# Cholestasis
## Analyzing a bile duct ligation experiment

Andreas Hoppe
Institut für Biochemie
Charité – Universitätsmedizin Berlin

# Outline

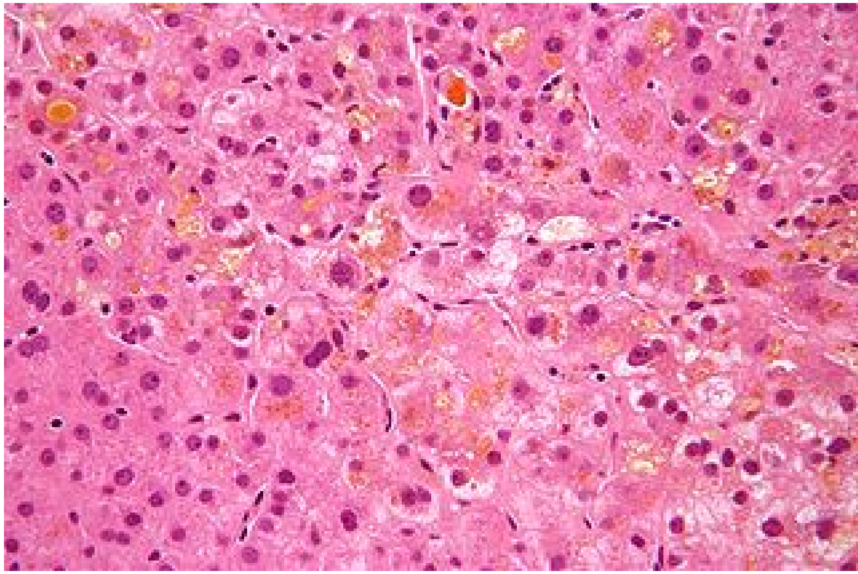
- Cholestasis
- Bile duct ligation experiment
- Correlation analysis
- Marker candidates
- Predictors of a time frame
- Decision trees for progress prediction

# Outline

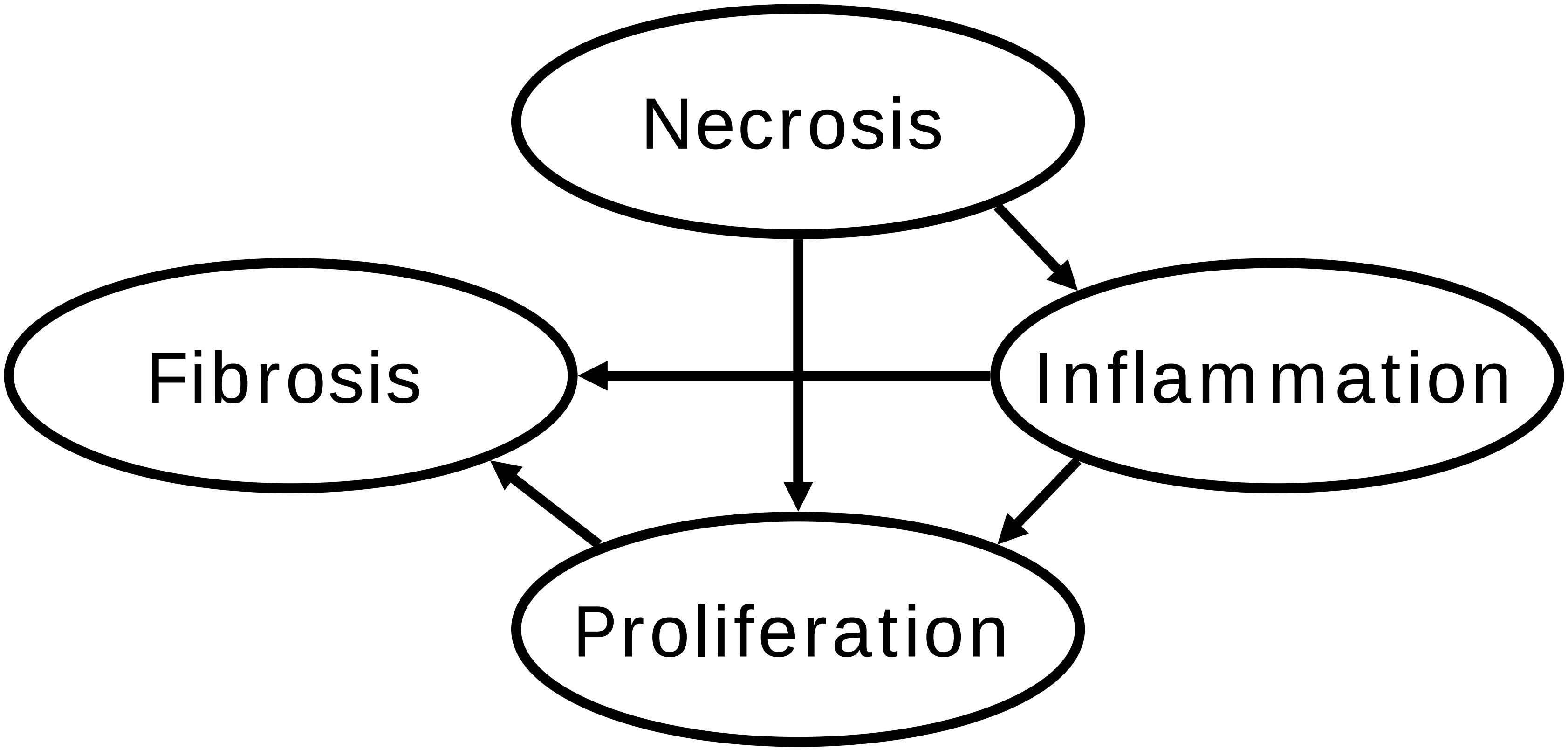
- Cholestasis
- Bile duct ligation experiment
- Correlation analysis
- Marker candidates
- Predictors of a time frame
- Decision trees for progress prediction

# Cholestasis

- Two types
  - obstructive: mechanical blockage
    - gallstone
    - malignancy
  - metabolic: disturbances in bile formation
    - side effect of many medications
    - genetic defects
- Bile stipples in hepatocytes
- Canalicular bile plugs
- Bile duct rupture
  - hepatic necrosis
  - bile lakes



# Processes



Necrosis

Fibrosis

Inflammation

Proliferation

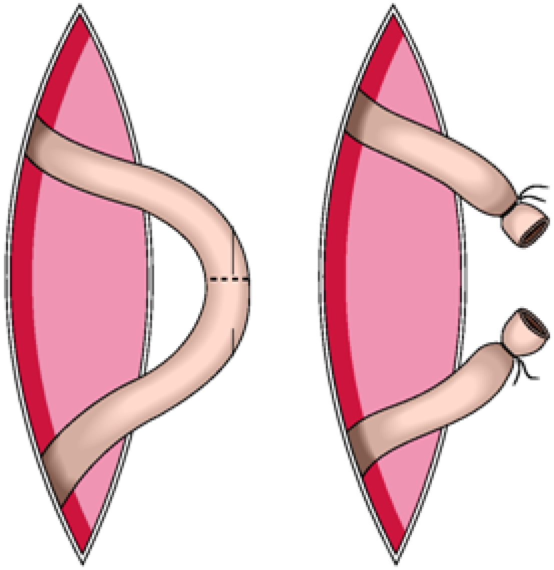
# Outline

- Cholestasis
- Bile duct ligation experiment

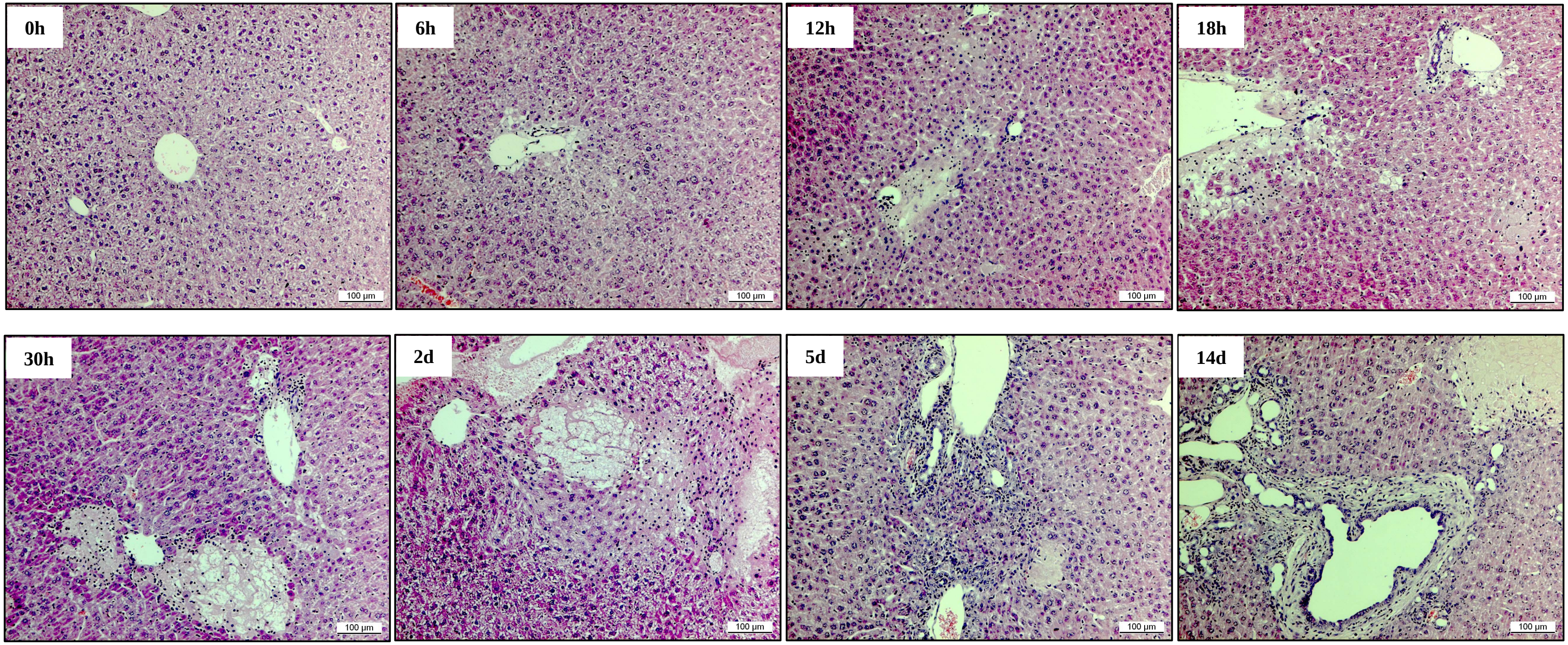
- Correlation analysis
- Marker candidates
- Predictors of a time frame
- Decision trees for progress prediction

# Bile duct ligation

- Mouse experiment
- Main bile duct is closed
- 7 Time points 6h-14d
- 5 repeats (3 for some factors)
- Mouse sacrificed
  - histological images
  - cell separation
  - serum concentrations
  - RNA

