samples. Sparc (secreted acidic cysteine rich glycoprotein), a known indicator of chronic liver disease [53] and a mediator of fibrosis [61], also has a large separation gap.

For the 5d/14d transition, CTGF-positive cell numbers and mRNA expression of Ctgf are the only separators, which is an argument in favor of the combination of the 5 and 14 day time point in the progression phase.

Figure 10 shows a summary of the disease processes, the serum factors or histopathological parameters they represent, and the most correlated factors.

Decision trees for disease progression

Our next aim was to find combinations of factors allowing a prediction of the time after the ligation for a putative mouse. In a first approach a MANOVA combining two factors followed by a linear discriminant analysis (LDA) was performed. The combination of the parameters CTGF cells and Cyp1a2 yields high recall of 0.88 in the leave-one-out cross-validation in the discriminant analysis, however, CTGF cells having only 3 repeats, violated the minimal requirement for MANOVA of this pair. The combination of Cyp1a2 and Tgfb1 obtained a recall of 0.72 and high significance. It shows, that these parameters along with other top scorers as Tnfrsf1a, Hk2, Fn1, Cyp3a11, Il1b, Il10rb are suitable marker candidates. See Supplementary File 4 for more information.

However, using MANOVA and LDA with so few repeats is problematic. For instance, the check on normality of the parameter values for each time point is not reliable. Therefore we applied a more direct approach which uses fewer assumptions. As many factors possess disjoint value ranges at specific time points, an assessment with binary decision trees is possible where to each node a question is attached, whether a specific factor has a smaller or larger value than a given threshold value.

Computationally, the decision tree is determined by the rule that the partial separator with the largest relative gap is used. A partial separator is a combination of a factor and a threshold value, which splits a time frame in subordinate time frames, if the respective value ranges are disjoint. The thresholds of the decision tree are defined as the median of the gap between the ranges. For a partial separator, it is sufficient to split a time frame (except for the root of the decision tree), e.g. ALT separates very well 0h from 6h (with respect to the time frame 0h-6h), although the values at 14d are similar to the values at 0h. A partial separator is measured by the quotient of the gap range and the value range, the relative gap.

The decision tree to predict the time point computed from all available data is shown in Figure 11A. This decision tree recalls the available data, i.e. the time point of every mouse is correctly predicted. It relies mainly on the factors CTGF- and α-SMA-positive cells. This is consistent with data from other studies, which observed a correlation of increased CTGF levels with histological fibrosis stages [62; 63], suggesting CTGF as a valuable diagnostic target, since it can be measured in patients' blood and maybe used to follow-up on patients suffering from chronic liver diseases [64]. Other factors identified are Tnfrsf1a, Gstm1, and Il28b. Tnfrsf1a (tumor necrosis factor 1, see Figure 6L) aggravates steatohepatitis [65] and is essential for HSC proliferation and ECM remodeling [66]. Polymorphisms of Gstm1 (glutathione-S-transferase mu 1, see Figure 6C) are a risk factor in alcoholic liver cirrhosis [67]. Il28b (interleukin 28β, see Figure 6K) polymorphisms are associated with fibrosis progression in patients with chronic hepatitis C [68].

As the decision trees are automatically generated, their predictive capacity can be tested with the leave-one-out approach. For each mouse in the experiment, a decision tree is generated under the exclusion of data from this mouse. Then, the decision tree is evaluated with the factor values of this mouse and compared with its true time point.

Doing so, 24 of 40 mice were not accountable by the decision tree, because the factors CTGF- and α-SMA-positive cells used in all trees were only measured for 24 mice (3 of 5 repeats). Of the 24 accountable mice, the predictions were correct for 15. For those 9 wrong predictions, the predicted time point was adjacent to the true time point. 4 of them misclassified between 18h and 30h, 3 between 2d and 5d, one between 5d and 14d, as well as between 0h and 6h. The reason for misclassification is always an extreme value of the left out mouse.

As most of the misclassifications occurred between 18h and 30h, reliability is presumably improved if these time points are combined. Thus, for the next series of decision trees, phases instead of time points shall be predicted. As the histological factor “CTGF-positive cells” is the only separator between 18h and 30h, decision trees for time points must use this factor, while for the prediction of phases, decision trees are possible that do not depend on histologic parameters.

The decision tree to predict the disease phase deduced from all available data is shown in Figure 11B. Il28b and CTGF are contained as decision parameters, like in the decision tree for time points. Additionally, Fn1 (Fibronectin, see Figure 6G) is used, a constituent of the ECM in liver fibrosis. The leave-one-out approach yielded decision trees, for which 29 of 40 mice could be assessed because the parameter CTGF-positive cells was measured for the left out mouse. They predicted the phase for 28 of 29 mice correctly. The wrong prediction was caused by the transcript value of Il10rb for the left-out mouse, which is far from the values of the other mice of the same time point, and the algorithm did not select this factor as a predictor for all other decision trees.

