

ANOVA results

Factor	p-value		p-value (Holm adjusted)	
Cyp1a2	1.91342264023676E-016	***	2.94667086596461E-014	***
bilirubin	3.27395955367953E-014	***	5.00915811712969E-012	***
Il10rb	7.63982399859513E-014	***	1.16125324778646E-011	***
Tgfb1	2.19724754091711E-013	***	3.31784378678484E-011	***
Ccl2	2.32001535448977E-013	***	3.48002303173466E-011	***
Cd86	4.43398094033834E-013	***	6.60663160110413E-011	***
Ccr2	4.7207988695294E-013	***	6.98678232690351E-011	***
Mrc1	4.76201704229549E-013	***	7.00016505217438E-011	***
Tnfrsf1b	5.44303026552875E-013	***	7.94682418767197E-011	***
Cxcl5	4.34564483500159E-012	***	6.3011850107523E-010	***
CTGF	5.46635298072514E-012	***	7.8715482922442E-010	***
Il10ra	1.15477077793218E-011	***	1.65132221244301E-009	***
Gstm1	6.51380591994191E-011	***	9.24960440631751E-009	***
Ccl7	2.43088335234461E-010	***	3.4275455268059E-008	***
Ccr5	3.13926202190641E-010	***	4.39496683066897E-008	***
Hgf	4.20334360409302E-010	***	5.8426476096893E-008	***
Osmr	7.38949941761375E-010	***	0.000000102	***
Ccl4	7.67090680670405E-010	***	1.05091423251846E-007	***
Nr0b2	0.000000001	***	1.30672786047445E-007	***
Tgfb2	1.27699003792023E-009	***	1.72393655119231E-007	***
BrdU_HSC	1.56963992826678E-009	***	2.10331750387749E-007	***
Ccl5	2.10407584924822E-009	***	2.79842087950013E-007	***
Col1a1	3.35663234495581E-009	***	4.43075469534167E-007	***
Ifnar1	5.93309462206502E-009	***	7.77235395490518E-007	***
Sl100A4	7.12298809471991E-009	***	0.000000926	***
Sparc	8.38363234548618E-009	***	1.08148857256772E-006	***
Cyp2e1	1.18106363259919E-008	***	1.51176144972696E-006	***
Cxcr2	1.39031297789528E-008	***	1.76569748192701E-006	***
Ccr3	1.52418929058291E-008	***	1.92047850613446E-006	***
Cd69	0.000000022	***	2.75034504981931E-006	***
Cyp2c29	2.38412801732629E-008	***	2.9563187414846E-006	***
Gsta2	3.18037362655874E-008	***	3.91185956066725E-006	***
Tnf	0.000000036	***	0.000004395	***
Gdf2	4.95804294192854E-008	***	5.99923195973354E-006	***
Il1b	6.12872857427699E-008	***	7.35447428913239E-006	***
Ifng	6.36729003696979E-008	***	0.000007576	***
Osm	6.36637501453233E-008	***	0.000007576	***
Ccl3	7.3549197590383E-008	***	8.60525611807481E-006	***
Il13	8.08157001605371E-008	***	9.37462121862231E-006	***
Cxcr1	9.41283110829526E-008	***	1.08247557745396E-005	***
Cyp2c37	9.65026773866885E-008	***	1.10013052220825E-005	***
Cd14	1.0281684108002E-007	***	1.16183030420422E-005	***
Col3a1	1.8165609631441E-007	***	2.03454827872139E-005	***
Tnfrsf1a	3.02825864088077E-007	***	3.36136709137765E-005	***
Il2	4.43711365153111E-007	***	4.88082501668422E-005	***
Ifnb1	4.50419343820152E-007	***	4.90957084763966E-005	***
Egf	4.55680005521404E-007	***	4.92134405963116E-005	***
Il4	4.60199593762361E-007	***	4.92413565325726E-005	***
Il28b	4.64575161516398E-007	***	0.000049245	***
Il10	4.6663710756229E-007	***	0.000049245	***
Slc10a1	5.10222512673073E-007	***	5.30631413179996E-005	***
Timp2	6.11614358775742E-007	***	6.29962789539014E-005	***
Cxcl3	0.0000000675	***	6.88516389428801E-005	***
Ccl8	1.23279663549998E-006	***	0.0001245125	***
Ctgf	1.33622451654719E-006	***	0.0001336225	***
Gstp1	1.40917342020316E-006	***	0.0001395082	***
Ppara	1.68310618760453E-006	***	0.0001649444	***
Ifnar2	1.89821207406025E-006	***	0.0001841266	***
Il6	2.35209617540156E-006	***	0.0002258012	***
Il17a	2.58037790633237E-006	***	0.0002451359	***
Bad	4.12522559223528E-006	***	0.0003877712	***
Timp1	4.67845283519821E-006	***	0.0004350961	***
Cdh1	0.00000492	***	0.0004526391	***
Cebpa	5.47343865919207E-006	***	0.0004980829	***
alpha.SMA	0.000005638	***	0.0005074204	***
BrdU_NHC	0.000006005	***	0.0005344464	***
Cdh2	6.17914742457919E-006	***	0.000543765	***
BrdU_SiriusRed	8.81983941044958E-006	***	0.000767326	***
Pdgfb	9.03245409952277E-006	***	0.0007767911	***
Il6st	1.03707632879727E-005	***	0.0008815149	***
Fn1	1.45776163221202E-005	***	0.0012245198	**
Mki67	1.77462324214441E-005	***	0.0014729373	**
Ifna1	1.80502947443697E-005	***	0.0014801242	**
Egfr	1.98831180442088E-005	***	0.0016105326	**
BrdU_Kupffer	2.47061167978283E-005	***	0.0019764893	**