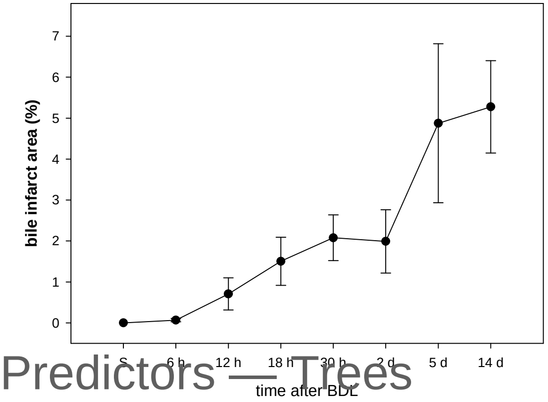


# Hematoxylin and eosin stain highlighting living cells

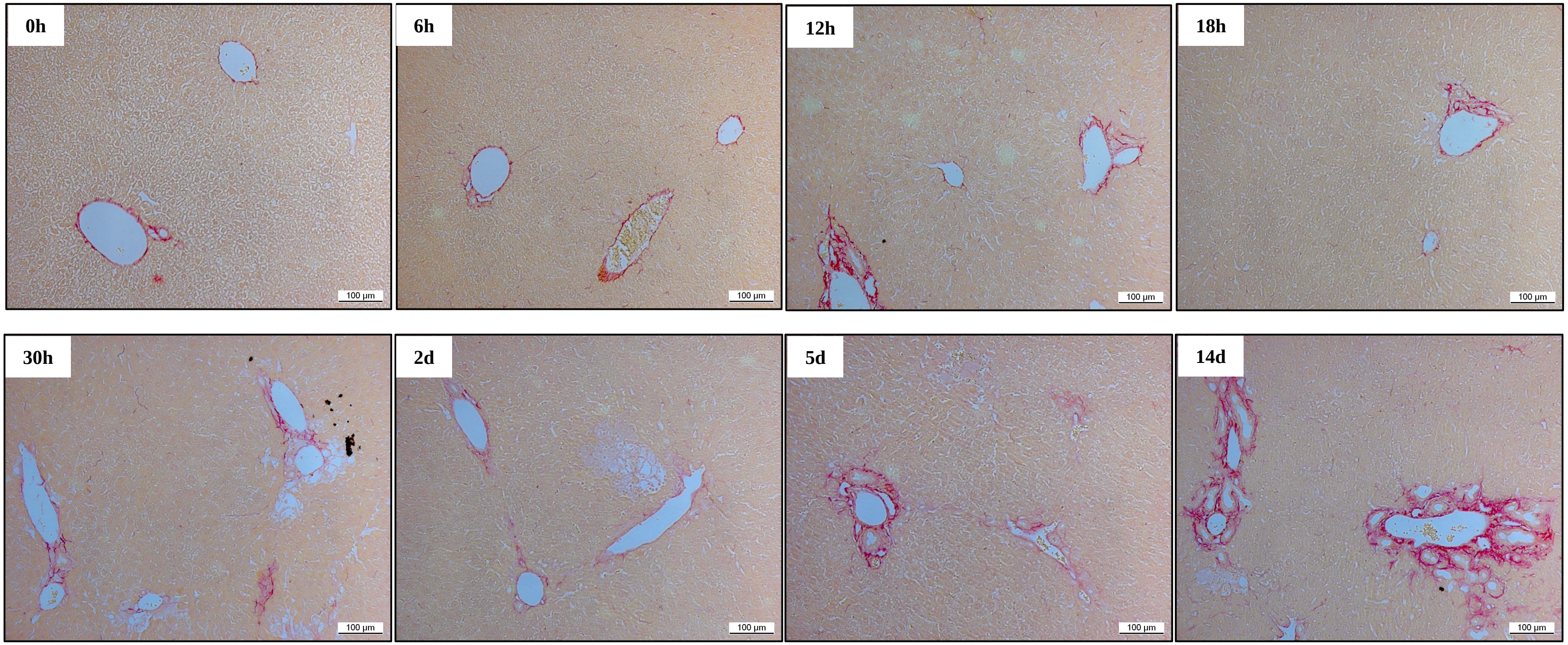


HE staining in BDL mice (100X)

- Area of dead tissue estimated

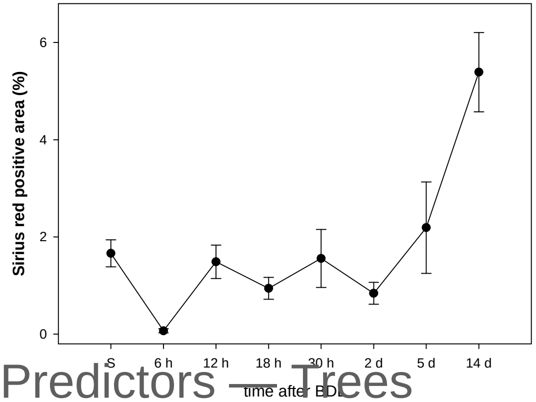


# Collagen marked by Sirius red

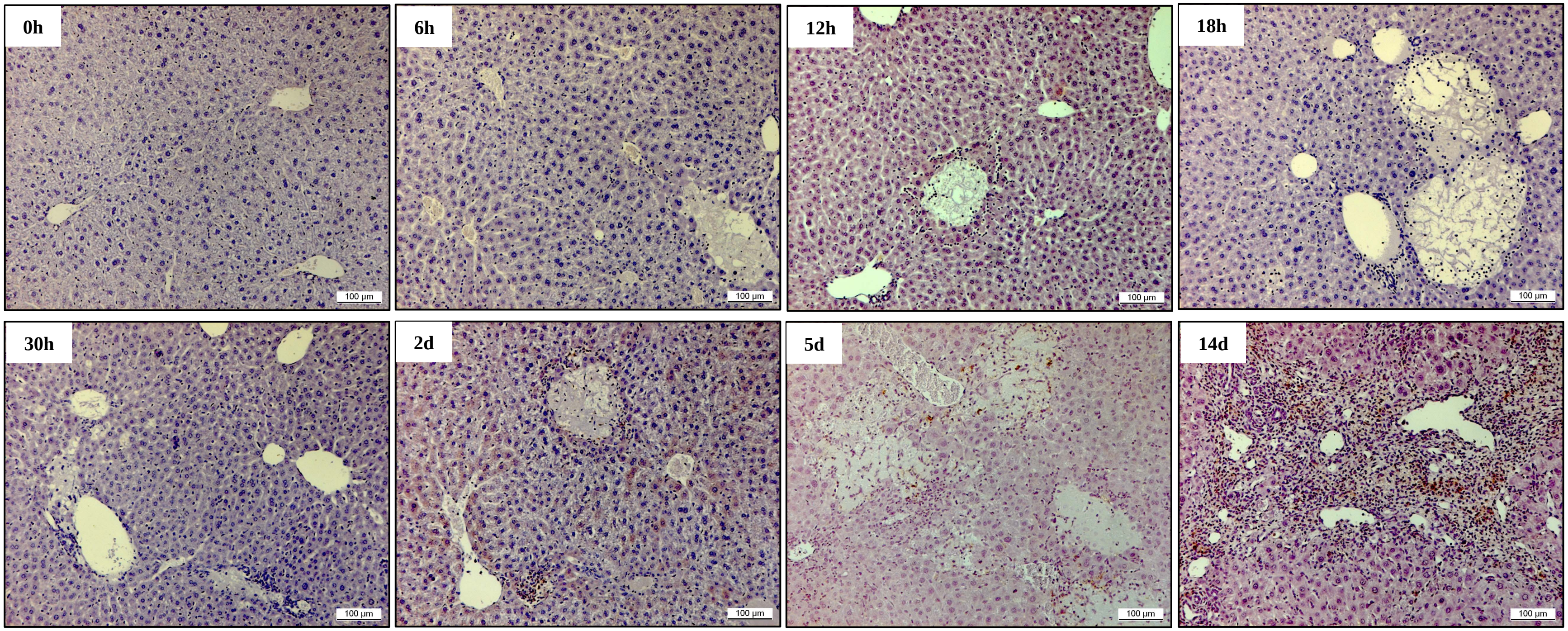


Sirius red staining in BDL mice (100X)

- ○ Area of fibrotic tissue measured

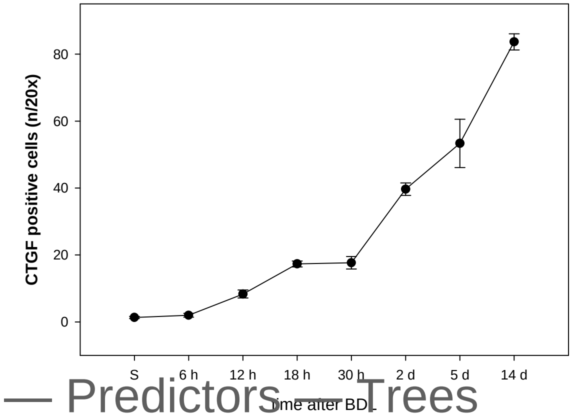


# Connective tissue growth factor

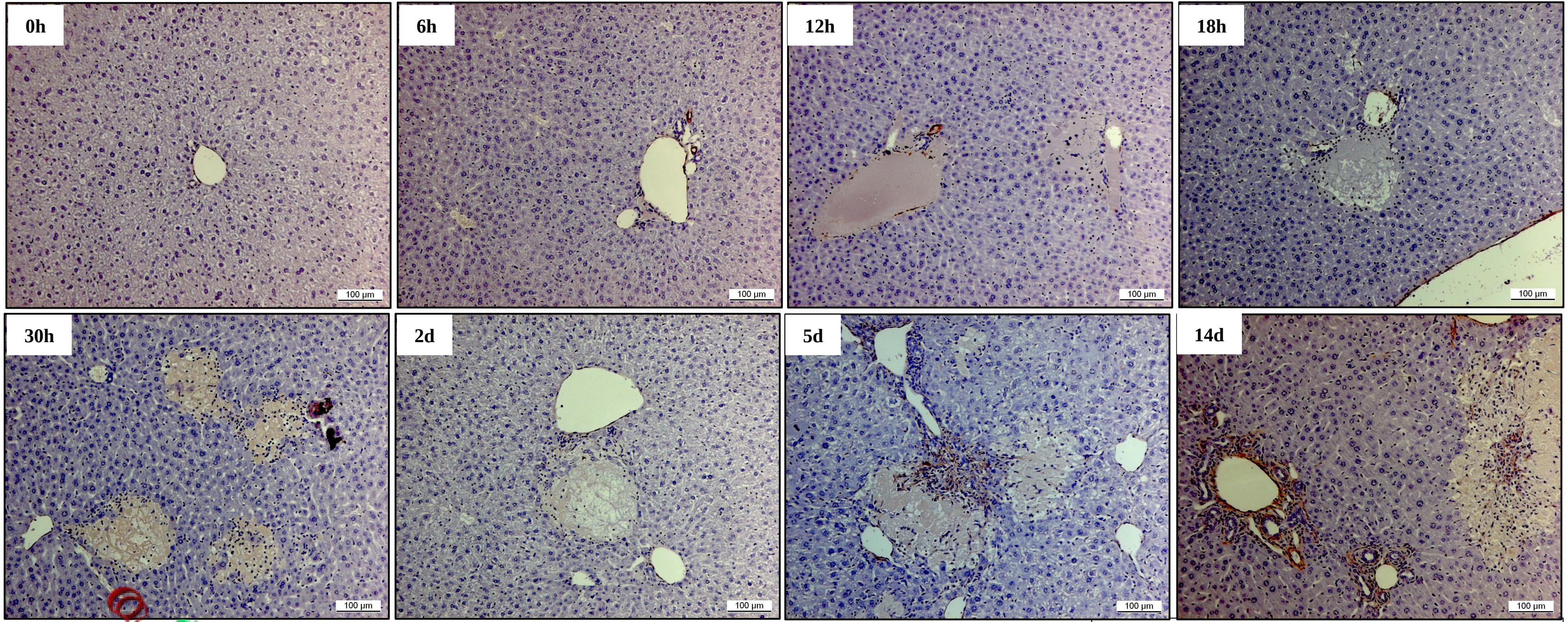


CTGF staining in BDL treated mice

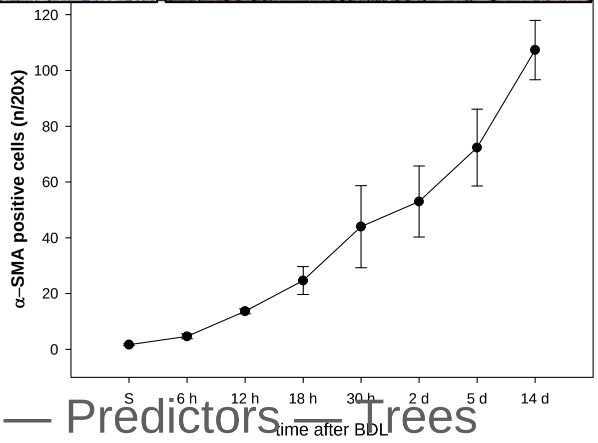
- associated with wound healing and fibrotic pathology
- cooperates with TGF-$\beta$ to induce sustained fibrosis



