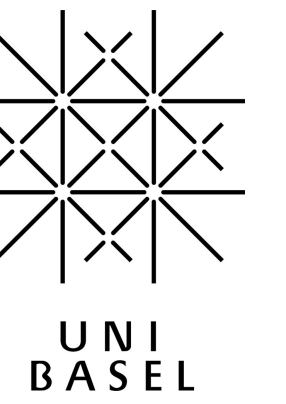


CC-SPMD  
Competence Center for  
Systems Physiology  
and Metabolic Diseases

# Data-driven Modelling of Hepatic Insulin Resistance

Julia E. Vogt<sup>1</sup>, Alberto Giovanni Busetto<sup>2</sup> and Daniel J. Stekhoven<sup>3</sup>  
<sup>1</sup>Computer Science, University of Basel; <sup>2</sup>Machine Learning Lab and <sup>3</sup>Seminar for Statistics, ETH Zurich

ETH  
Eidgenössische Technische Hochschule Zürich  
Swiss Federal Institute of Technology Zurich



## Overview

### Main Goal:

Data-driven generation and testing of hypothetical models aimed at understanding the mechanisms of hepatic insulin resistance.

### Approach:

- Filtering:** screening and selection of genes & proteins.
- Clustering:** finding structures of genes & proteins.
- Network Identification:** inferring interaction between structures.
- Behavior Identification:** selecting hypothetical mechanisms.
- Enrichment Analysis:** mapping models to biological functions.

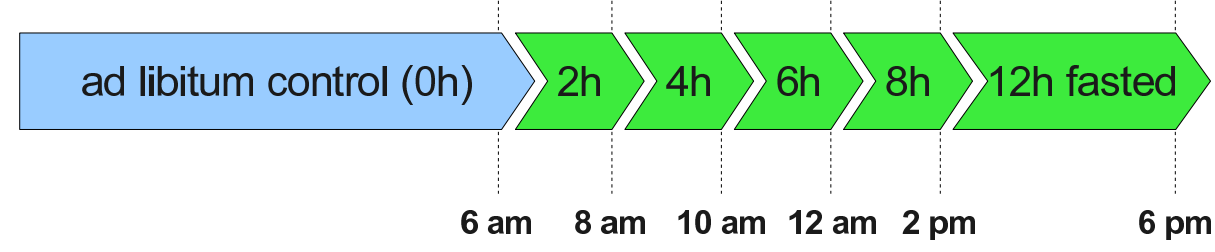
## Data & Framework

### Available Data:

LiverX short-term experiment on *M. musculus* (male C57 BL/6J strain) providing exon and protein measurements.

### Experimental Layouts:

#### Feeding-to-fasting transition



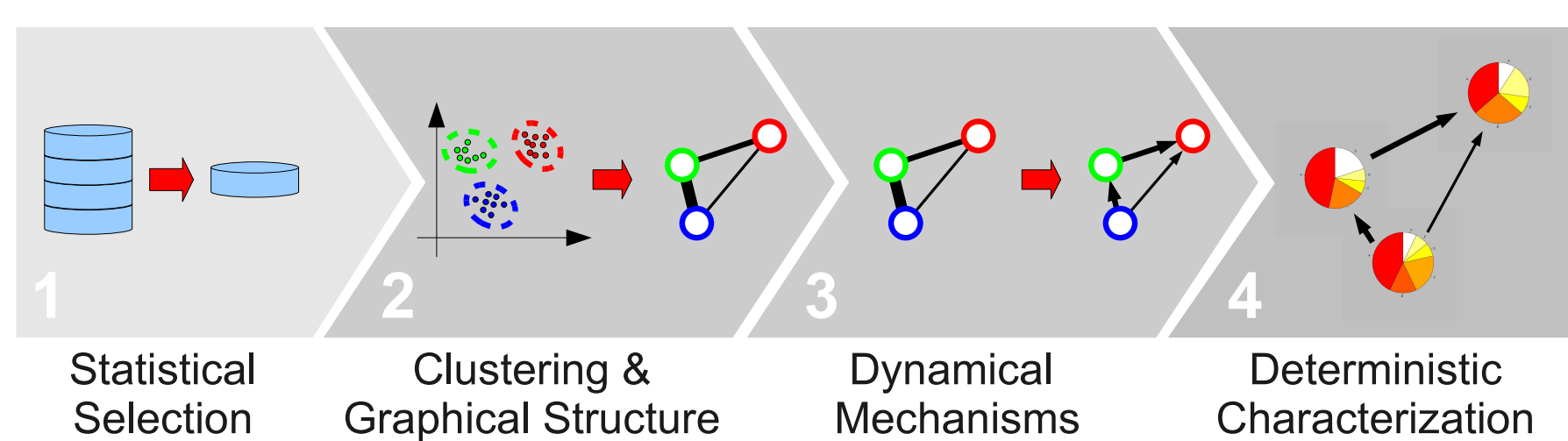
#### Fasting-to-feeding transition



23'081 genes obtained after gene-wise median aggregation of the exon expressions.  
855 proteins available from proteomics.