Tnc	2.61121496644436E-005	***	0.0020628598	**
Ugt1a1	3.57068529161741E-005	***	0.0027851345	**
Sult1a1	3.92227409891161E-005	***	0.0030201511	**
GLDH	5.71185646136636E-005	***	0.0043410109	**
Notch1	0.000060164	***	0.0045123018	**
Met	6.80346774563997E-005	***	0.0050345661	**
Cyp7a1	0.00013651	***	0.0099652314	**
Cyp24a1	0.0001391831	***	0.0100211829	*
Tgfb2	0.0001542364	***	0.010950784	*
Birc5	0.0002454637	***	0.0171824558	*
Actb.y	0.000280664	***	0.019365815	*
Bak1	0.0004077548	***	0.0277273274	*
Bax	0.0004371327	***	0.0292878876	*
Rarres1	0.0005491381	***	0.0362431116	*
bileInfarcts	0.0005904178	***	0.0383771539	*
Cyp3a11	0.0008410363	***	0.0538263241	0
Sult1b1	0.0009257764	***	0.058323915	0
Cyp4a10	0.0011244464	**	0.069038815	0
Pparg	0.0011135293	**	0.069038815	0
Hk2	0.0012711725	**	0.0762703529	0
ALT	0.0015159716	**	0.0894423215	0
Smad6	0.0016656545	**	0.0966079586	0