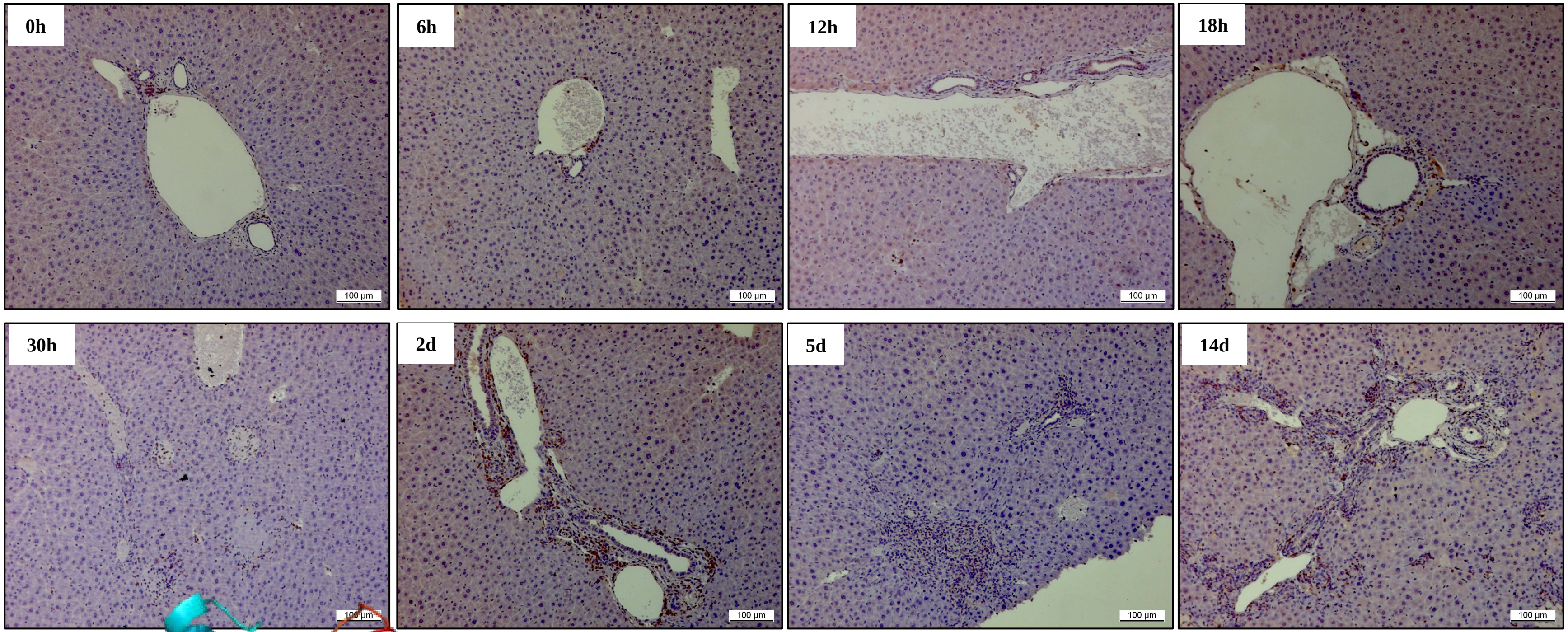
# Alpha smooth muscle actin, marker of myofibroblast formation



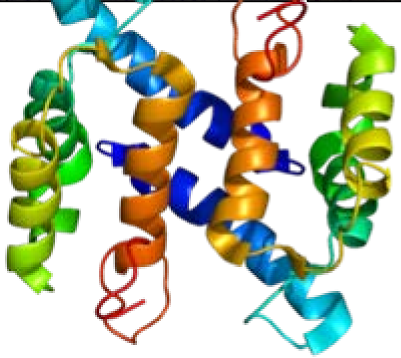
α-SMA staining in BDL mice (X100)

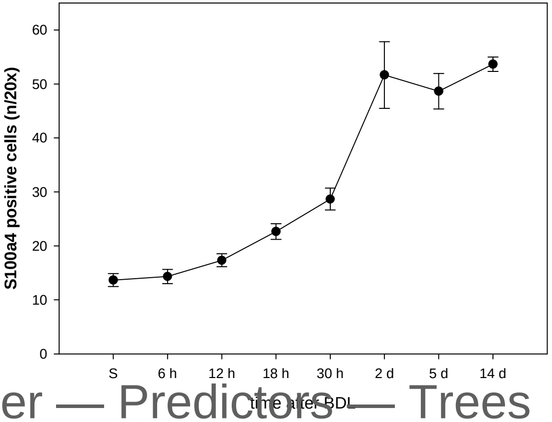


# S100 calcium binding protein A4, fibroblast-specific marker



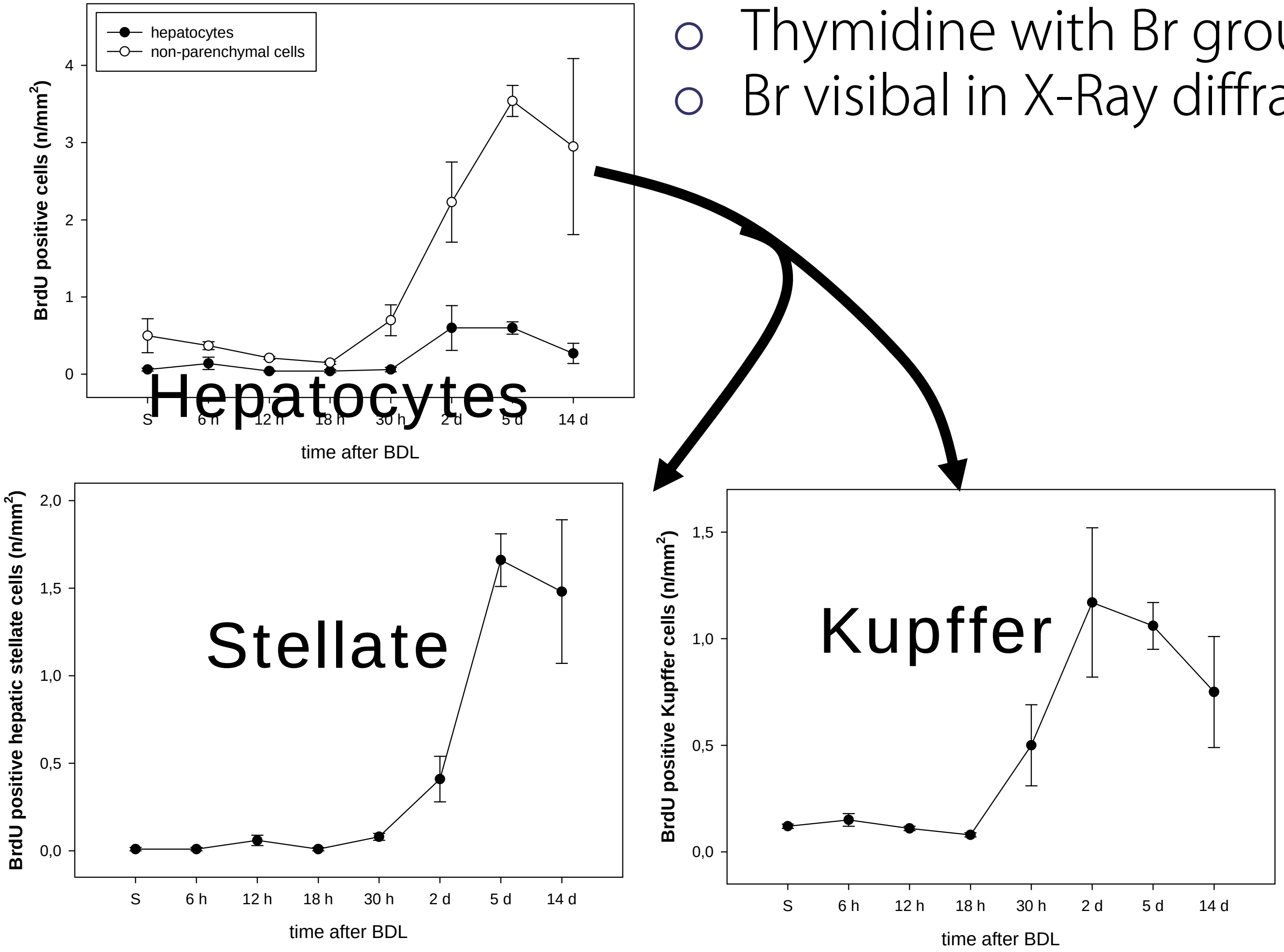
S100A4 staining in BDL treated mice





# BrdU proliferation marker for hepatic cell types

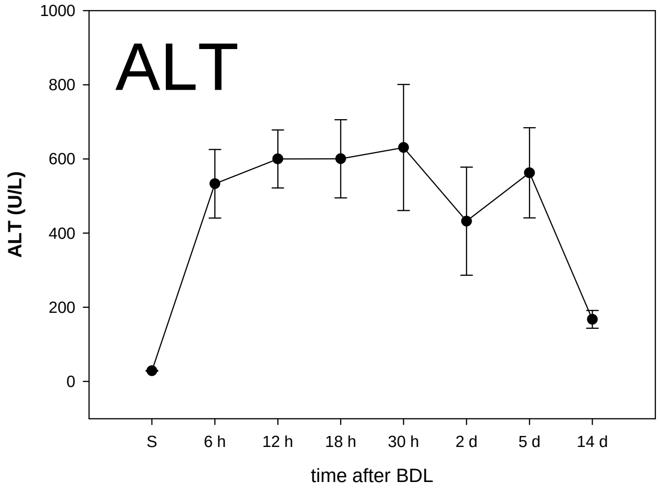
- Thymidine with Br group for $CH_3$
- Br visibal in X-Ray diffraction

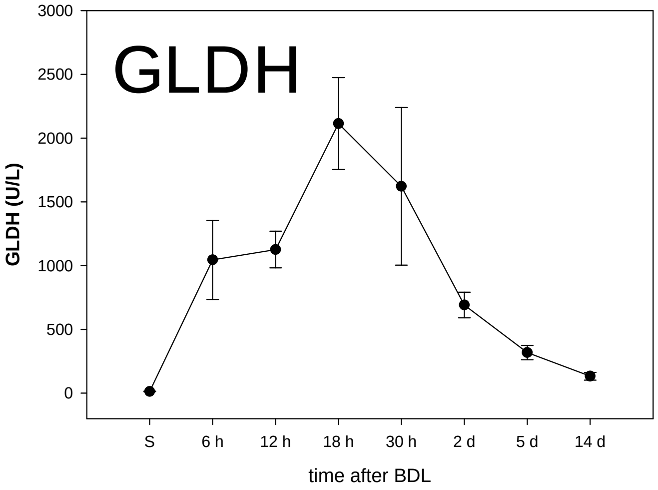


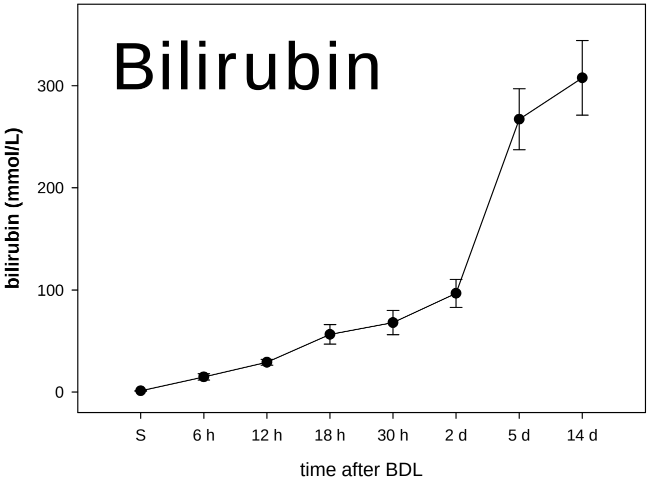
# Fluidigm quantitative PCR

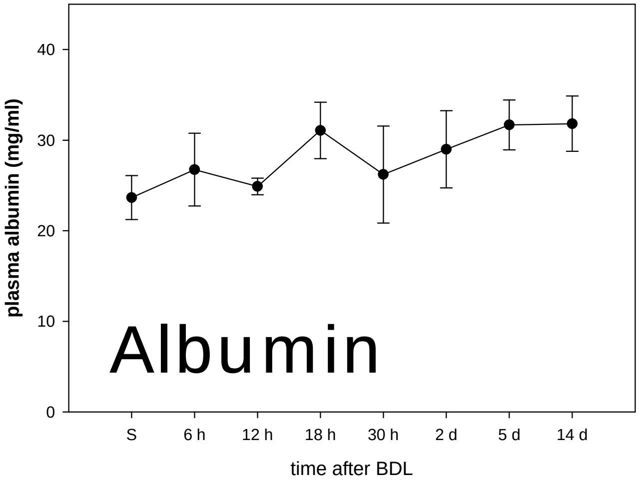
- 3 panels, 141 genes, by Zanger group (Stuttgart)
  - ADME: cytochromes p450 etc.
  - cytokines: e.g. TGFβ, receptors, etc.
  - fibrosis: collagens, differentiation factors, etc.
- relative to GAPDH

# Serum markers: ALT GLDH Bilrubin Albumin









# Outline

- Cholestasis
- Bile duct ligation experiment
- Correlation analysis
- Marker candidates
- Predictors of a time frame
- Decision trees for progress prediction

# 1. Correlation of time averages

- Advantage
  - shows correlations in disease progression
- Time averages (5 repeat mice) huge variance
  - Infarct area at 5d: 0.9...12%
  - Collagen area at 5d: 0.8...5.9%
- Hypothesis: mice are different
  - for mechanistic correlations: better single mice!