The histopathological factors (cell counts) were only measured only for 24 mice. After their exclusion, all 40 mice can be subjected to the leave-one-out test, in which 37 of 40 predictions were correct. Wrong predictions were the result of outlying values for Cyp1a2 (Figure 6A) and Il10rb (which were chosen as the decision factors for the decision tree in these cases), and the largest value for Cdh2. As cadherin is not excreted and presumably not suited as a serum marker, it was also excluded. Finally, a decision tree of Il28b, Fn1, and Il2 (Figures 6K, G, and J) was formed, genes whose gene products are likely to be excreted and more easily observable (Figure 11C). The leave-one-out test again yielded 37 of 40 correct predictions. See Supplementary File 3 for more details.

Alternate progression routes

The large variation in the parameters infarct area (Figure 2) and collagen area (Figures 4B and C) in mice from the late experimental time points are an intriguing finding, considering the homogeneity of the experimental system. Indeed, they are negatively correlated with each other, and suggest that alternate routes of disease progression exist. One route is characterized by a strong increase in necrotic tissue and a weaker activation of HSC and lower expression change of inflammation factors. The other route is

characterized by strong activation of fibrogenesis factors, and finally macroscopic collagen tissue. Both routes are similarly connected to loss of liver function, however, the histopathological display is quite different: the former contains a large amount of necrotic tissue while the latter contains large fibrotic tissue. Once verified also for human livers, the hypothesis of alternate disease routes can be of far-reaching importance for an individualized therapy, as obviously the medical interventions avoiding necrosis differ from interventions reducing overshooting fibrosis. The design of the study, which included the sacrifice of mice after a specific time, did not allow to answer whether the alternate developments, shown macroscopically in later time points, can also be observed in earlier time points.

Conclusion

The time-resolved analysis of a wide range of parameters (6,313 in total) in bile duct ligated mice has shown that many factors share the pattern of an increase throughout disease progression (Figure 10), however, most of them display a large variability, which maybe the major reason why translation of a set of such parameters into diagnostic approaches have not proven sufficient robustness for valid predictions in human patients with chronic liver diseases. Count of CTGF-positive cells and expression of Il28b are among the few exceptions and their diagnostic potential is promising. The most correlated factor to CTGF-positive cell numbers is Tgf-β2 (Figure 6I). Both, Ctgf and Tgf-β2 encode excreted cytokines, and therefore represent good candidates as serum marker to monitor disease progression.

For the regulatory program, the largest change of quality of the disease progress occurs between days 2 and 5, and is characterized by large expression changes of factors, such as Il2, Il28b or Il13. This information has strong clinical relevance, as it indicates a switching point, and human homologs of the respective interleukines are top candidates to be used as clinical markers.

Our data suggest fibronectin (Fn1, Figure 6G) as the most sensitive marker for the onset of the disease process. Oncostatin M and its receptor are both highly correlated with the area of deposited collagen/ECM, suggesting it as a sensitive diagnostic marker for ongoing fibrogenesis.

Our experiments also found previously unknown molecular events which are probably elements of a transcriptional program. For instance, there are strong transcript peaks for SHP (Nr0b2) at 6h or transin-2 (Mmp10, Figure 6H) at 18h, which now need a focused analysis to delineate the chain of molecular

interactions causing it and the functional consequences for disease onset and progression. This is a potential starting point in therapy.

To conclude, a detailed time-resolved transcriptional profiling of liver homogenates following BDL revealed a coordinated induction of detoxification processes immediately after surgery and an up-regulation of an inflammatory response along with activation of metabolically active genes, which can be explained by physiological recovery and adaptation of the mice to the bile acid exposure.

Based on our results, we propose binary decision trees from selected serum parameters, such as in Figure 11C as a promising tool for bedside assessment of disease progression. Therefore, as next step, suitability of the parameters selected from the mouse model need to be confirmed for human patients as in [69]. We additionally suggest to identify common transcription factors (TF)/TF-binding sites controlling the transcription of those genes with the most significantly correlated expression.

Availability of supporting data

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

Dataset S1. Measurements in bile duct ligated mice.

Dataset S2. Statistical analysis.