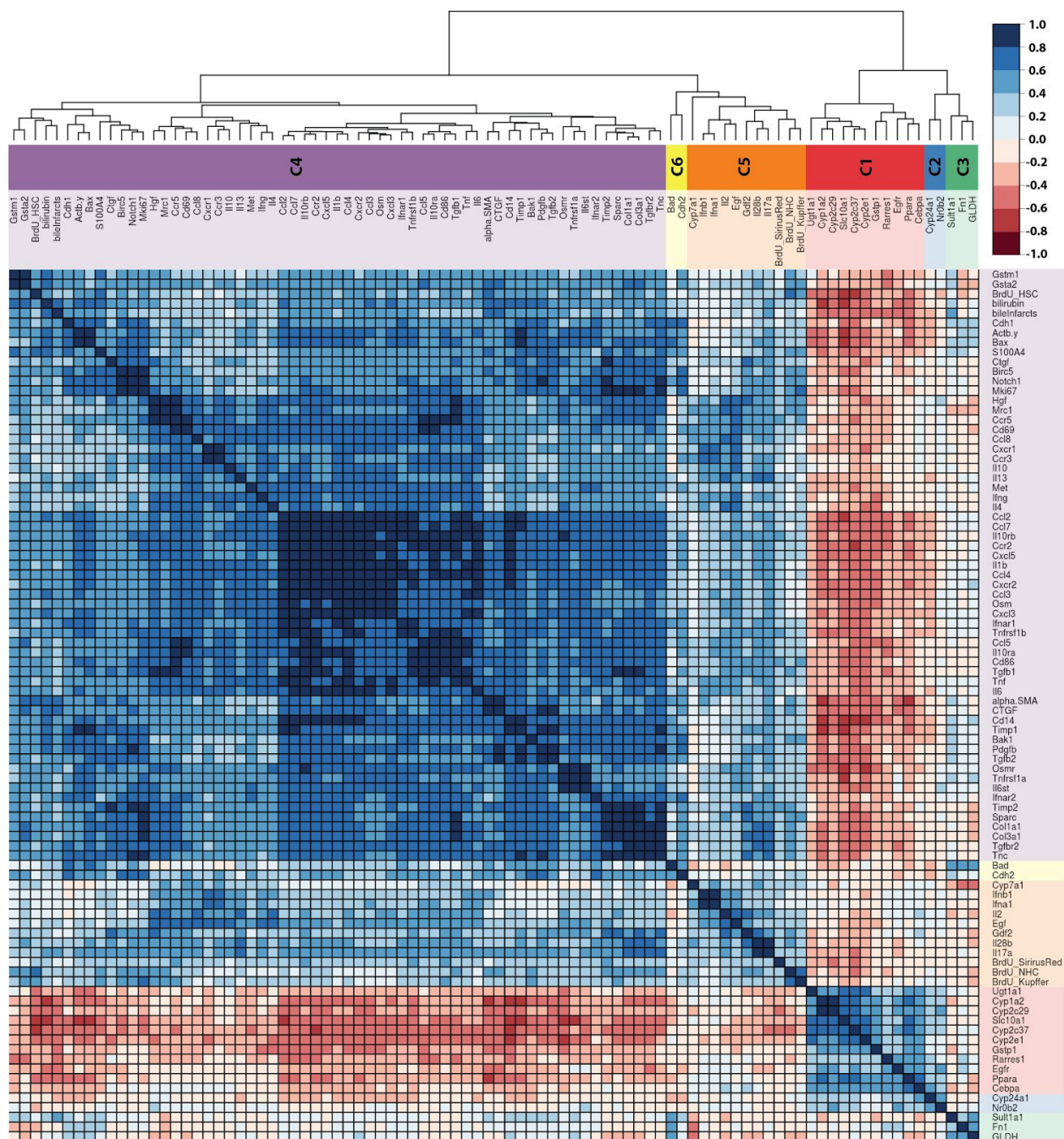
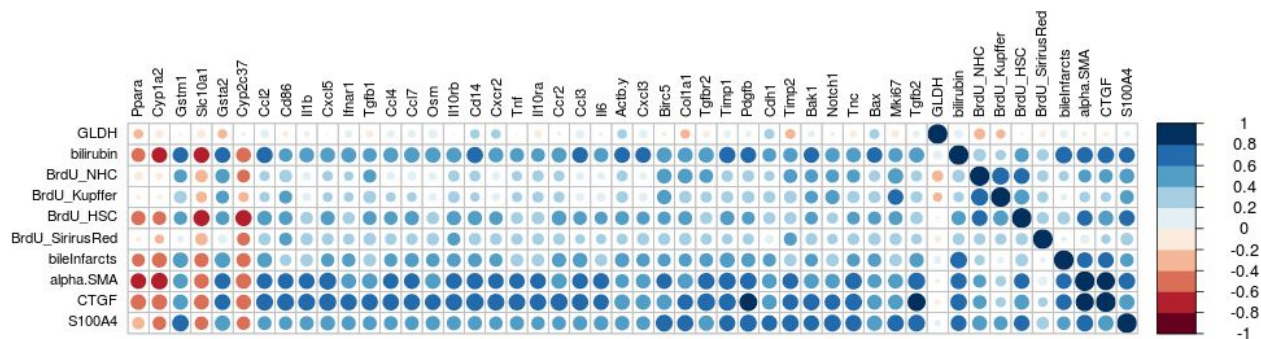