### Modeling Pipeline:

The modelling approach deploys as follows:



## Gene Screening and Selection

**Basic Assumption:** sparse structure of direct gene interactions in hepatic insulin resistance.

### Gene Selection Conditions:

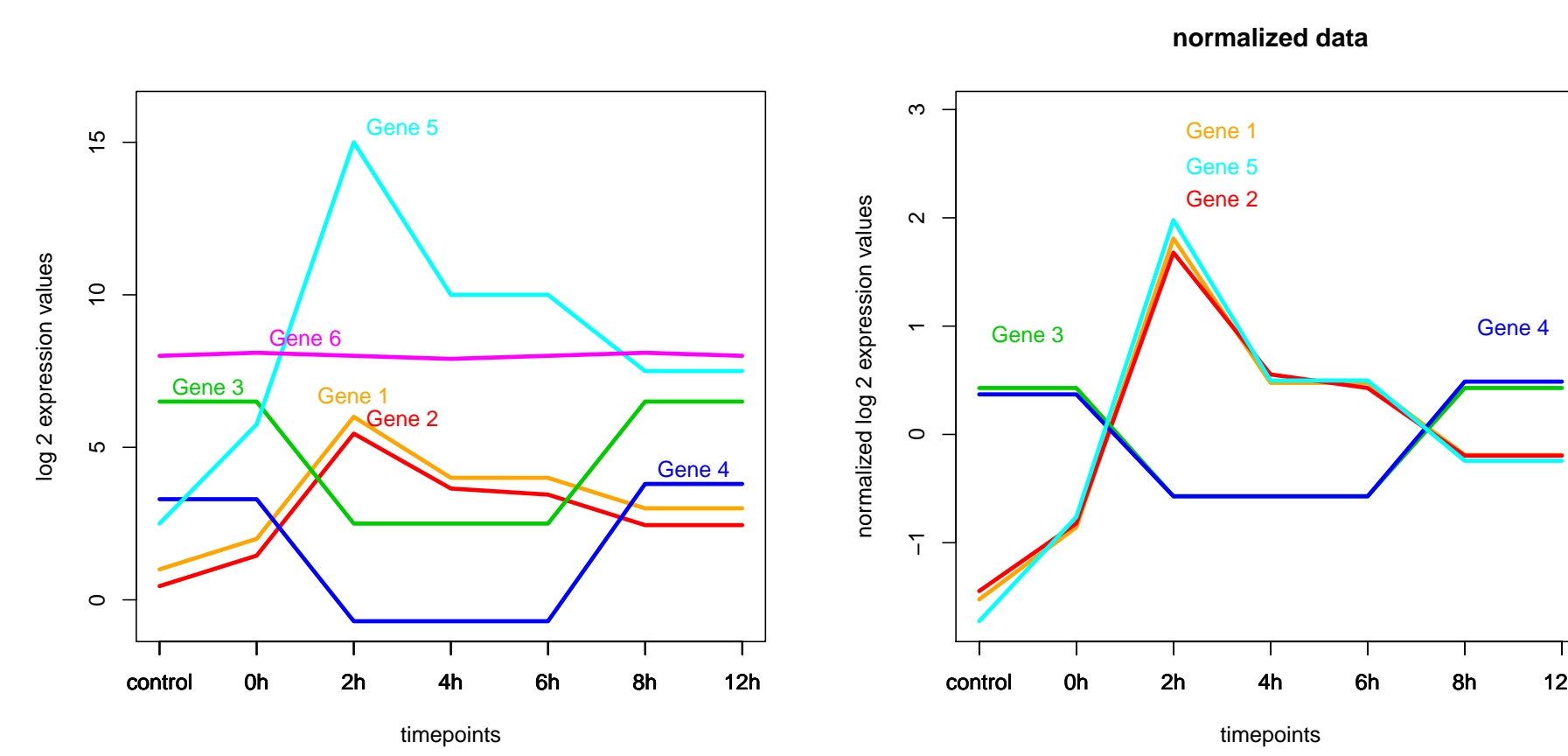
- presence on gene priority list (705 genes);
- high variance during the time-course (2'000 genes exhibiting highest variance);
- significant differential expression (4'000 genes).