# 2. Correlations of all mice

- Advantage:
  - more likely to exhibit mechanistic correlations
- Disadvantage:
  - Disregard of time
  - Unbalance:
    - many factors roughly near base level for majority of points

# 3. Correlations for mice of single time points

- Advantage
  - shows correlations only for similar mice
- Disadvantage
  - only 5 (for some factors 3) data points
  - low robustness
  - low significance

# 4. Correlations of phases:

- Phases according to clinical symptoms
  - initial (6-12h)
  - perpetuation (18h-2d)
  - progression (5-14d)
- Advantage
  - Correlations of similar mice
  - 6-15 data points
- Disadvantage
  - Arbitrariness of phases

# 5. Correlations of all 2/3 time point mice set

- Advantage
  - Arbitrariness reduced
  - Robustness
- Disadvantage
  - Lots of data
  - Overlap

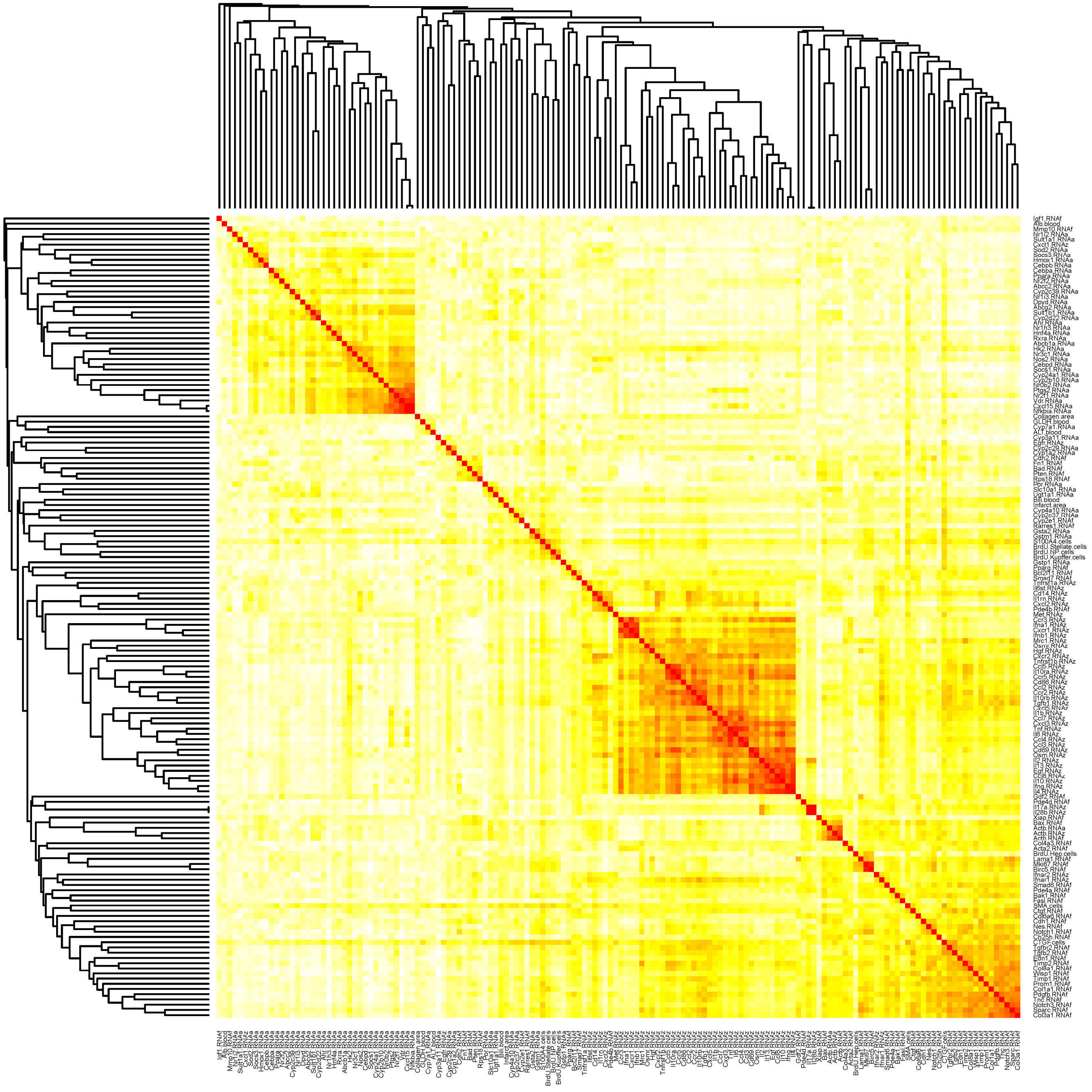
# Consensus correlation

$$C(\alpha,\beta) = \sum_{i=1}^{25} w_i \frac{\mathrm{corr}_i(\alpha,\beta)}{\sqrt[5]{10 + \mathrm{rank}_i(\alpha,\beta)}}$$

- High weight (5) for
  - time average correlations
  - 3 phases correlation
  - all mice correlation
- Low weight (1) for
  - single time point correlations
  - 2/3 mice set correlations
- Rank modification factor
  - reduces the effect of the distribution of correlation values

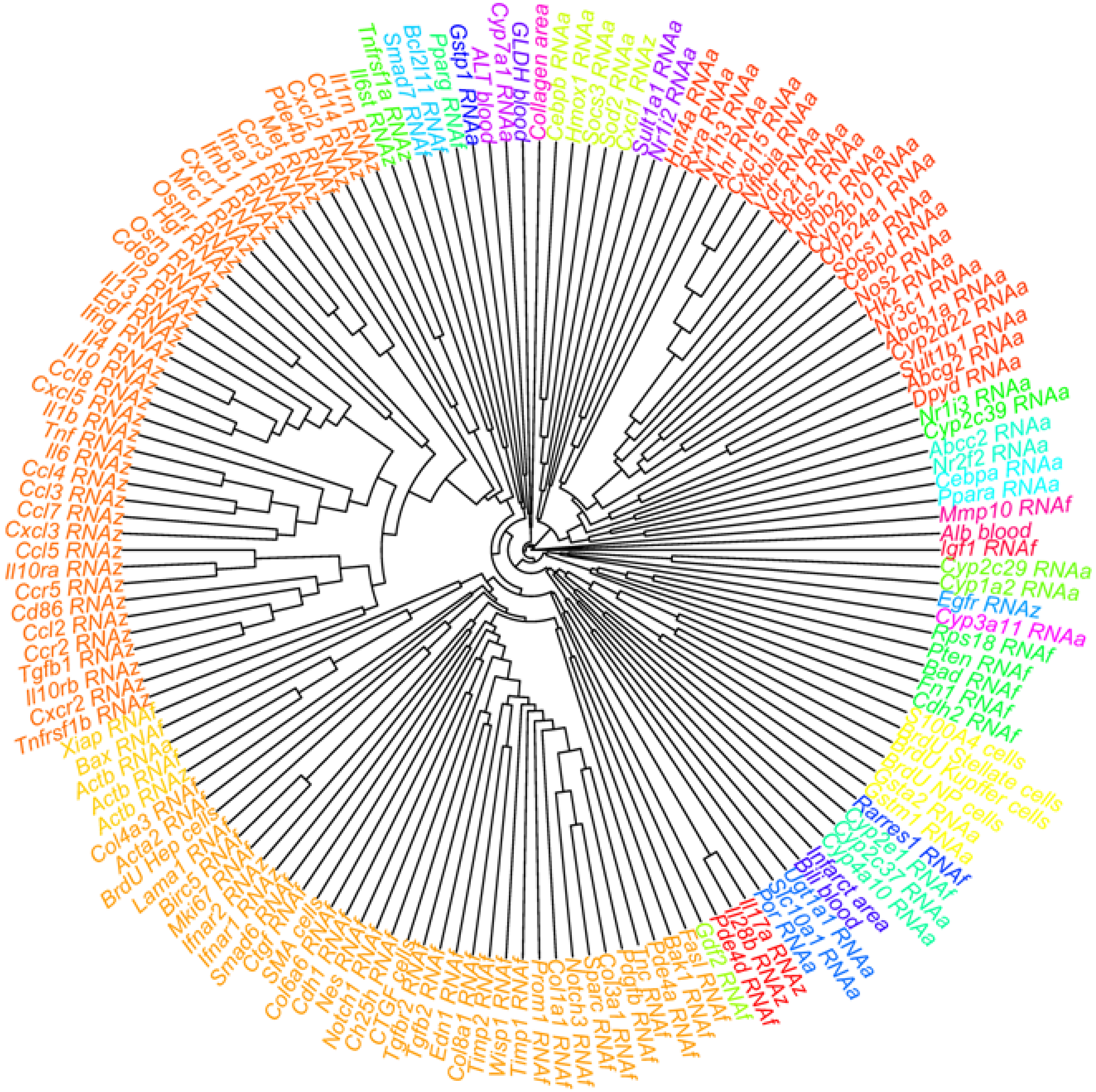
# Consensus correlation

- Advantage
  - Correlation artifacts countered
  - Effect of single mouse outliers reduced
  - Robustness
- Disadvantage
  - Heuristical