List of Abbreviations

CTGF: connective tissue growth factor; BDL: bile duct ligation; SHP: small heterodimer partner; CLD: chronic liver disease; HSC: hepatic stellate cells; KC: Kupffer cells; ADME: absorption, distribution, metabolism, and excretion; ALT: alanine aminotransferase; GLDH: glutamate dehydrogenase; EDTA: ethylenediaminetetraacetic acid; H&E: hematoxylin and eosin;

Competing Interests

The authors declare that they have no conflicts of interest.

Author Contributions

Conceived and designed the experiments: KA AH MT IM BV SD. Performed the experiments: KA MT IM HW. Analyzed the data: KA MK AH MT IM ME HW HH UMZ BV SD. Wrote the paper: KA MK AH MT ME HW HH UMZ BV SD. All authors read and approved the final manuscript.

Acknowledgments

The authors kindly thank Dorothea Frenz, Berit Blendow, Maren Nerowski, Doris Butzlaff and Eva Lorbeer-Rehfeldt (Institute for Experimental Surgery, University of Rostock) and Igor Liebermann (Institute of Clinical Pharmacology, Stuttgart) for their excellent technical assistance. The authors thank Pablo Jaskowiak for helpful discussions. The study was supported by grants from the German Research Foundation to Kerstin Abshagen (AB 453/1-1), by the BMBF program “Virtual Liver” (Grants 0315755, 0315764) and by the Robert Bosch Foundation, Stuttgart, Germany. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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.

13. Leask A, Abraham DJ. All in the CCN family: essential matricellular signaling modulators emerge from the bunker. J Cell Sci. 2006;119:4803-10.
14. Gressner AM, Weiskirchen R, Breitkopf K, Dooley S. Roles of TGF-beta in hepatic fibrosis. Front Biosci. 2002;7:d793-807.
15. Gupta S, Stravitz RT, Dent P, Hylemon PB. Down-regulation of cholesterol 7alpha-hydroxylase (CYP7A1) gene expression by bile acids in primary rat hepatocytes is mediated by the c-Jun N-terminal kinase pathway. J Biol Chem. 2001;276:15816-22.
16. Russell DW. Nuclear orphan receptors control cholesterol catabolism. Cell. 1999;97:539-42.
17. Schacter BA, Joseph E, Firneisz G. Effect of cholestasis produced by bile duct ligation on hepatic heme and hemoprotein metabolism in rats. Gastroenterology. 1983;84:227-35.
18. Hattori T, Ohoka N, Hayashi H, Onozaki K. C/EBP homologous protein (CHOP) up-regulates IL-6 transcription by trapping negative regulating NF-IL6 isoform. FEBS Lett. 2003;541:33-9.
19. Hara M, Kono H, Furuya S, Hirayama K, Tsuchiya M, Fujii H. Interleukin-17A plays a pivotal role in cholestatic liver fibrosis in mice. J Surg Res. 2013;183:574-82.
20. Chen YX, Weng ZH, Zhang SL. Notch3 regulates the activation of hepatic stellate cells. World J Gastroenterol. 2012;18:1397-403.
21. Sancho-Bru P, Altamirano J, Rodrigo-Torres D, Coll M, Millán C, José Lozano J, et al. Liver progenitor cell markers correlate with liver damage and predict short-term mortality in patients with alcoholic hepatitis. Hepatology. 2012;55:1931-41.
22. Fujii T, Fuchs BC, Yamada S, Lauwers GY, Kulu Y, Goodwin JM, et al. Mouse model of carbon tetrachloride induced liver fibrosis: Histopathological changes and expression of CD133 and epidermal growth factor. BMC Gastroenterol. 2010.10:79.
23. Reister S, Kordes C, Sawitza I, Häussinger D. The epigenetic regulation of stem cell factors in hepatic stellate cells. Stem Cells Dev. 2011;20:1687-99.
24. Page S, Birerdinc A, Estep M, Stepanova M, Afendy A, Petricoin E, et al. Knowledge-based identification of soluble biomarkers: hepatic fibrosis in NAFLD as an example. PLoS One. 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.
67. Khan AJ, Choudhuri G, Husain Q, Parmar D. Polymorphism in glutathione-S-transferases: a risk factor in alcoholic liver cirrhosis. Drug Alcohol Depend. 2009;101:183-90.

68. Noureddin M, Wright EC, Alter HJ, Clark S, Thomas E, Chen R, et al. Association of IL28B genotype with fibrosis progression and clinical outcomes in patients with chronic hepatitis C: a longitudinal analysis. Hepatology. 2013;58:1548-57.
69. Gadd VL, Skoien R, Powell EE, Fagan KJ, Winterford C, Horsfall L, et al. The portal inflammatory infiltrate and ductular reaction in human nonalcoholic fatty liver disease. Hepatology. 2014;59:1393-1405.