**Differential Expressions:** analysis obtained through hierarchical models with empirical Bayes approach (G.K. Smyth, 2004). The composite null-hypotheses were tested using a moderated F-statistic. All gene selections are significant after multiple testing correction controlling false discovery rate (Benjamini & Hochberg, 1995).

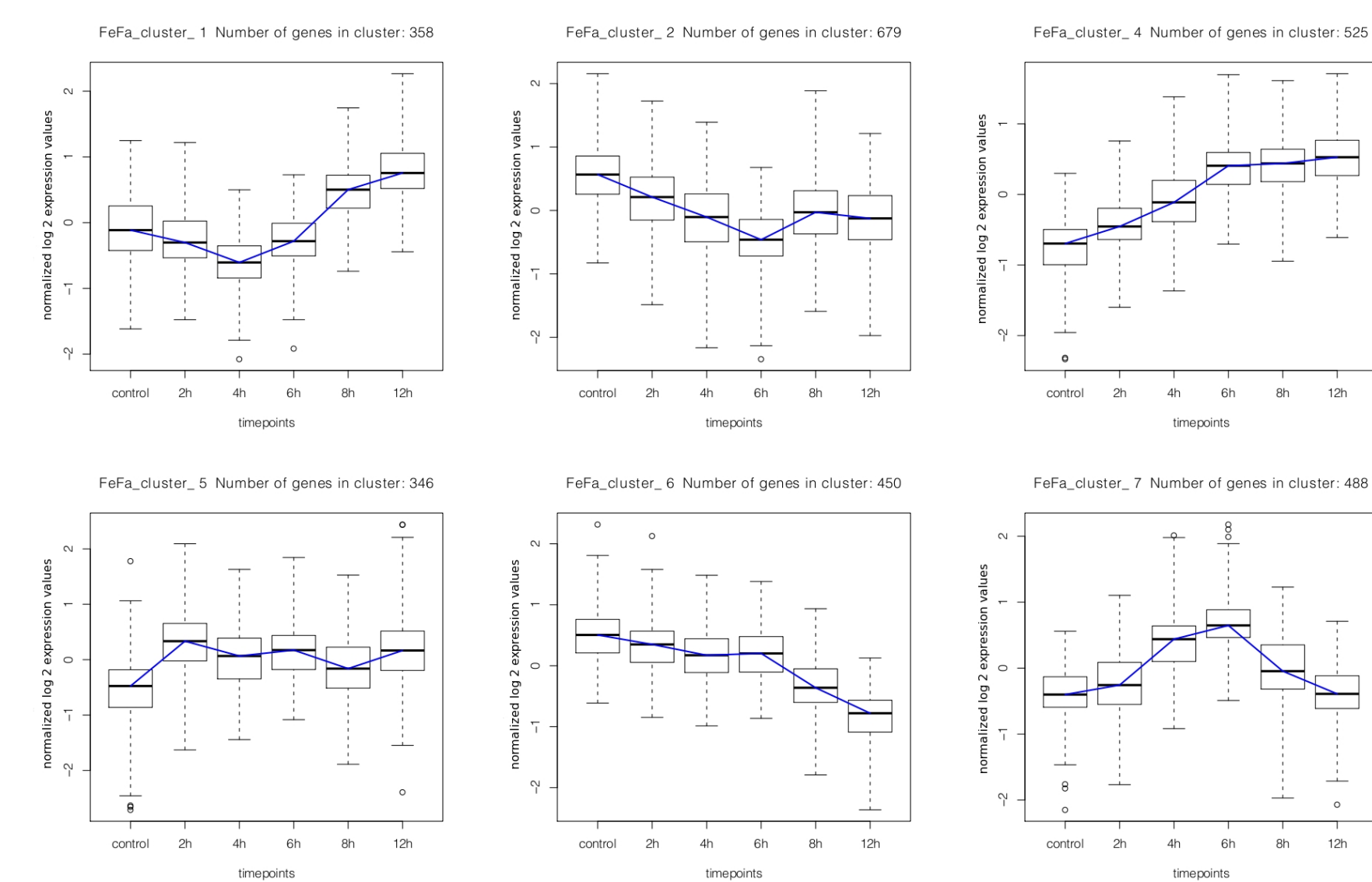
## Functional Network Reconstruction

### Clustering:

genes and proteins grouped (separately) using an infinite Gaussian mixture model (J. Pitman, 2006).