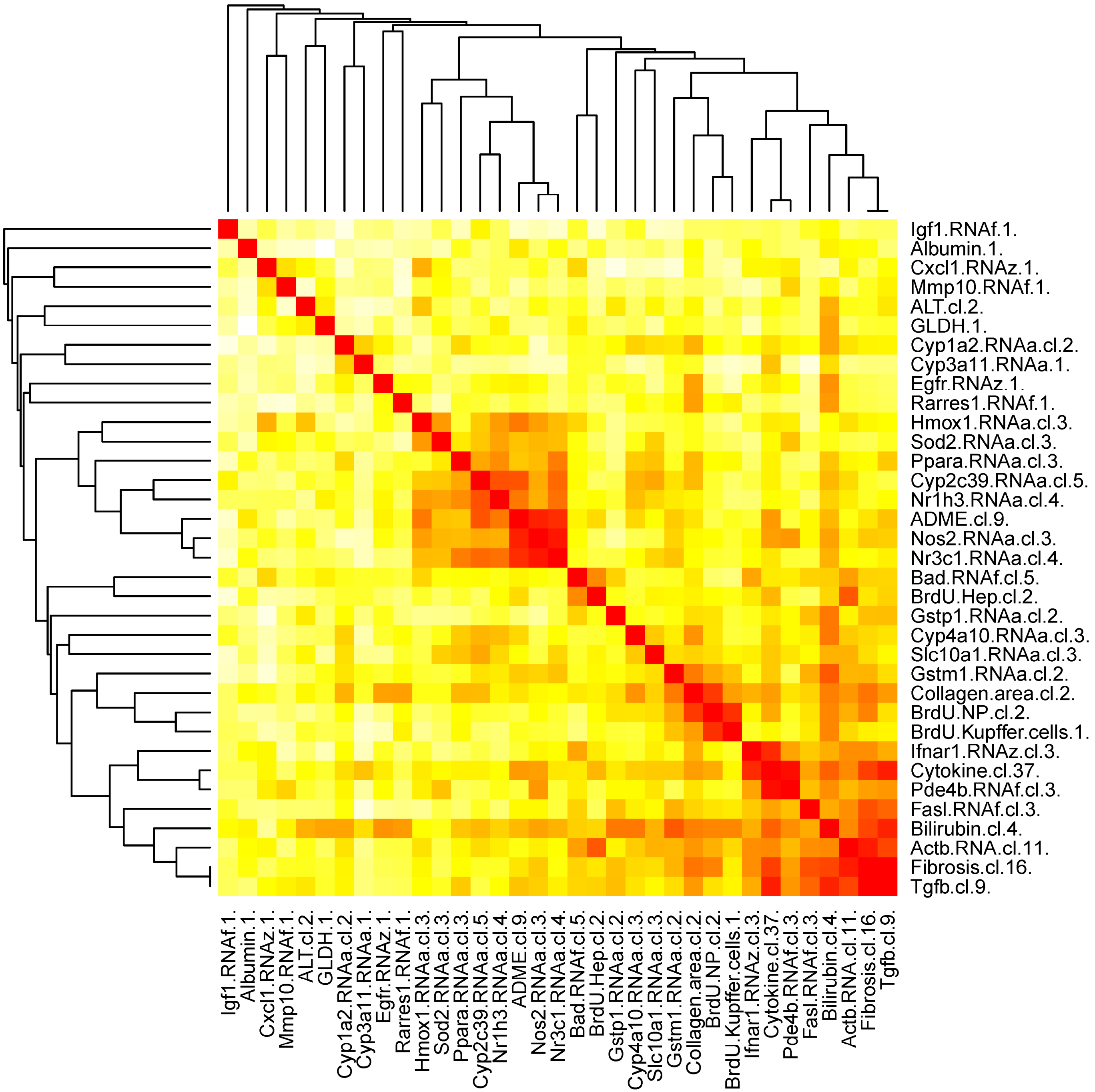


# Consensus correlation heatmeap

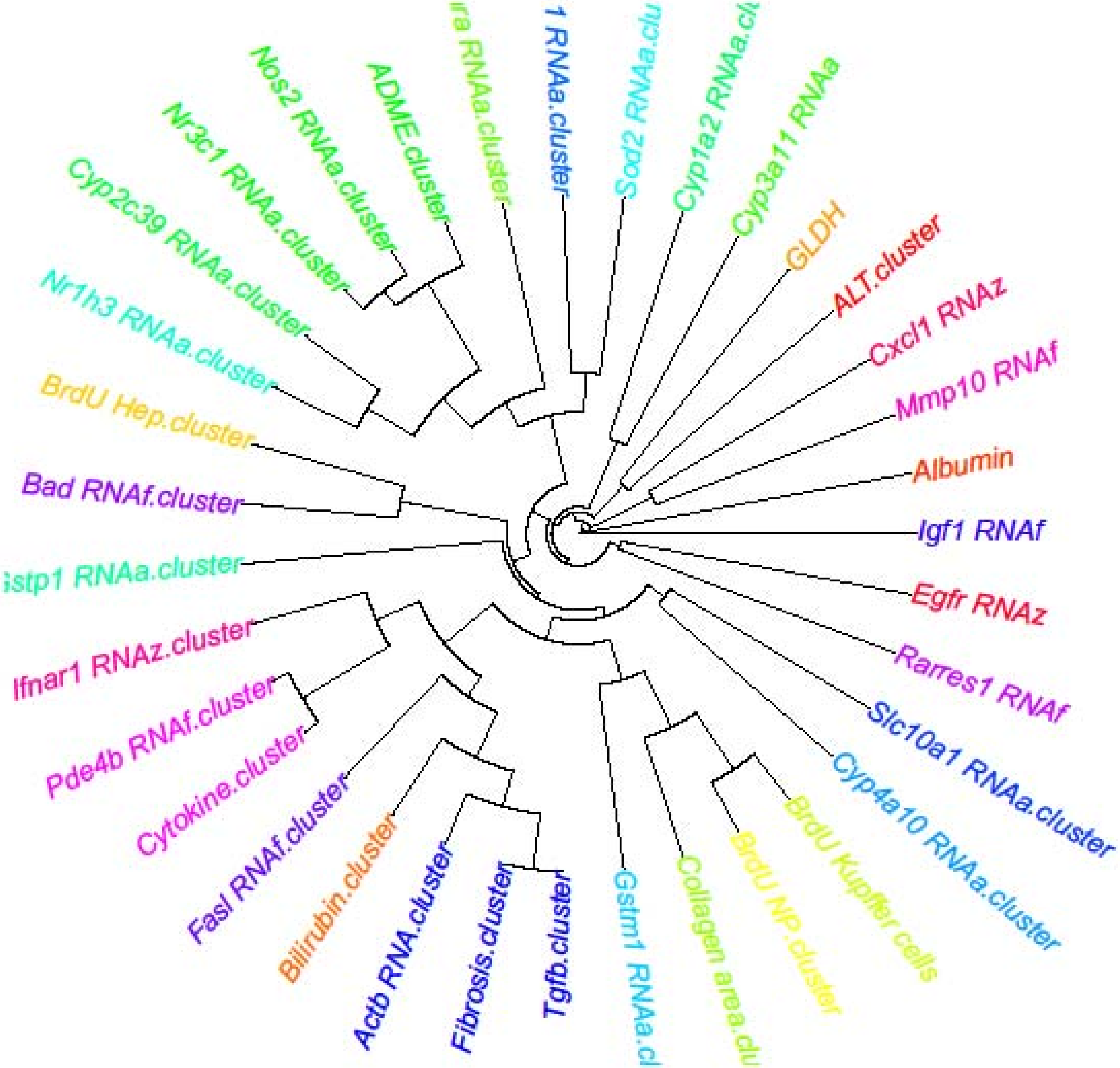
- Blocks: related RNA
- Stripes
  - white low variance
  - yellow: high variance



- ○ Cluster of ALT and Cyp7a1
- ○ Bilirubin & Infarct area cluster
- ○ SMA cluster with fibrosis RNA
- ○ Proliferation marker cluster with RNA such as Gstm1
- ○ GLDH, Collagen area solitude



- Blocks of similar RNA collapsed
- Bili & Collagen dark stripes
- Collagen area & proliferation Kupffer/Stellate close


ra RNAa.cluster
1 RNAa.cluster
Sod2 RNAa.clu
Cyp1a2 RNAa.cl
Cyp3a11 RNAa
GLDH
ALT.cluster
Cxcl1 RNAz
Mmp10 RNAf
Albumin
Igf1 RNAf
Egfr RNAz
Rarres1 RNAf
Slc10a1 RNAa.cluster
Cyp4a10 RNAa.cluster
BrdU Kupffer cells
BrdU NP.cluster
Collagen area.clu
Gstm1 RNAa.cl
Tgfb.cluster
Fibrosis.cluster
Actb RNA.cluster
Bilirubin.cluster
Fasl RNAf.cluster
Cytokine.cluster
Pde4b RNAf.cluster
Ifnar1 RNAz.cluster
stp1 RNAa.cluster
Bad RNAf.cluster
BrdU Hep.cluster
Nr1h3 RNAa.cluster
Cyp2c39 RNAa.cluster
Nr3c1 RNAa.cluster
Nos2 RNAa.cluster
ADME.cluster

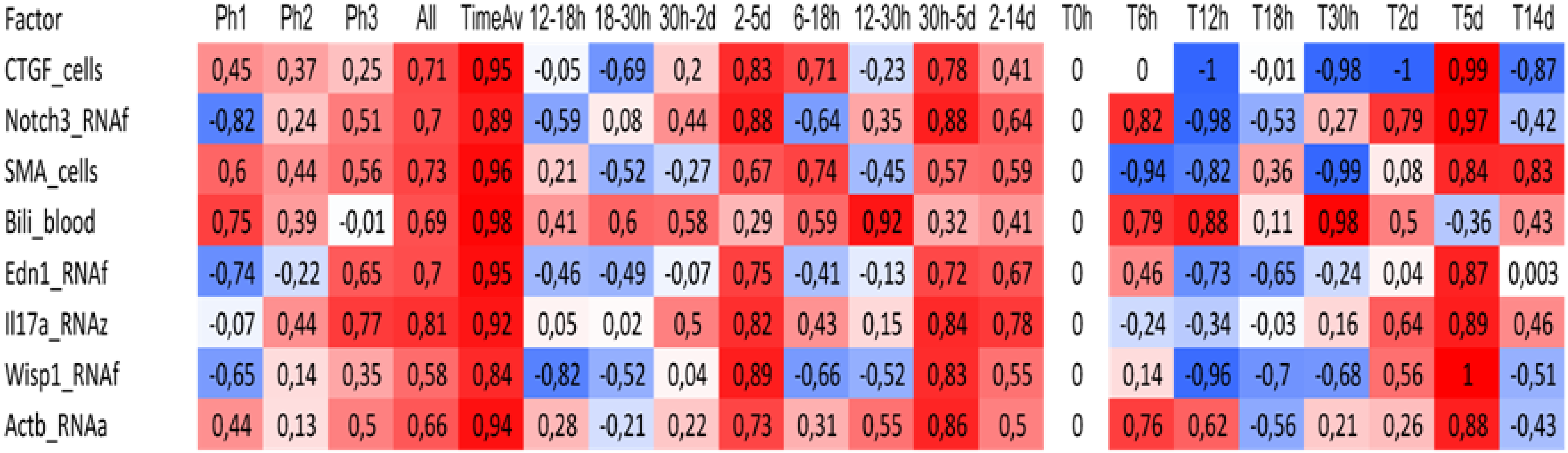
# Outline

- Cholestasis
- Bile duct ligation experiment
- Correlation analysis
- Marker candidates
- Predictors of a time frame
- Decision trees for progress prediction

# Most correlated factors

- One factor marker for another
- Histologic parameters invasive, expensive
- Parameters of interest
  - Infarct area
  - Collagen area
- Markers
  - serum concentrations
  - RNA coding secreted peptide
  - RNA triggering
  - intracellular

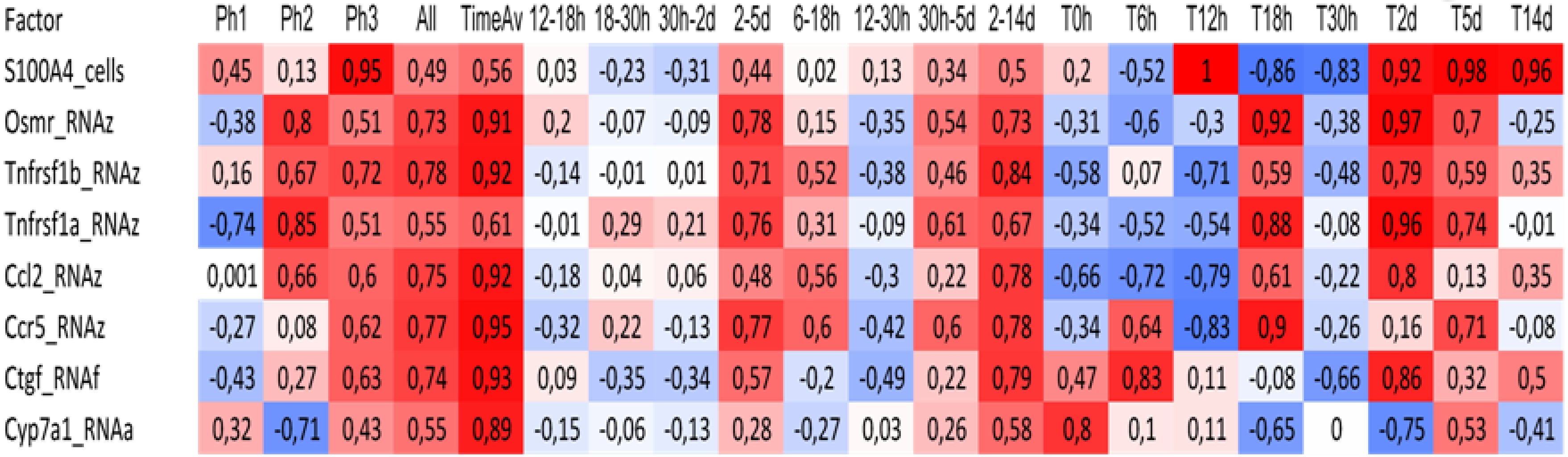
# Infarct area



| Factor | Ph1 | Ph2 | Ph3 | All | TimeAv | 12-18h | 18-30h | 30h-2d | 2-5d | 6-18h | 12-30h | 30h-5d | 2-14d | T0h | T6h | T12h | T18h | T30h | T2d | T5d | T14d |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CTGF_cells | 0,45 | 0,37 | 0,25 | 0,71 | 0,95 | -0,05 | -0,69 | 0,2 | 0,83 | 0,71 | -0,23 | 0,78 | 0,41 | 0 | 0 | -1 | -0,01 | -0,98 | -1 | 0,99 | -0,87 |
| Notch3_RNAf | -0,82 | 0,24 | 0,51 | 0,7 | 0,89 | -0,59 | 0,08 | 0,44 | 0,88 | -0,64 | 0,35 | 0,88 | 0,64 | 0 | 0,82 | -0,98 | -0,53 | 0,27 | 0,79 | 0,97 | -0,42 |
| SMA_cells | 0,6 | 0,44 | 0,56 | 0,73 | 0,96 | 0,21 | -0,52 | -0,27 | 0,67 | 0,74 | -0,45 | 0,57 | 0,59 | 0 | -0,94 | -0,82 | 0,36 | -0,99 | 0,08 | 0,84 | 0,83 |
| Bili_blood | 0,75 | 0,39 | -0,01 | 0,69 | 0,98 | 0,41 | 0,6 | 0,58 | 0,29 | 0,59 | 0,92 | 0,32 | 0,41 | 0 | 0,79 | 0,88 | 0,11 | 0,98 | 0,5 | -0,36 | 0,43 |
| Edn1_RNAf | -0,74 | -0,22 | 0,65 | 0,7 | 0,95 | -0,46 | -0,49 | -0,07 | 0,75 | -0,41 | -0,13 | 0,72 | 0,67 | 0 | 0,46 | -0,73 | -0,65 | -0,24 | 0,04 | 0,87 | 0,003 |
| Il17a_RNAz | -0,07 | 0,44 | 0,77 | 0,81 | 0,92 | 0,05 | 0,02 | 0,5 | 0,82 | 0,43 | 0,15 | 0,84 | 0,78 | 0 | -0,24 | -0,34 | -0,03 | 0,16 | 0,64 | 0,89 | 0,46 |
| Wisp1_RNAf | -0,65 | 0,14 | 0,35 | 0,58 | 0,84 | -0,82 | -0,52 | 0,04 | 0,89 | -0,66 | -0,52 | 0,83 | 0,55 | 0 | 0,14 | -0,96 | -0,7 | -0,68 | 0,56 | 1 | -0,51 |
| Actb_RNAa | 0,44 | 0,13 | 0,5 | 0,66 | 0,94 | 0,28 | -0,21 | 0,22 | 0,73 | 0,31 | 0,55 | 0,86 | 0,5 | 0 | 0,76 | 0,62 | -0,56 | 0,21 | 0,26 | 0,88 | -0,43 |

