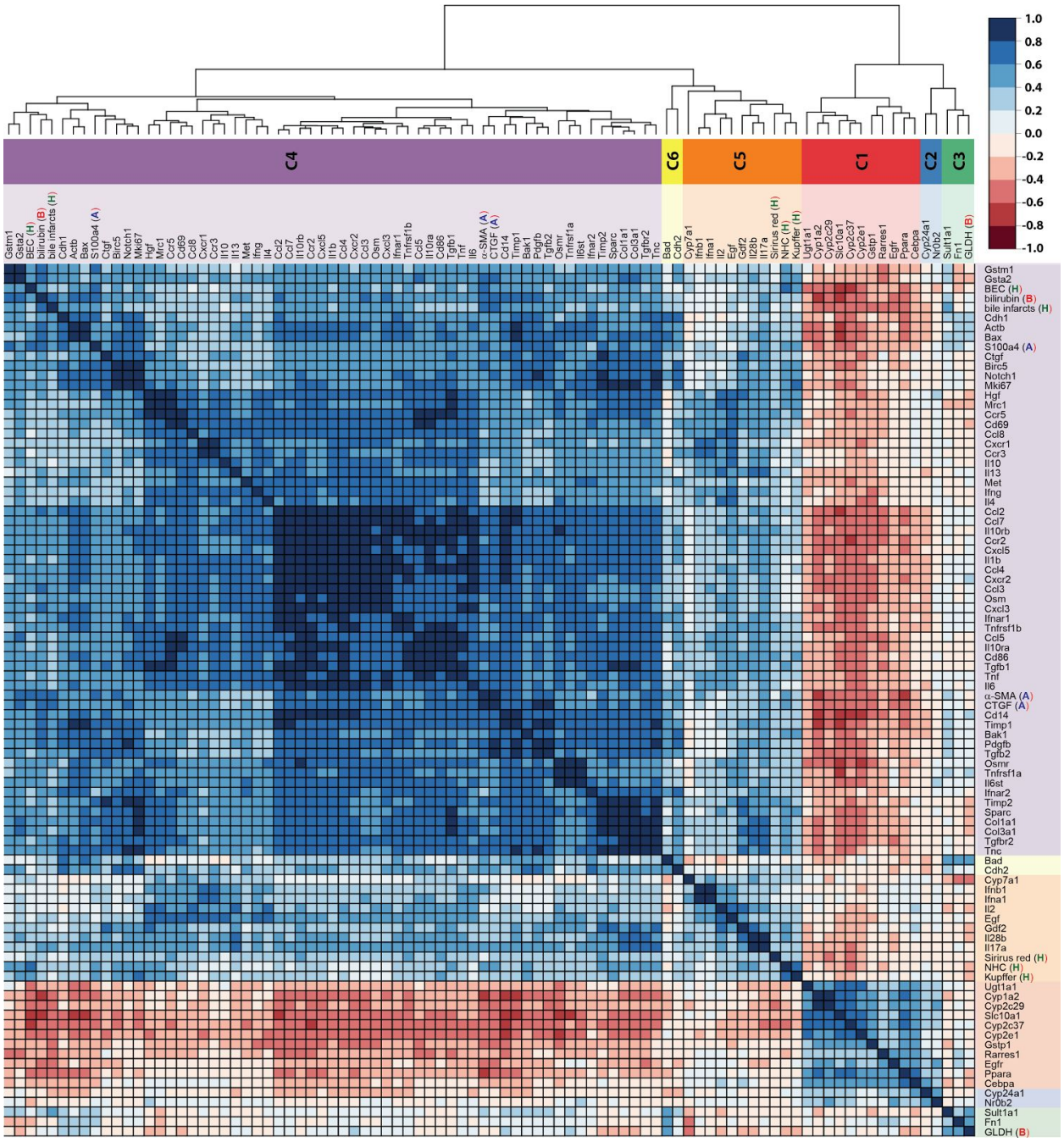

Figure 7: Correlation matrix of factors. Correlation matrix (YS3) of the subset of significantly changed factors determined by ANOVA. Positive correlation depicted in blue, negative correlation in red according to color key. Side dendrogram depicts the hierarchical clustering results with the 6 time course clusters marked in the color sidebar (see Figure 9 for the time courses corresponding to clusters c1-c6). Histological factors marked with **H**, immunostainings with **A**, and biochemical factors with **B**.