### Examples of clusters (feeding-to-fasting transition)

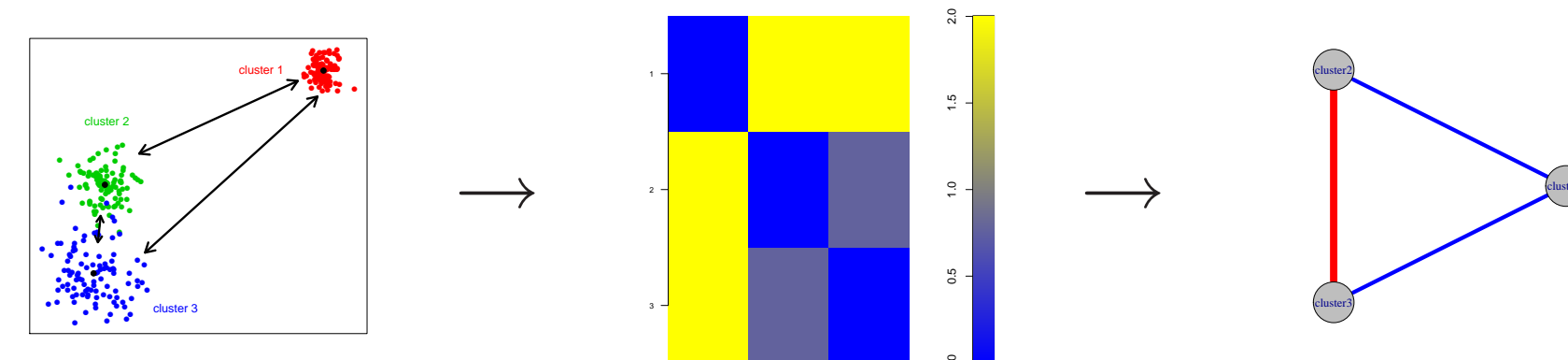


### Graphical Structure:

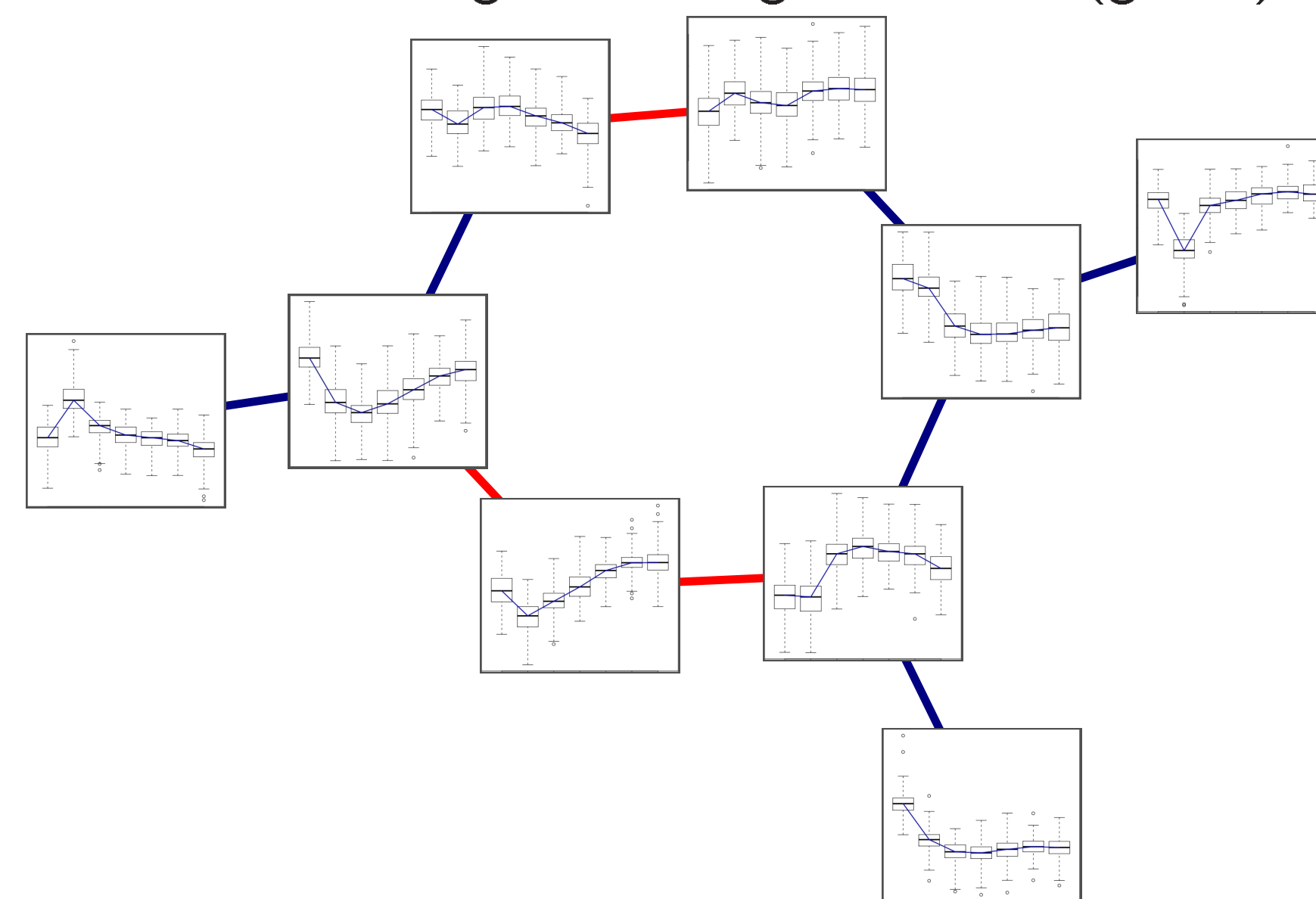
reconstruction of the underlying cluster interaction network.