- Notch3 … Neurogenic locus notch homolog protein 3
  - mutations underlying cause of Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy
  - associated with Alzheimer's disease
- Edn1 … preproendothelin-1
  - proteolytically processed to **release** a secreted peptide endothelin 1
  - vasoconstrictor
  - antagonists are used in the treat of pulmonary hypertension

# Collagen area



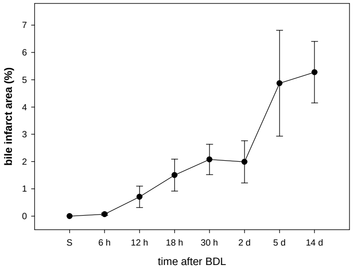
| Factor | Ph1 | Ph2 | Ph3 | All | TimeAv | 12-18h | 18-30h | 30h-2d | 2-5d | 6-18h | 12-30h | 30h-5d | 2-14d | T0h | T6h | T12h | T18h | T30h | T2d | T5d | T14d |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S100A4_cells | 0,45 | 0,13 | 0,95 | 0,49 | 0,56 | 0,03 | -0,23 | -0,31 | 0,44 | 0,02 | 0,13 | 0,34 | 0,5 | 0,2 | -0,52 | 1 | -0,86 | -0,83 | 0,92 | 0,98 | 0,96 |
| Osmr_RNAz | -0,38 | 0,8 | 0,51 | 0,73 | 0,91 | 0,2 | -0,07 | -0,09 | 0,78 | 0,15 | -0,35 | 0,54 | 0,73 | -0,31 | -0,6 | -0,3 | 0,92 | -0,38 | 0,97 | 0,7 | -0,25 |
| Tnfrsf1b_RNAz | 0,16 | 0,67 | 0,72 | 0,78 | 0,92 | -0,14 | -0,01 | 0,01 | 0,71 | 0,52 | -0,38 | 0,46 | 0,84 | -0,58 | 0,07 | -0,71 | 0,59 | -0,48 | 0,79 | 0,59 | 0,35 |
| Tnfrsf1a_RNAz | -0,74 | 0,85 | 0,51 | 0,55 | 0,61 | -0,01 | 0,29 | 0,21 | 0,76 | 0,31 | -0,09 | 0,61 | 0,67 | -0,34 | -0,52 | -0,54 | 0,88 | -0,08 | 0,96 | 0,74 | -0,01 |
| Ccl2_RNAz | 0,001 | 0,66 | 0,6 | 0,75 | 0,92 | -0,18 | 0,04 | 0,06 | 0,48 | 0,56 | -0,3 | 0,22 | 0,78 | -0,66 | -0,72 | -0,79 | 0,61 | -0,22 | 0,8 | 0,13 | 0,35 |
| Ccr5_RNAz | -0,27 | 0,08 | 0,62 | 0,77 | 0,95 | -0,32 | 0,22 | -0,13 | 0,77 | 0,6 | -0,42 | 0,6 | 0,78 | -0,34 | 0,64 | -0,83 | 0,9 | -0,26 | 0,16 | 0,71 | -0,08 |
| Ctgf_RNAf | -0,43 | 0,27 | 0,63 | 0,74 | 0,93 | 0,09 | -0,35 | -0,34 | 0,57 | -0,2 | -0,49 | 0,22 | 0,79 | 0,47 | 0,83 | 0,11 | -0,08 | -0,66 | 0,86 | 0,32 | 0,5 |
| Cyp7a1_RNAa | 0,32 | -0,71 | 0,43 | 0,55 | 0,89 | -0,15 | -0,06 | -0,13 | 0,28 | -0,27 | 0,03 | 0,26 | 0,58 | 0,8 | 0,1 | 0,11 | -0,65 | 0 | -0,75 | 0,53 | -0,41 |

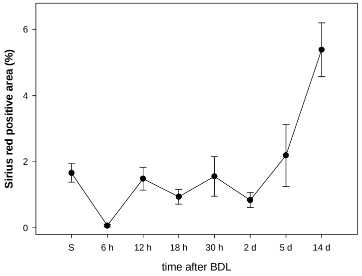
- Osmr … Oncostatin-M specific receptor subunit β
  - associated with primary cutaneous amyloidosis
  - transduces oncostatin M and interleukin 31 induced signaling events
- Tnfrsf1a/b … tumor necrosis factor receptor, CD120
  - mediates TNFα-induced sickness behavior
  - involved in neurotoxic processes
  - elevated in severe mental disorders

# Infarct/Collagen
# Alternate fate hypothesis

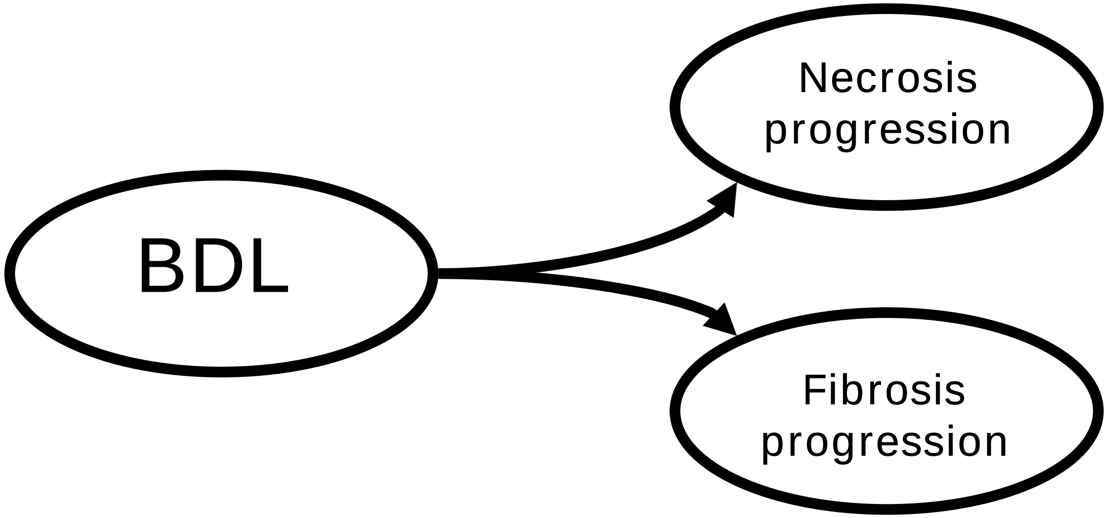
| Factor1 | Factor2 | Ph1 | Ph2 | Ph3 | All | TimeAv | 12-18h | 18-30h | 30h-2d | 2-5d | 6-18h | 12-30h | 30h-5d | 2-14d | T0h | T6h | T12h | T18h | T30h | T2d | T5d | T14d |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Infarct_area | Collagen_area | 0,06 | -0,04 | -0,46 | 0,22 | 0,76 | -0,52 | -0,12 | 0,01 | -0,25 | 0,07 | -0,19 | -0,39 | -0,03 | 0 | 0,75 | -0,5 | -0,43 | -0,15 | 0,3 | -0,64 | -0,81 |

- Positive correlation of time averages





- Strong negative correlations in later time points



# Outline

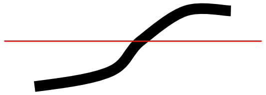
- Cholestasis
- Bile duct ligation experiment
- Correlation analysis
- Marker candidates
- Predictors of a time frame
- Decision trees for progress prediction

# Outlook for the clinical application

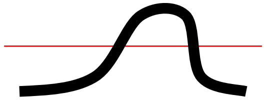
- Marker for
  - the severness of cholestasis
  - duration of cholestasis
- Limits of this experiment
  - complete cholestasis
  - no individual mouse development

# Predictors

- Factor+Time frame+Threshold
- Time frame can
  - start at 0h/end at 14d (bipartition)



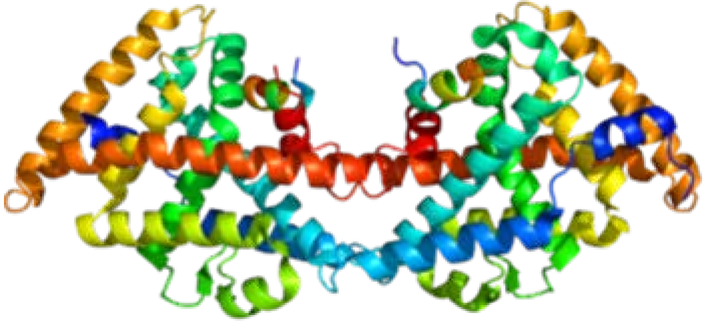
  - in the middle (tripartition)

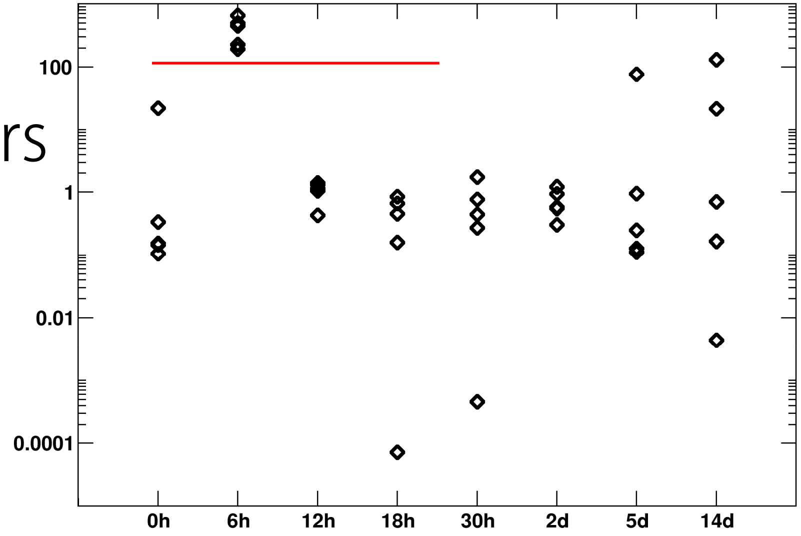


- 41 predictors with separation
  - perfect: gap between *in* and *out*, no outliers
  - 35 bipartition
  - 6 tripartition

# Nr0b2–small heterodimer partner (SHP)

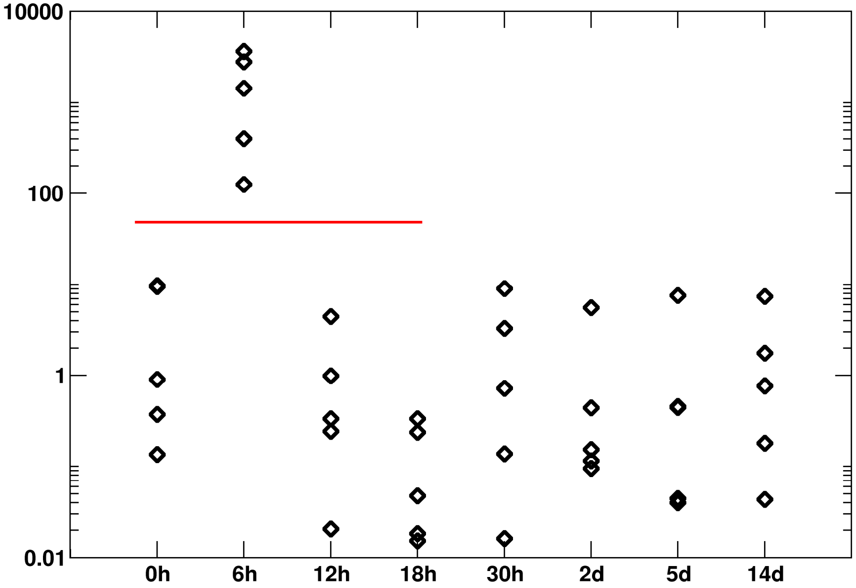
- strongly increases only for mice in the 6h timepoint
- homologous to nuclear receptor family but lacks DNA binding domain
- repression of other nuclear receptors
  - retinoid and thyroid hormone receptors
  - estrogen receptors
- competition with co-activators