Figure 7: Correlation matrix of factors. Correlation matrix (YS3) of the subset of significantly changed factors over time (determined via ANOVA). Positive correlation depicted in blue, negative correlation in red according to color key. Side dendrogram depicts results of hierarchical clustering with the 6 time course clusters marked in color sidebar (see Figure 9 for respective time courses of clusters c1-c6).

A



B

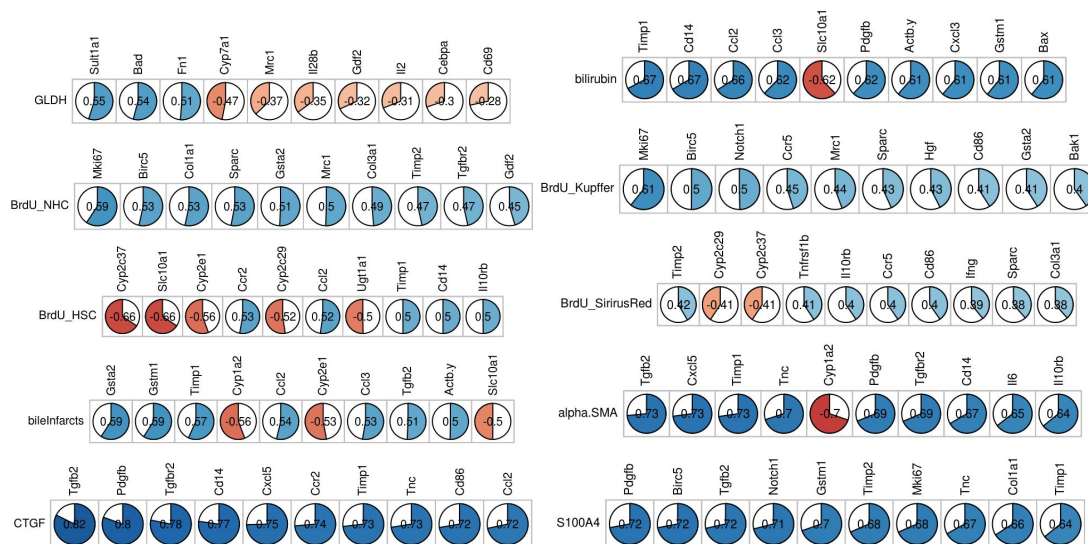


Figure 8: Correlations of histological, biochemical and antibody factors. A Correlation between histological, biochemical and antibody factors and gene expression data. YS3 correlation was calculated between histopathological factors (rows) and the subset of ANOVA filtered factors (data corresponds to a subset of Figure 7). Only columns with at least one YS3 correlation of $\text{abs}(\text{YS3}) \geq 0.6$ are shown. Positive correlation in blue, negative correlation in red analog to Figure 7, with the area of circles corresponding to the correlation coefficients. Numerical values are provided in Supporting Information S2. **B** Highest absolute correlations between the individual histological, biochemical and antibody factors and the gene expression data. Data sorted from left to right by $\text{abs}(\text{YS3})$. Color and size of the filled pie correspond to the respective correlation.