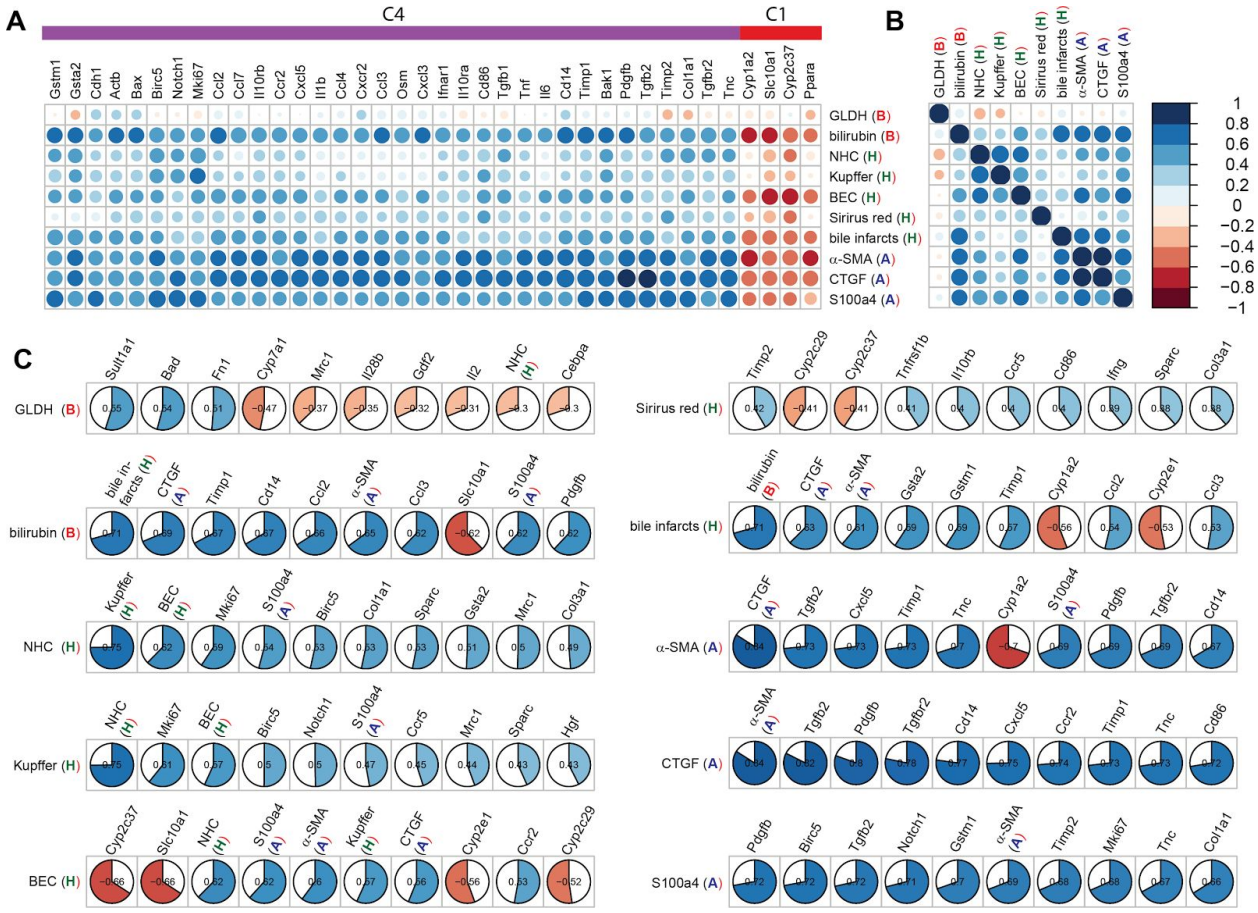

Figure 8: Histological (H), biochemical (B) and immunostaining (A) correlations. All correlations are YS3 correlation between the respective factors with numerical values provided in Supplement 2. **A** Correlation between histological, biochemical and antibody factors and gene transcripts. Only genes with at least one YS3 correlation of abs(YS3)>=0.6 are shown. Positive correlation in blue, negative correlation in red analog to Figure 7, with the area of circles corresponding to the correlation coefficients. **B** Correlation between histological, biochemical and antibody factors. **C** Highest absolute correlations between histological, biochemical and antibody factors and all other ANOVA filtered factors. Data sorted from left to right by absolute value of correlation. Color and size of the filled pie corresponds to the respective correlation value.