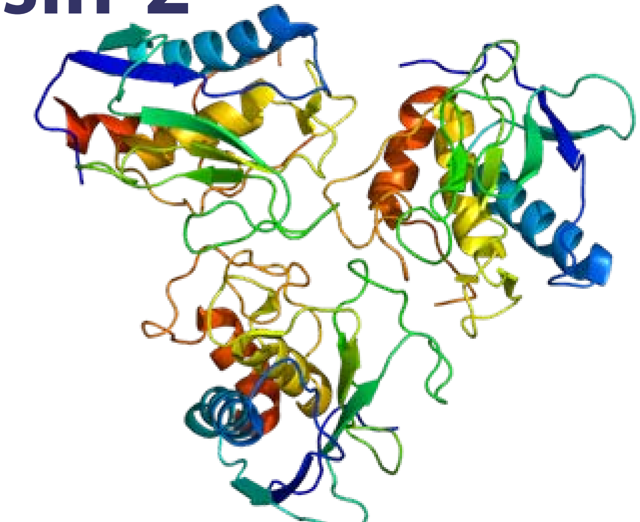
# Cyp24a1–mitochondrial 1,25-dihydroxyvitamin D3 24-hydroxylase

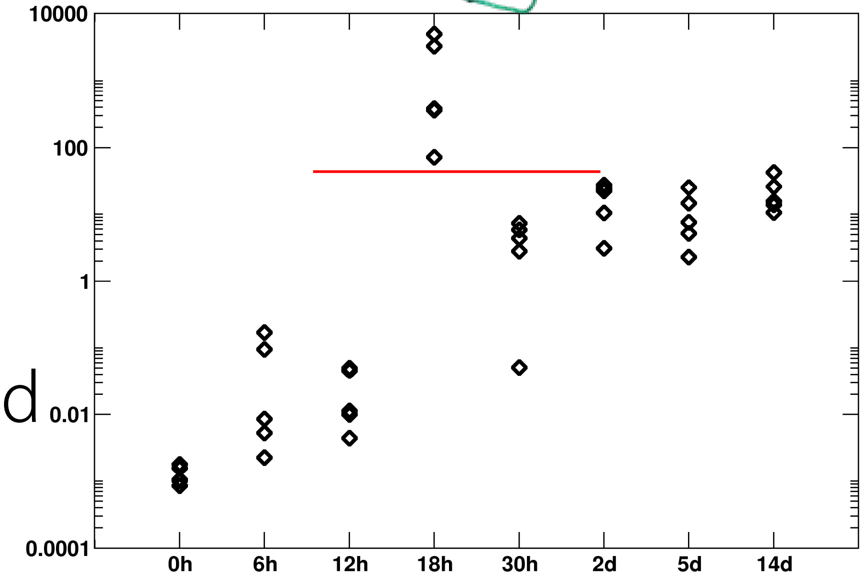
- strongly increases only for mice in the 6h timepoint
- initiates the degradation of 1,25-dihydroxyvitamin D3
- regulating the level of vitamin D3
- calcium homeostasis
- vitamin D endocrine system.
- involoved in Severe Infantile Hypercalcemia



# Mmp10–Stromelysin-2/matrix metalloproteinase-10/transin-2

- strongly increases at 18h
- breakdown of extracellular matrix
  - embryonic development, reproduction, and tissue remodeling
  - arthritis, metastasis.
- secreted as inactive proproteins, activated by extracellular proteinases
- degrades proteoglycans and fibronectin

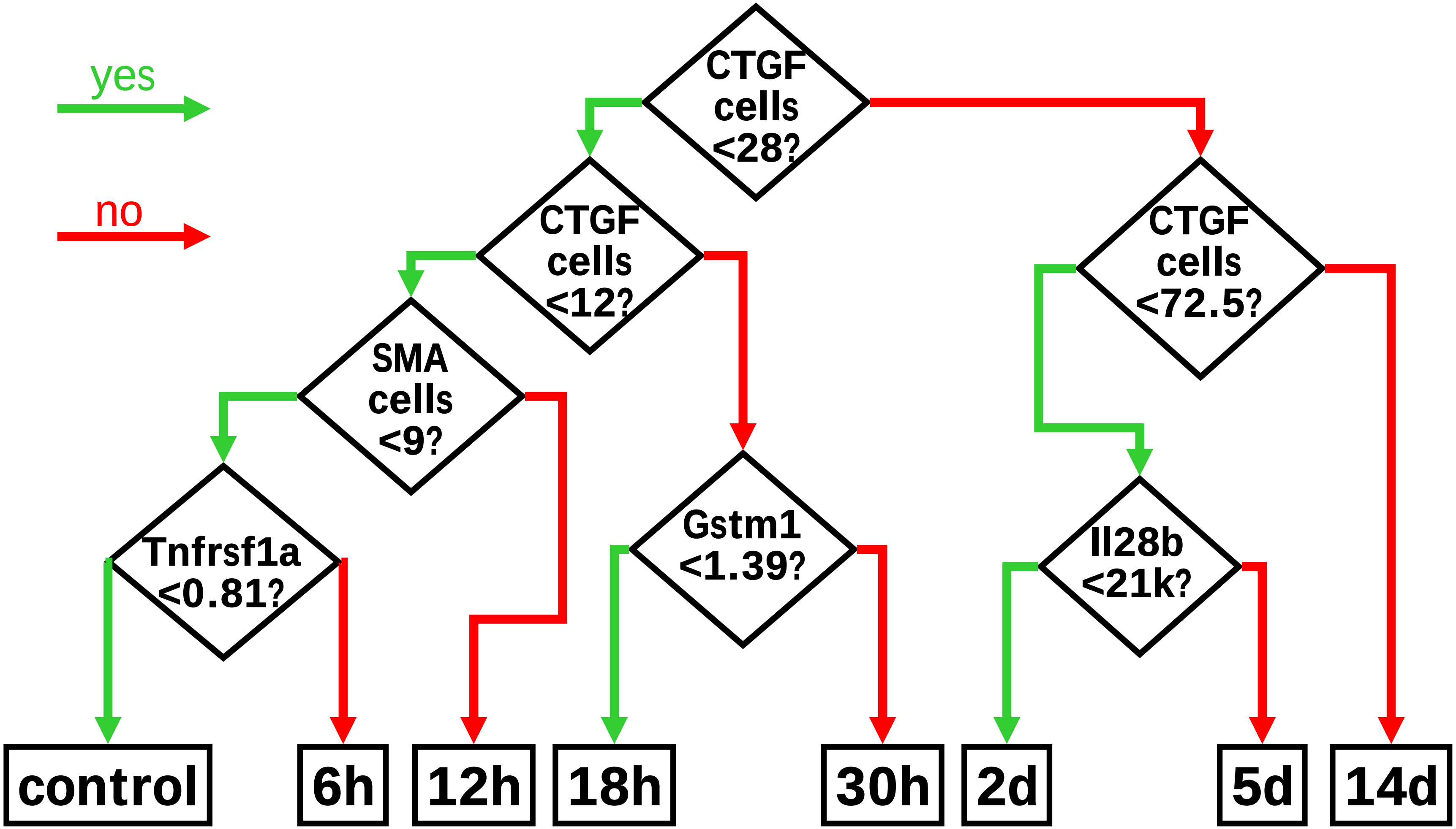




# Outline

- Cholestasis
- Bile duct ligation experiment
- Correlation analysis
- Marker candidates
- Predictors of a time frame
- Decision trees for progress prediction

# Decision tree for time points



yes

no

CTGF cells <28?

CTGF cells <12?

CTGF cells <72.5?

SMA cells <9?

Tnfrsf1a <0.81?

Gstm1 <1.39?

Il28b <21k?

control

6h

12h

18h

30h

2d

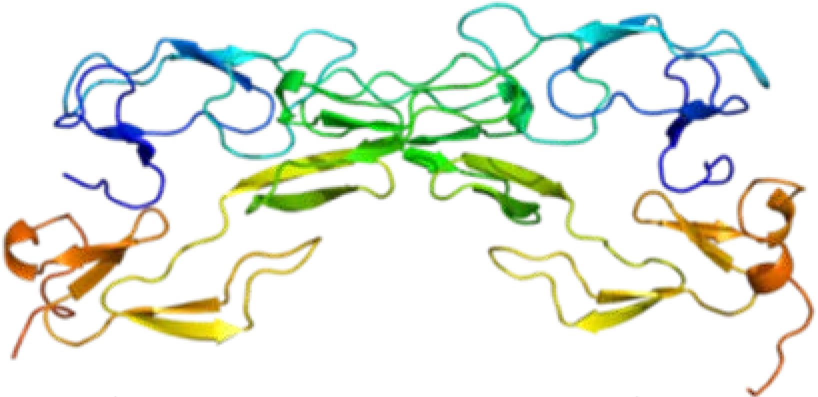
5d

14d

# Quality test

- Decision tree algorithm
  - order factors by largest separation for each split
  - bipartite splits with best factor
- Leave-one-out test
  - 15 of 24 correct

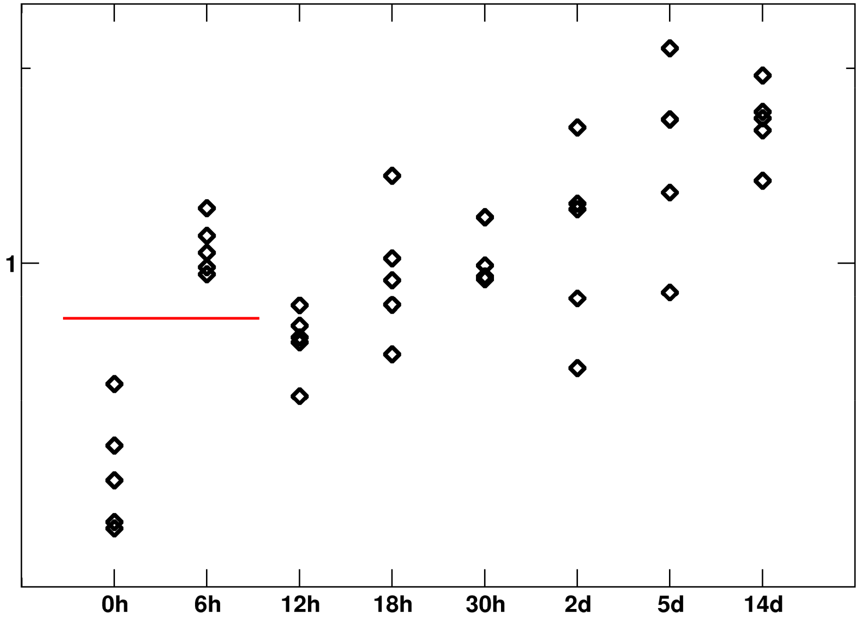
# Tnfrsf1a



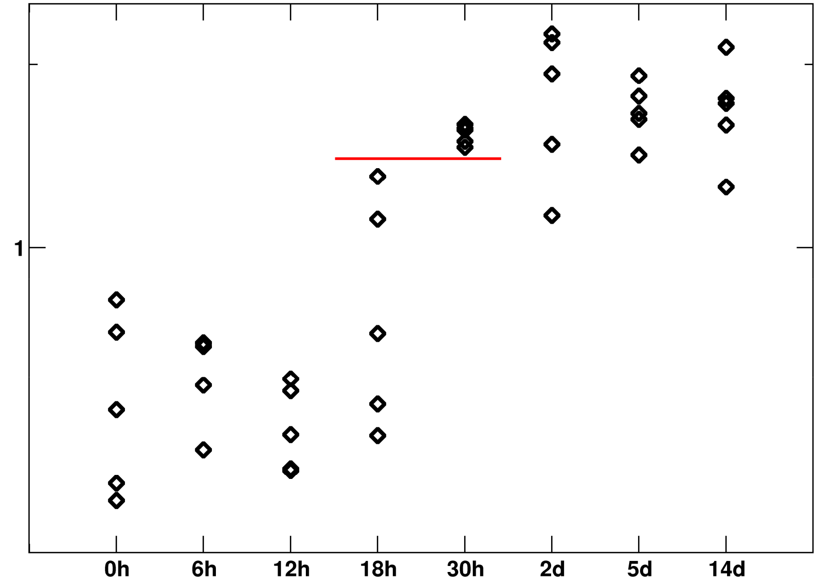
Tnfrsf1a ··· Tumor necrosis factor receptor superfamily member 1A

- one of the major receptors for the tumor necrosis factor-alpha
- can activate the transcription factor NF-κ B