- Nodes:** clusters
- Edges:** partial correlation

- red edge = positive partial correlation
- blue edge = negative partial correlation



### Network fasting-to-feeding transition (genes)

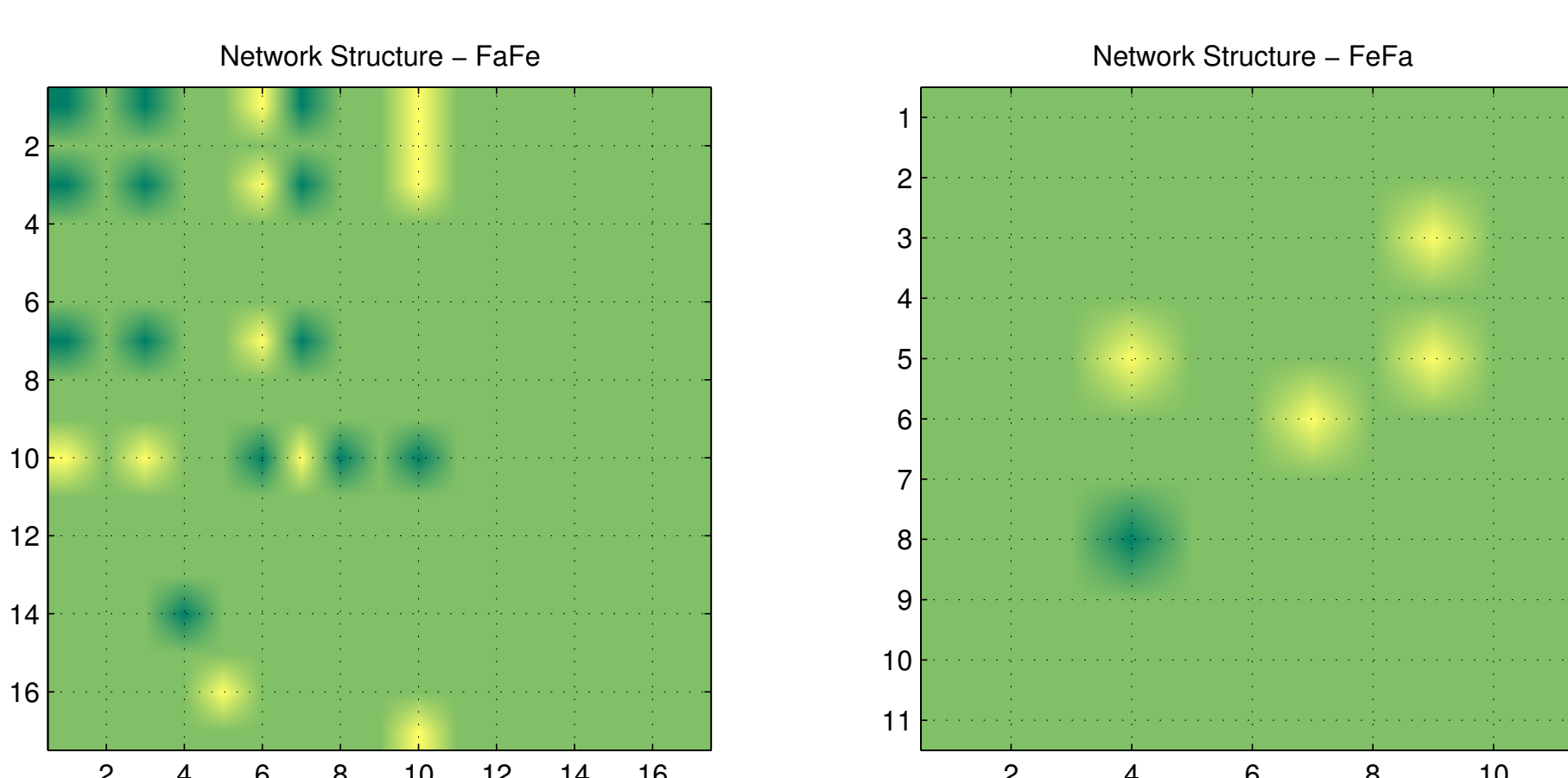