TODO: add colors for the factor type

TODO: sort by the order of the clusters (i.e. use the hierarchical clustering on ys3 for the sorting)

TODO: Adobe: nice layout + add continuous color legend

TODO: add cluster colors to the individual entries

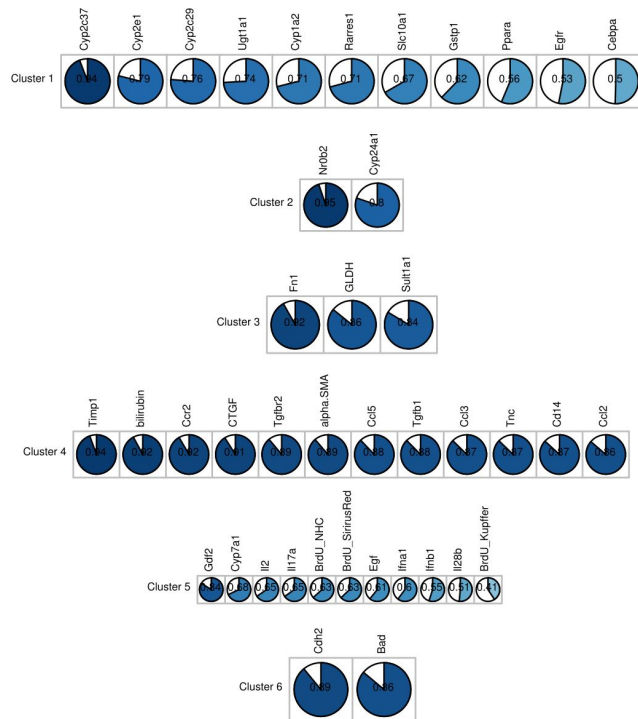
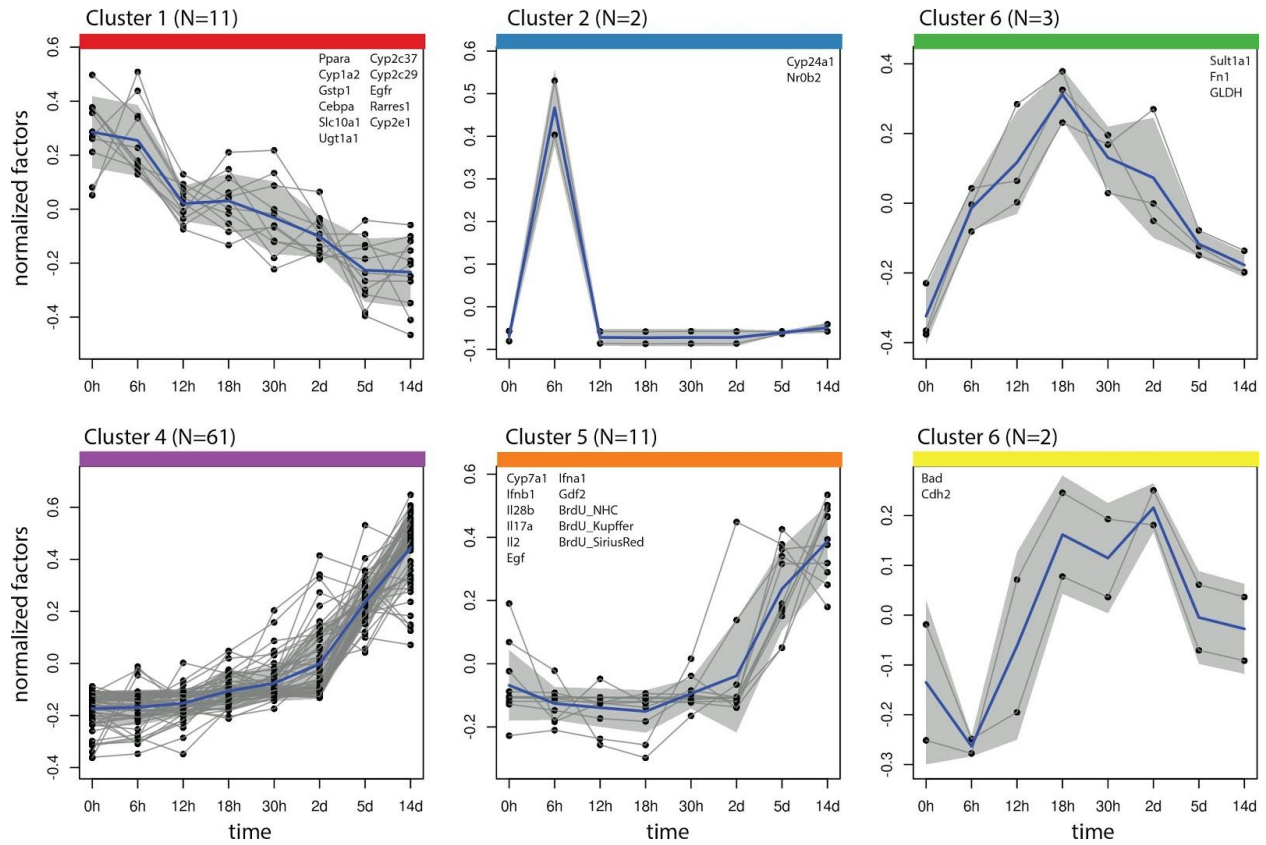