Tables

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

Figure legends

Figure 1: Analysis of liver injury and-function. Plasma activities of alanine aminotransferase (ALT) (**A**) and glutamate dehydrogenase (GLDH) (**B**) and concentrations of plasma bilirubin (**C**) and albumin (**D**) at multiple time points after BDL. Values are given in means ± SEM of five independent experiments per time point.

Figure 2: Quantification of bile infarcts in H&E-stained liver sections at multiple time points after BDL (A). Values are given in means ± SEM of five independent experiments per time point. Representative H&E stainings of paraffin-embedded liver sections for each time point after BDL (**B**; arrows indicate bile lakes; magnification x10) with higher magnifications (x40) in (**C**), displaying cellular infiltrates (asterisk) and formation of artificial bile ductules (arrowhead).

Figure 3: Analysis of the proliferative and cellular response at multiple time points after BDL. Quantitative immunohistochemical analysis of BrdU-positive biliary epithelial cells (**A**), liver cells positive for α-SMA (**B**) and S100a4 (**C**), BrdU-positive hepatocytes (**D**) and Kupffer cells (**E**) and CTGF-positive cells (**F**). Values are given in means ± SEM of five independent experiments per time point. Corresponding representative immunohistochemical stainings are shown in the right panel (magnifications x40).

Figure 4: Analysis of proliferation and extracellular matrix accumulation. mRNA quantification of the proliferation marker Ki67 (**A**) by Fluidigm real-time PCR. Values are given in means ± SEM of five independent experiments per time point. Quantitative analysis of extracellular matrix deposition (**B**) and representative histological images (**C**; magnification x10) of Sirius red-positive areas at multiple time points after BDL. Values are given in means ± SEM of five independent experiments per time point.

Figure 5: Heat maps displaying gene expression pattern at multiple time points after BDL. Gene expression relative to Gapdh gene, obtained from Fluidigm qPCR, are shown as fold changes to sham mice (0 h) and are displayed in log2 scale. Red colour indicates up-regulation (2), blue down-regulation (-2) and

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

Figure 6: mRNA quantification of different selected genes by Fluidigm real-time PCR displayed in log2 scale. (**A**) Cyp1a2, (**B**) Cyp24a1, (**C**) Gstm1, (**D**) Nr0b2, (**E**) Col1α1, (**F**) Col3α1, (**G**), Fn1, (**H**) Mmp10, (**I**) Tgfb2, (**J**) Il2, (**K**) Il28b, (**L**) Tnfrsf1a. Values are given in means ± SEM of five independent experiments per time point.

Figure 7: Correlation between factors. Correlation matrix (YS3) for ANOVA filtered subset of significantly changed factors over time. Hierarchical clustering was performed using Complete Linkage with 6 resulting clusters depicted by colors. The side dendrogram depicts the full hierarchical clustering. Blue corresponds to positive correlation between factors, red to negative correlation based on YS3. Column colors correspond to the respective clusters, row colors for the respective type of factor (Fluidigm ADME panel, Fluidigm fibrosis panel; Fluidigm cytokine panel).

Figure 8: Correlations of histological, biochemical and antibody factors. YS3 correlation was calculated and the factors filtered to factors with a correlation of at least YS3>=0.6 or YS3<=-0.6. Top correlations between histological and biochemical markers and the gene expression data set. Positive correlation in blue, negative correlation in red with shade corresponding to strength of correlation. Area of the circles corresponds to YS3 correlation. Numerical values are provided in Supporting Information S1.

Figure 9: Time course clusters in BDL. Main clusters based on YS3 correlation with hierarchical clustering based on complete linkage. Cluster colors correspond to Figure 9. Mean cluster time course is depicted in blue with grey area corresponding to SD for the cluster members at the respective time points. Cluster members are enumerated for all classes with exception for the largest cluster, cluster 4 with N=61 members.

Figure 11: Decision trees. A) Regression tree for the prediction of time points where the latter shows the decision tree solely based on gene expression. The decision rules on the clusters and the best single factor representatives are depicted. The resulting time classes are shown with the information about the mean point, ranges of the classes and number of samples falling into the class (based on mean cluster data). **B-G**) Performance of regression tree. The predictive performance of the regression tree was evaluated using mean cluster data, all single factors from the individual clusters and a random sample of 2 factors from each cluster.

Figure 12: Outline of the disease process. Each box is dedicated to a specific disease aspect (first line) which is represented by a commonly known marker (second line) or several markers. Below (in small font) the genes are shown whose expression is correlated to the factor above.