## Activation/Inhibition Detection

### Time Series Analysis:

- detection of **activation** processes (light green)
- detection of **inhibition** processes (dark green)

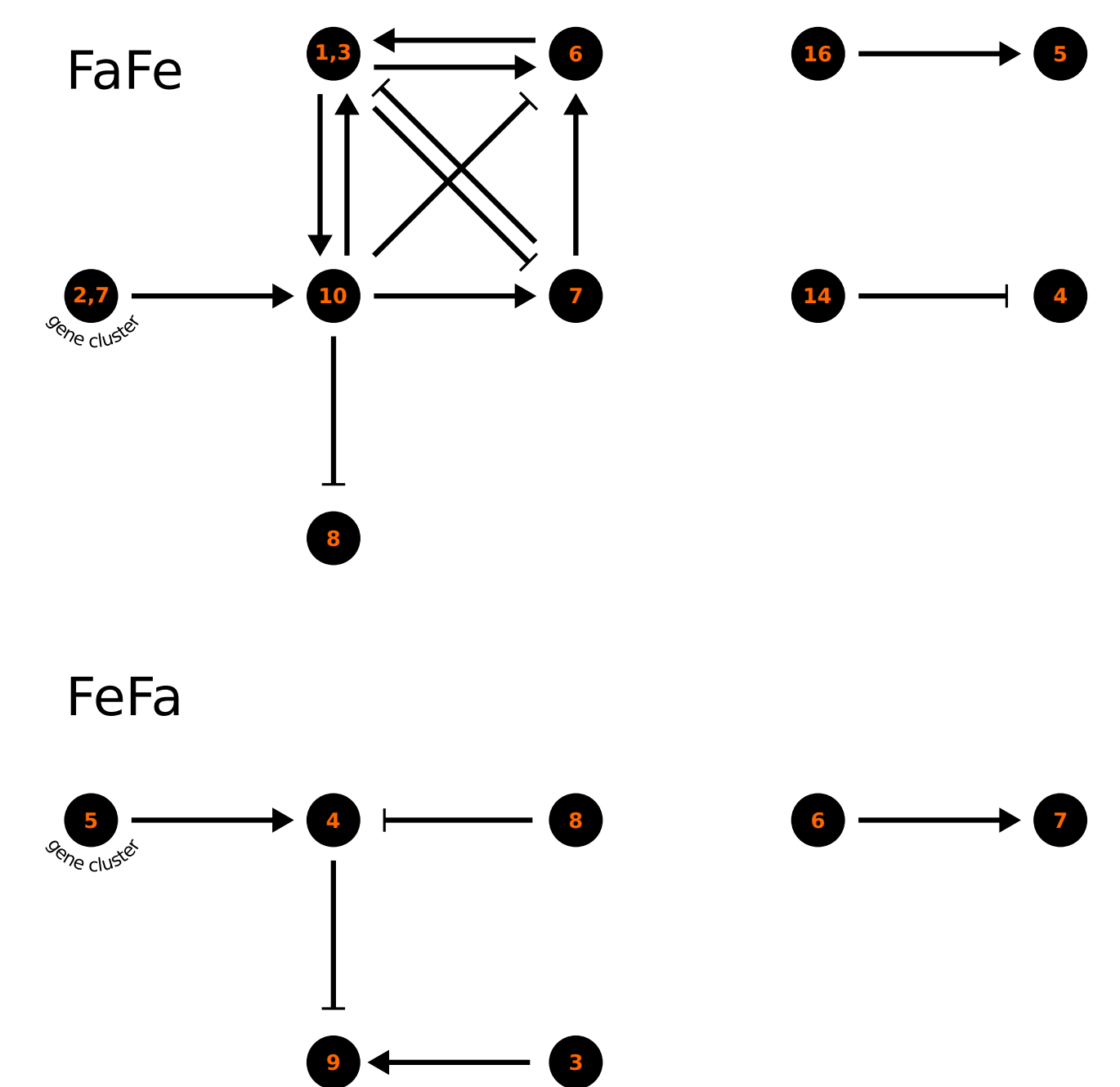
from the correlation between relative intensities and rates of change both for genes and proteins.