Figure 9: Main time course clusters in BDL. The 6 top time course clusters based on YS3 correlation with hierarchical clustering based on complete linkage. Cluster colors analog to Figure 7. The mean cluster time course (averaged over all factors and repeats) is depicted in blue with grey area corresponding 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 (Gstm1, Gsta2, Ccl2, Cd86, Met, Tnfrsf1a, Il1b, Cxcl5, Cxcr1, Ifnar1, Osmr, Tgfb1, Ifng, Ccl4, Ccl5, Ccr3, Il13, Tnfrsf1b, Ccl7, Osm, Cd69, Il10rb, Cd14, Cxcr2, Tnf, Il10ra, Il10, Ccr2, Hgf, Ifnar2, Mrc1, Ccr5, Ccl3, Il6, Actb.y, Il4, Ccl8, Cxcl3, Il6st, Birc5, Sparc, Col1a1, Tgfb2, Timp1, Pdgb, Cdh1, Timp2, Bak1, Ctgf, Notch1, Tnc, Bax, Mki67, Tgfb2, Col3a1, bilirubin, BrdU_HSC, bileInfarcts, alpha.SMA, CTGF, S100A4)

TODO: add the legend to figure (what is blue, what is grey)

TODO: better color (gray & gray not good)

TODO: add A-F markers for clusters