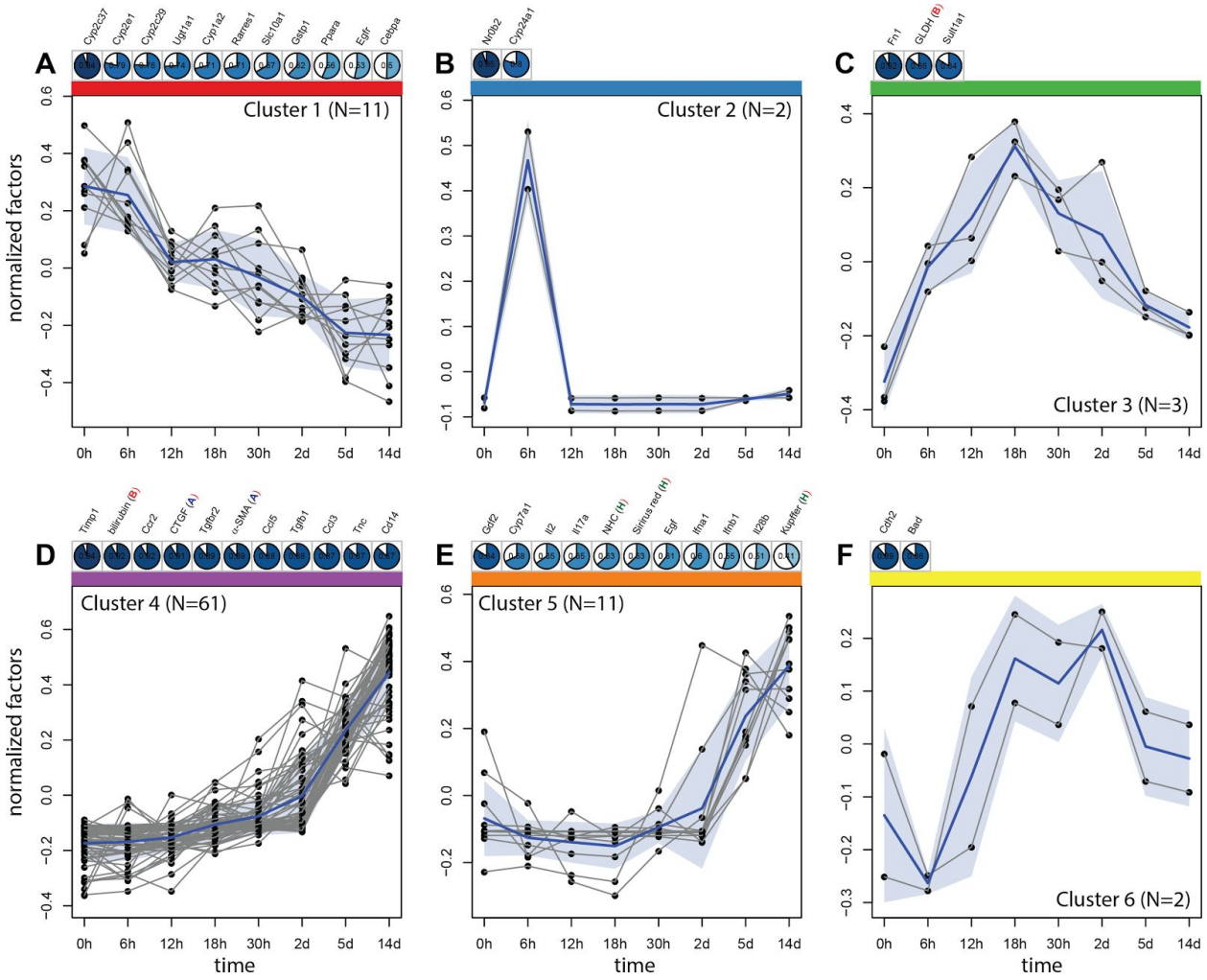

Figure 9: Time course clusters in BDL. Resulting time course clusters based on YS3 correlation with hierarchical clustering based on complete linkage corresponding to clusters in Figure 8. Mean cluster time course (averaged over all factors and repeats) depicted in blue, normalized factor representatives of cluster in grey. Shaded grey area corresponding to SD of all mean factor time courses. Top correlations based on YS3 between the factors in the cluster and the mean cluster time course are shown above the time course. Histological factors marked with **H**, immunostainings with **A**, and biochemical factors with **B**. Cluster members are fully enumerated in the figure for all classes with exception of cluster 4 with the members and respective correlation to mean cluster time course being: Timp1 (0.94), bilirubin (**B** 0.92), 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 (0.87), Cd14 (0.87), Ccl2 (0.86), Cd86 (0.86), Pdgfb (0.86), Col1a1 (0.86), Cxcl3 (0.86), Ccl4 (0.85), Cxcl5 (0.85), Il10ra (0.85), Col3a1 (0.85), Il10rb (0.84), Ccl7 (0.82), Cd69 (0.82), Ifnar1 (0.82), Tnf (0.82), Osm (0.81), Sparc (0.8), Il6 (0.8), Tnfrsf1b (0.8), Cxcr2 (0.78), Il1b (0.78), Timp2 (0.77), Ifnar2 (0.77), Ccr5 (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), Mrc1 (0.69), Tgfb2 (0.69), Ccr3 (0.68), Actb (0.68), S100a4 (**A** 0.66), Il13 (0.66), Met (0.66), bile infarcts (**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), Notch1 (0.54), Cxcr1 (0.51), Gstm1 (0.45), Cdh1 (0.42)