## Mechanism Identification

### Identification:

reconstruction of active physiological mechanisms: enabling automatic hypothesis generation and testing.

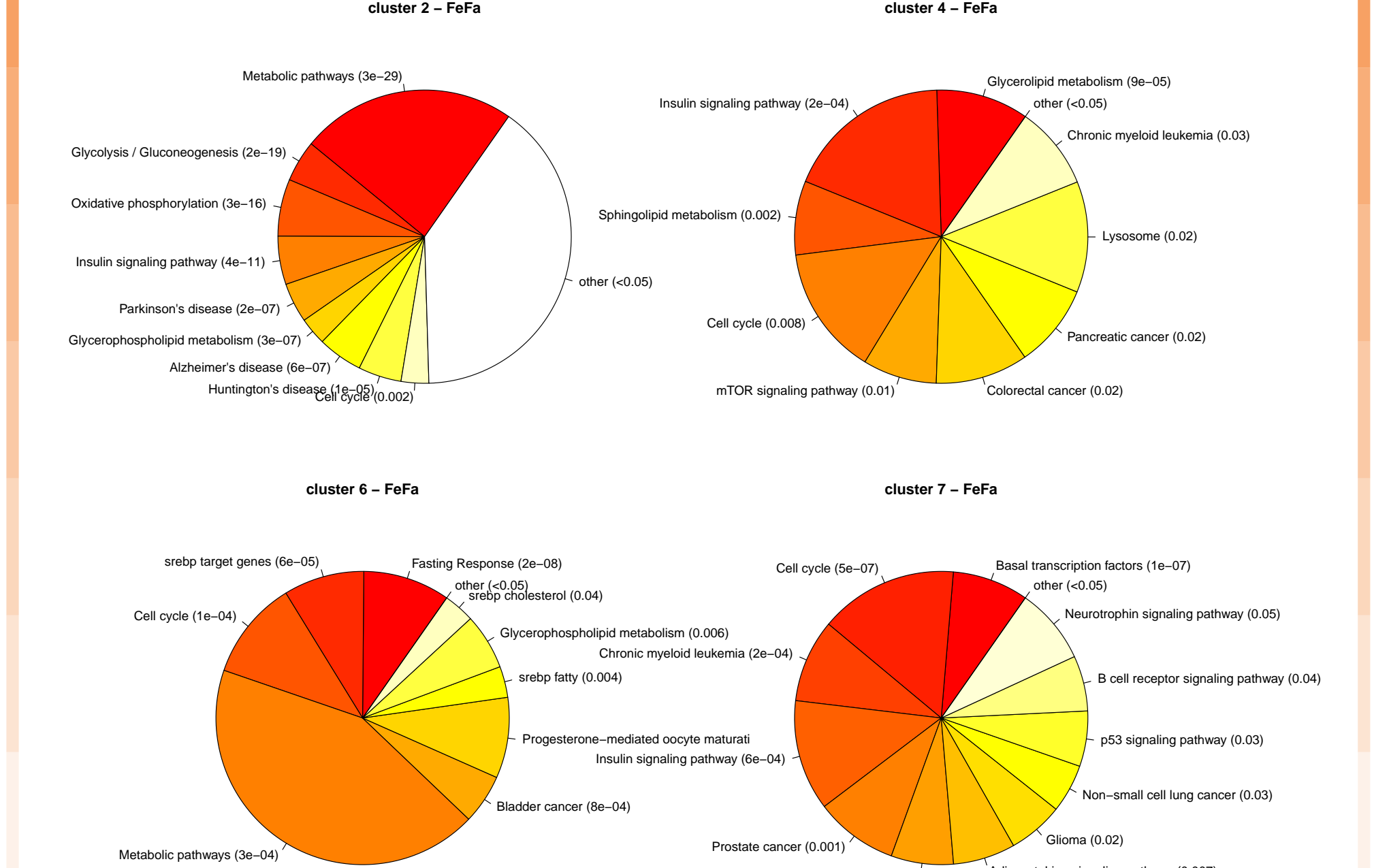


## Gene Enrichment

### Functional Enrichment:

mapping gene clusters to sets of pathways with information coming from multiple sources (such as KEGG, literature, etc.)