TODO: best representatives

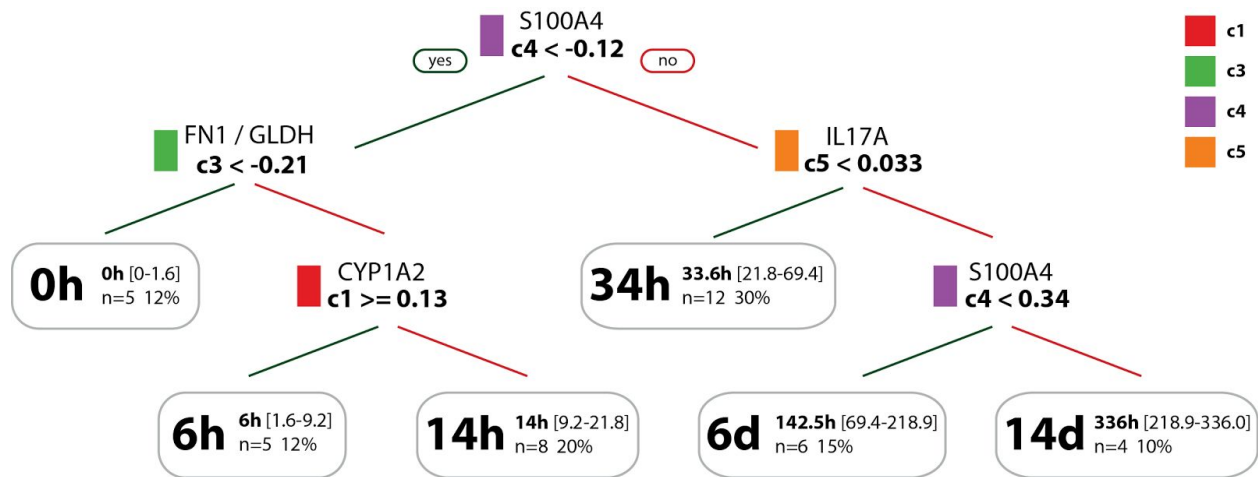


Figure 11: Decision trees based on clusters. A) Regression tree for the prediction of time phases after BDL based on the correlation time course clusters. The decision trees for the mean clusters and the best tree using single representatives from the clusters are depicted. The resulting time phases (0h, 6h, 14h, 24h, 6d, 14d) are depicted with the respective information about mean point, ranges and number of samples falling into the class (based on mean cluster data).

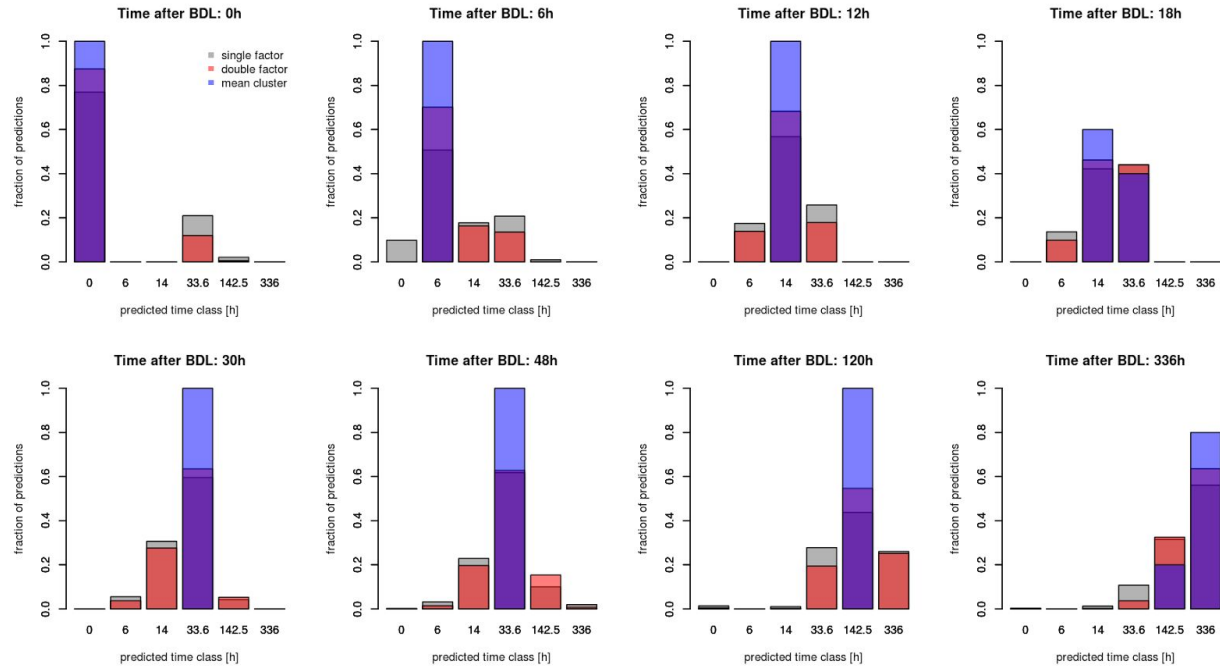


Figure 12 Predictive performance of decision tree. The predictive performance of the regression tree was evaluated using mean cluster data, all single factors from the individual clusters and a random sample of 2 factors from each cluster. In addition the results for the best single factor tree (depicted in Figure 10) are shown.