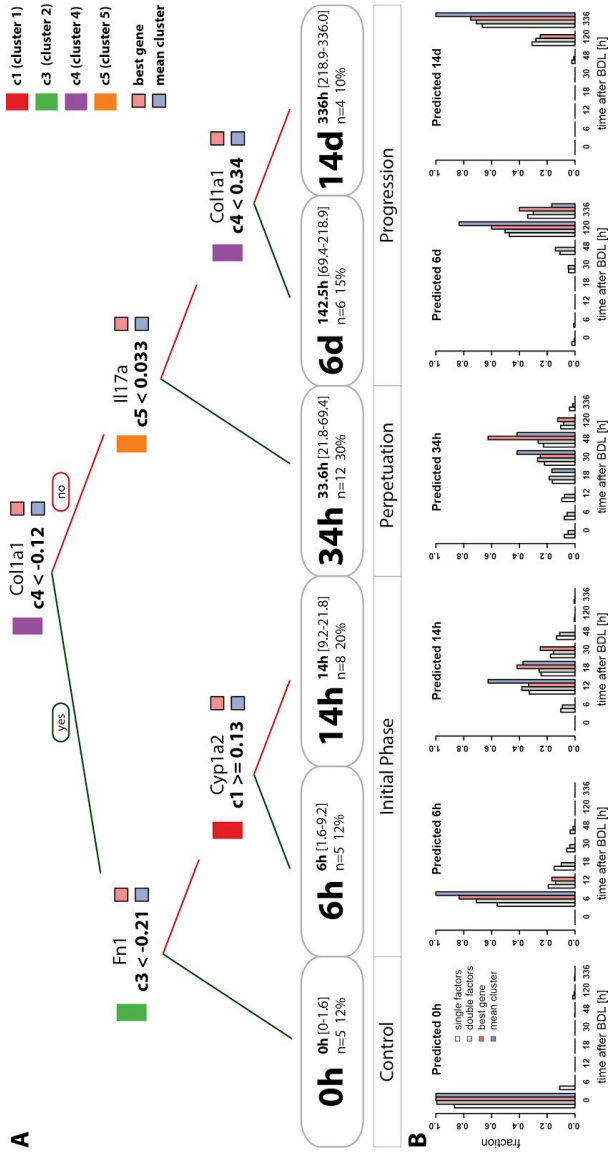

Figure 10: Decision tree for disease progression. A Regression tree for the prediction of time phases after BDL based on mean cluster time courses. The best gene representatives from the clusters are depicted above the decision rules. The regression tree classifies the data in 6 time classes 0h, 6h, 14h, 24h, 6d, 14d with information on mean time value, range, and number and percentage of samples falling given in the respective class. The best tree based on genes, histological, biochemical and antibody factors is highly similar to the best gene tree, with the single change of using S100a4 instead of Col1a1 for the decision on cluster c4. **B** Predictive performance of decision tree. The predictive performance of the regression tree was evaluated using all single factor combinations from the individual clusters (white), a random sample (N=10000) of two factors from each cluster (gray), the best gene representative tree (red), and the mean cluster data (blue, trainings data).

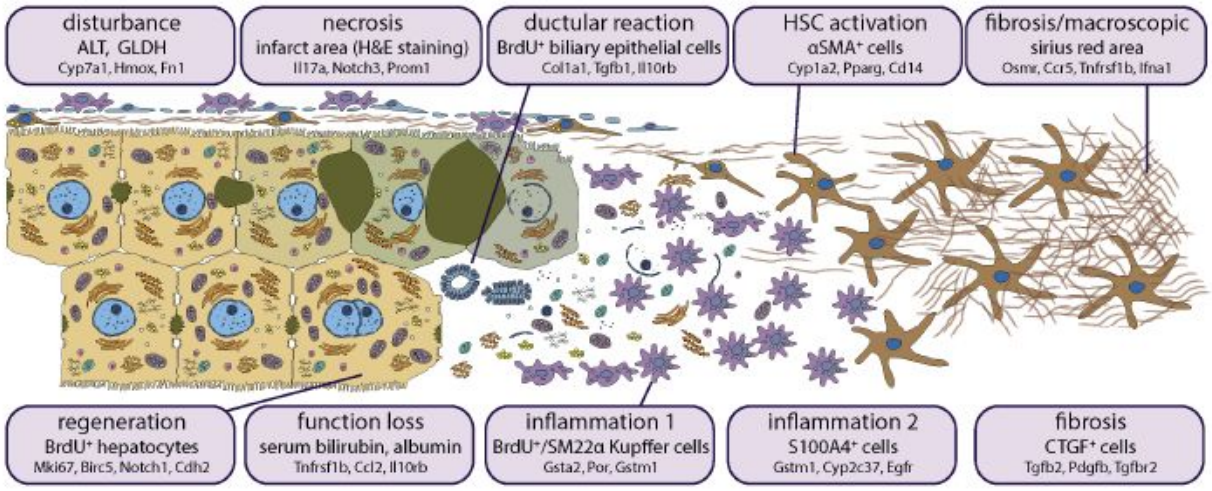

Figure 12: Outline of the disease process. Each box is dedicated to a specific disease aspect (first line) which is represented by a commonly known marker (second line) or several markers. Below (in small font) the genes are shown whose expression is correlated to the factor above.

TODO: Update the factors in the groups based on clustering results and correlation