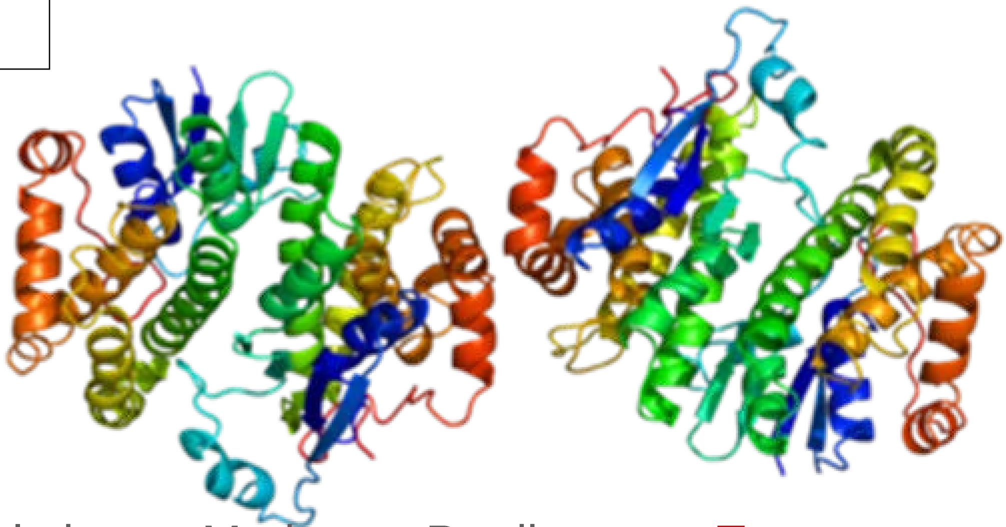
- mediate apoptosis
- regulator of inflammation
- **serum levels** of TNFRSF1A are elevated in schizophrenia and bipolar disorder
- high levels associated with more severe psychotic symptoms



# Gstm1–Glutathione S-transferase mu 1

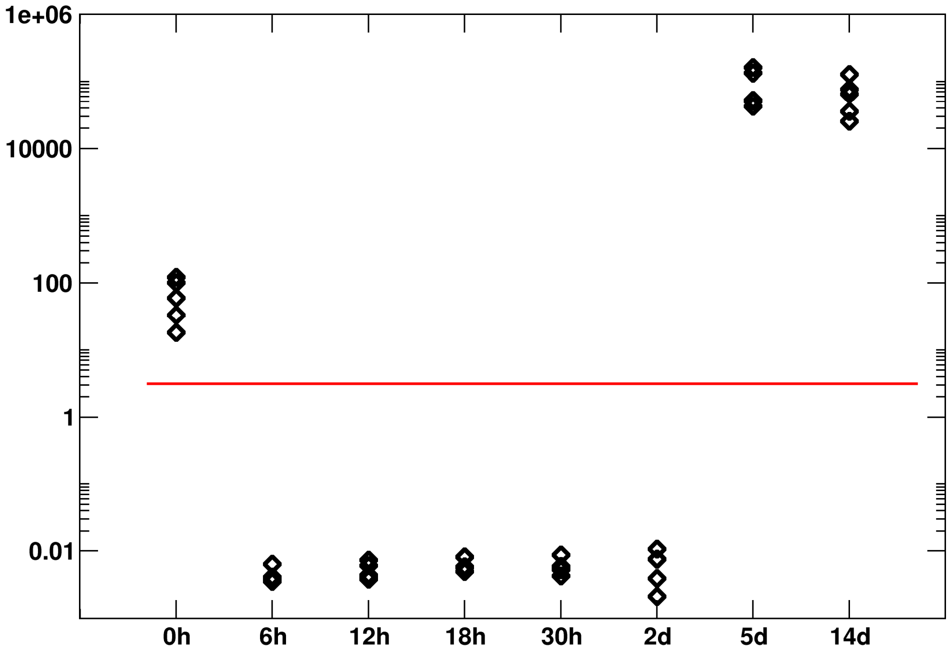


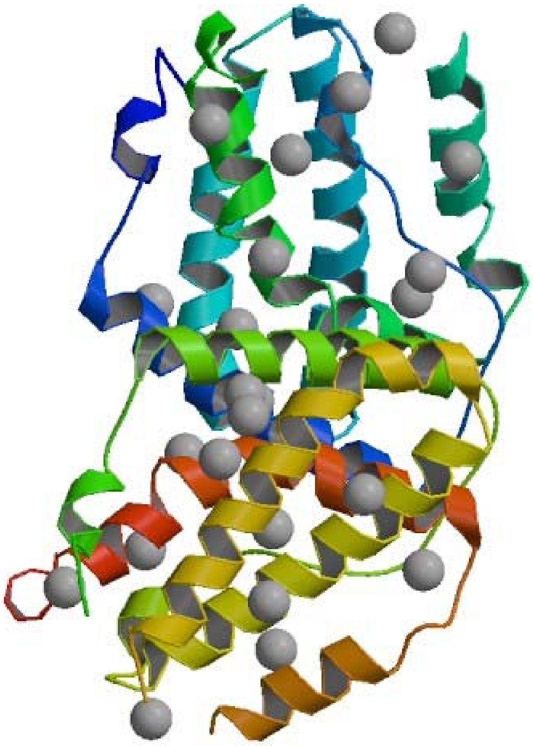
- detoxification of electrophilic compounds



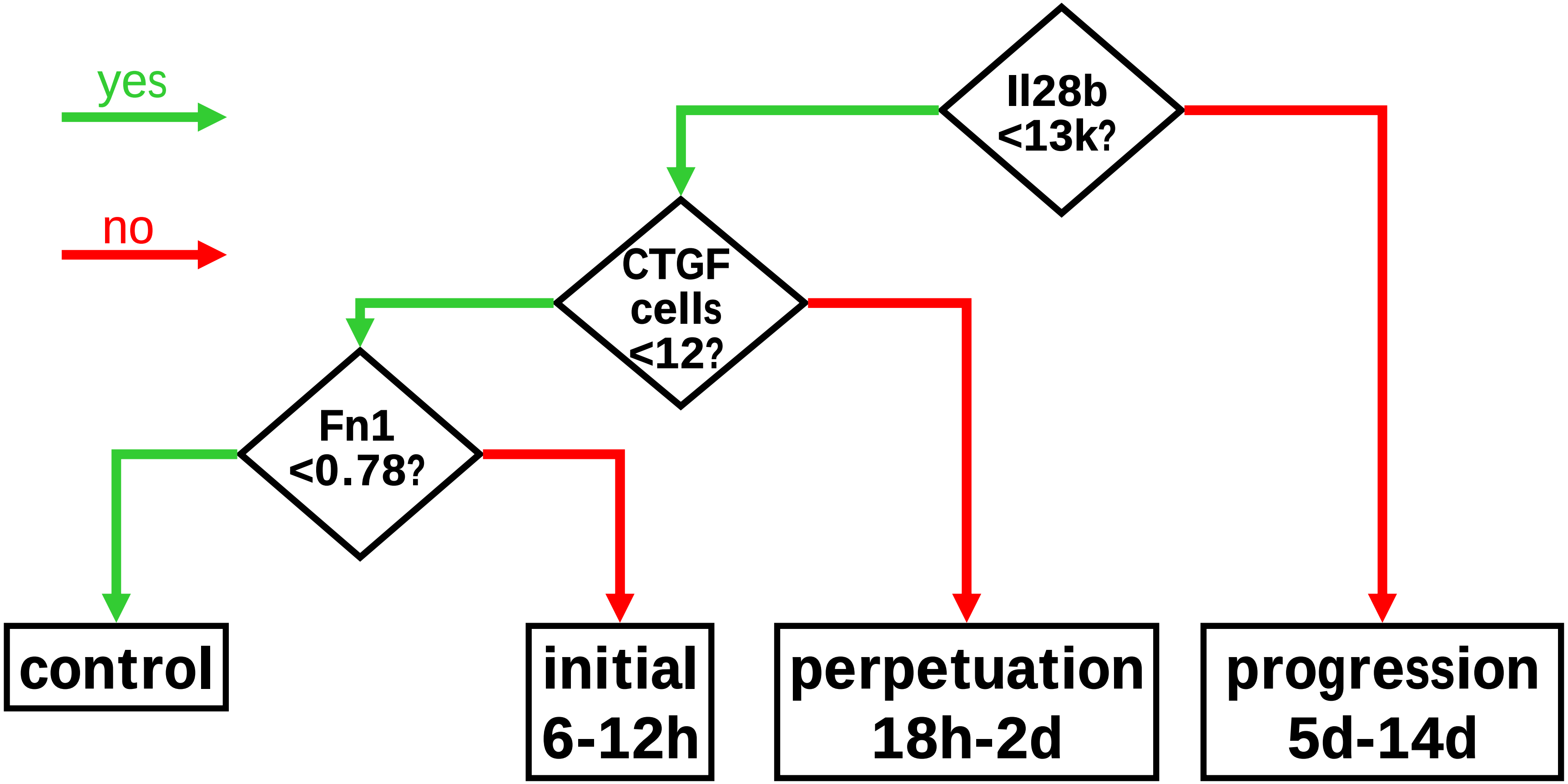
# Il28b–Interleukin 28B (interferon, lambda 3)

- cytokine
- induced by viral infection





# Decision tree for phase



yes

no

Il28b <13k?

CTGF cells <12?

Fn1 <0.78?

control

initial 6-12h

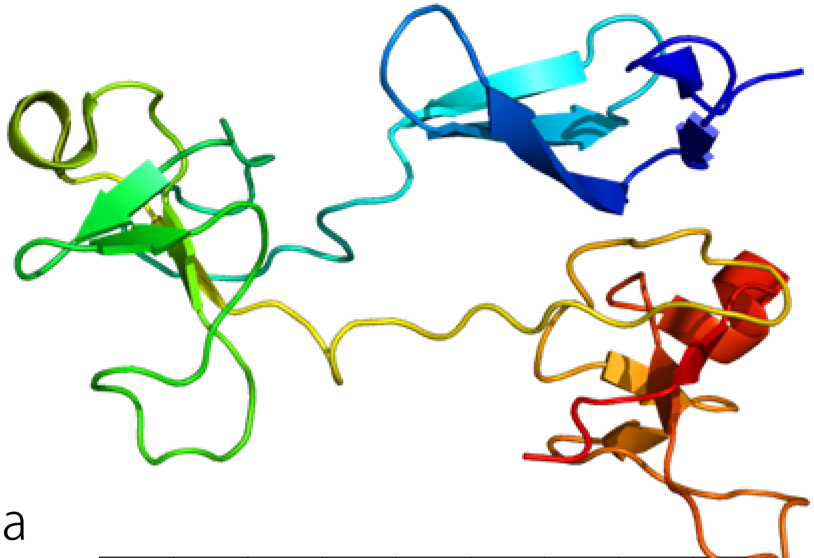
perpetuation 18h-2d

progression 5d-14d

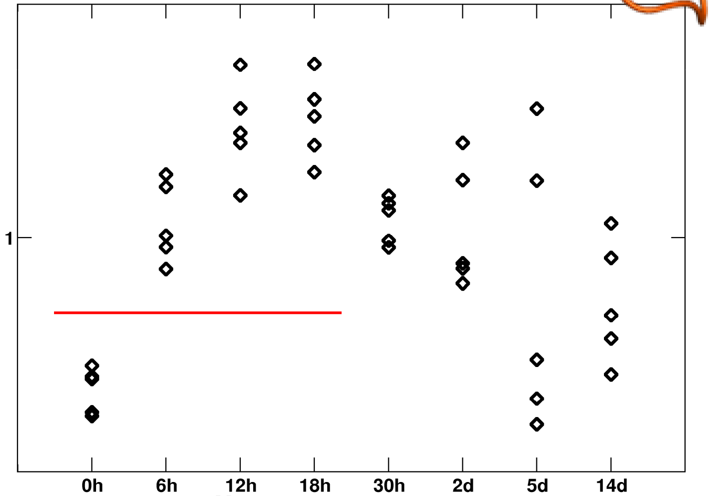
# Quality test

- Decision tree algorithm
  - order factors by largest separation for each split
  - bipartite splits with best factor
- Leave one-out test
  - 28/29 correct

# Fn1–fibronectin



- glycoprotein of the extracellular matrix that binds to membrane-spanning receptor proteins called integrins
- two forms
  - soluble plasma fibronectin, major protein component of blood plasma (300 μg/ml), produced by hepatocytes.
  - insoluble cellular fibronectin is a major component of the extracellular matrix, primarily secreted by fibroblasts



# Take-home message

- Cholestasis complex process
  - necrosis, inflammation, proliferation, fibrosis
- Correlation analysis
  - identifies groups of related factors
  - most correlated RNA for infarct/collagen area
- Predictors of stage
  - single factors time frame
  - decision trees
- 10 genes suggested as markers
  - partly gene products secreted

# Acknowledgements

- Kerstin Abshagen, Isabell Müller, Christian Eipel, Brigitte Vollmer
  - Institute for experimental surgery, Rostock university medical center
- Steven Dooley, Honglei Weng
  - Molecular Hepatology, Medical Faculty Mannheim, Heidelberg University
- Maria Thomas
  - Margarete Firscher-Bosch Institute for Clinical Pharmacology Stuttgart