TODO: barplot next to each other instead of overlay

TODO: add best single factor combination & best single factor & best single factor only using RNA information (i.e. the best trees)

TODO: add A-G numbers

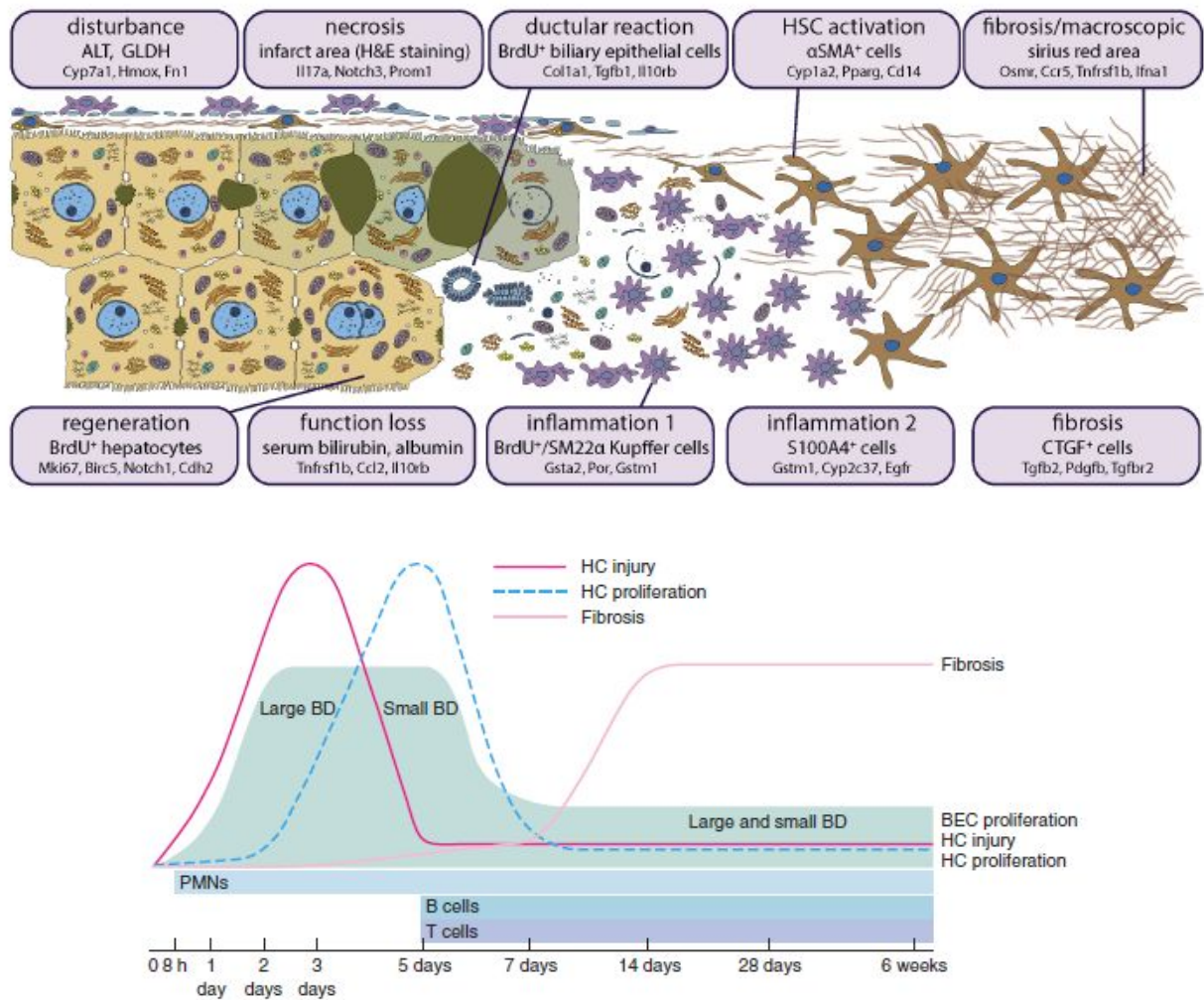


Fig. 9 Overview of dynamic changes following bile duct ligation in mice. BD, bile duct; BEC, biliary epithelial cell; HC, hepatocyte; PMN, polymorphonuclear leucocyte

{Georgiev2008}

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 (not necessarily the same results for all)