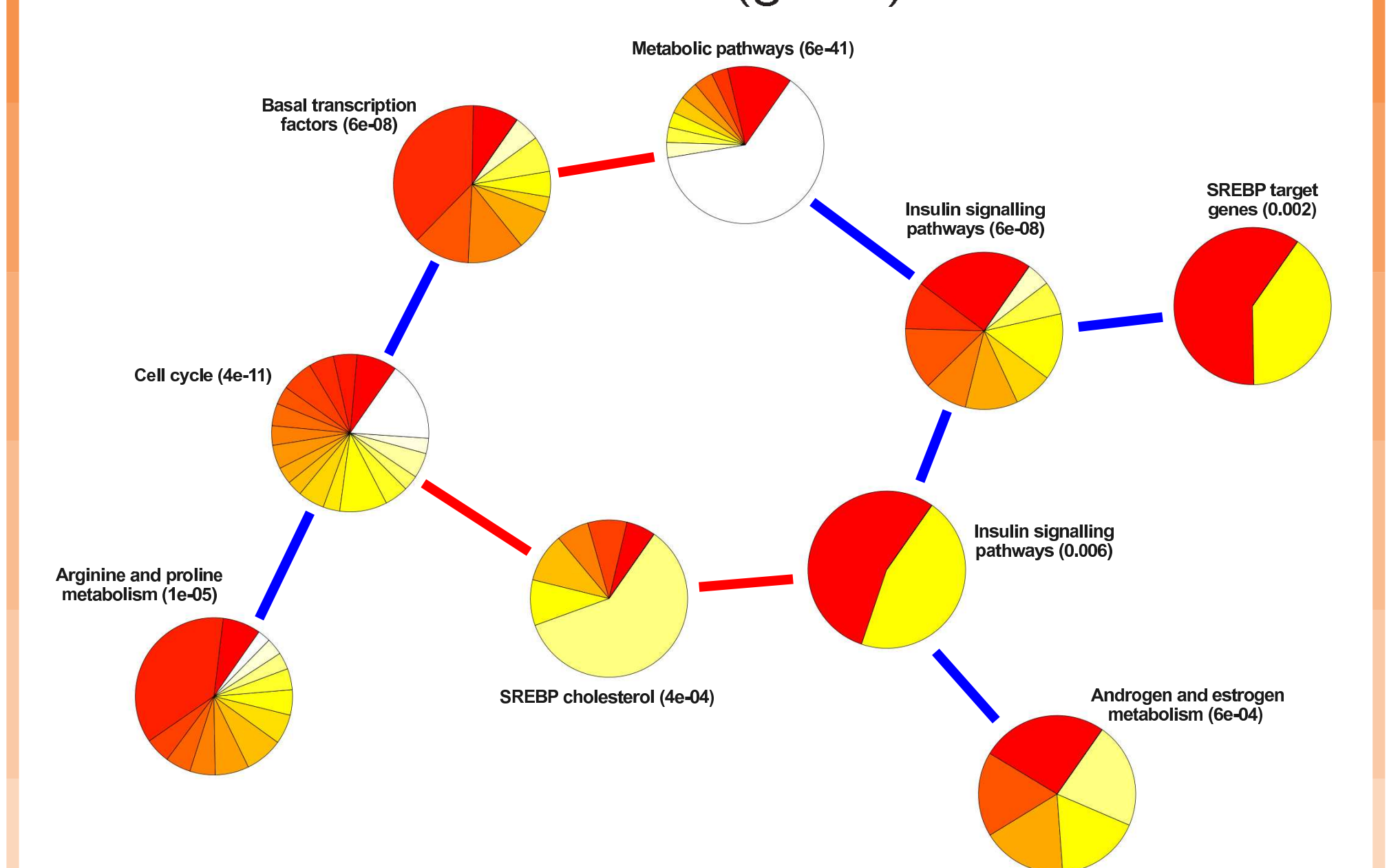
Annotation examples, feeding-to-fasting transition (p-value of exact Fisher test in brackets):



## Discussion and Outlook

Finally, mechanism identification is combined with annotation knowledge to obtain predictive models of the physiological process.

### Annotated causal network fasting-to-feeding transition (genes)



The introduced approach is general and its applicability goes beyond the target application. It enables automatic generation and testing of hypotheses from heterogeneous and highly uncertain data.

## Acknowledgments

The authors thank all the collaborators for insightful discussions and helpful interactions (in particular Niklaus Fankhauser who performed the enrichment analysis). This project was financed with a grant from the Swiss SystemsX.ch initiative, evaluated by the Swiss National Science Foundation.