

LiSym Pillar III: Theis Group

Richa Batra*, Nikola S. Mueller, Fabian J. Theis
*richa.batra@helmholtz-muenchen.de

Overview

For initial data analysis we will use standard differential expression analyses to identify regulation of proteome/phosphoproteome and transcriptome of ex vivo cultivated and stimulated hepatocytes. The temporal dynamics will be statistically analysed with generalized additive regression models to account for time and APAP or LPS treatment. To identify jointly regulated pathways the multi-omics ontology analysis method “MONA” will be applied to test for enriched pathway activity based on a Bayesian network. To infer the ACLF-specific multi-omics regulatory network, elastic net regression and support vector machines will be used to link molecular levels. To achieve translation of results to humans, bioinformatic analysis will be performed to identify orthologous genes using public databases and the best model will be validated based on measurements of human material.

Roadmap

Year 1

Standard analysis (proteome, phosphoproteome, transcriptome) of hepatocytes : Klingmüller Lab

Year 2

- Generalized additive regression model : APAP and LP induced changes

- Identify underlying pathways

- Pre-selection of ACLF biomarker

Human Liver Tissue: proteomics

Experiments: Barbieux, Klingmüller Lab

Human Liver Tissue (post surgery)
non- cirrhotic liver (3 samples) cirrhotic liver (3 samples)

Global Proteome Analysis (4K proteins)

Bioinformatics: Batra, Theis Lab

Log transformation

cirrhotic/ non-cirrhotic log fold change

Human Liver Tissue: proteomics

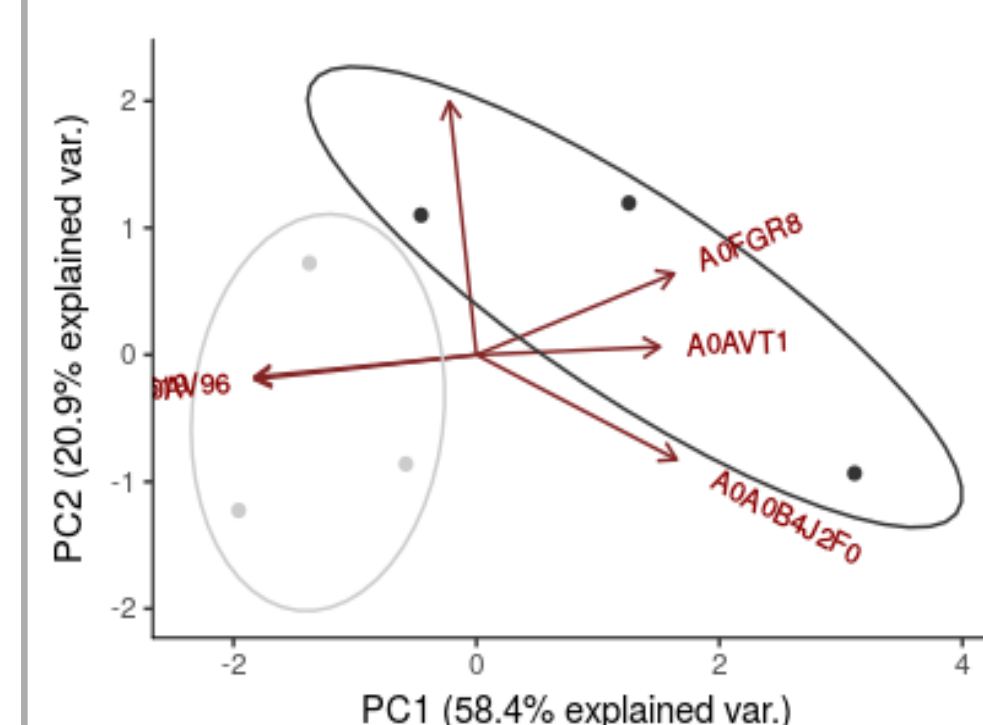


Figure1: Principal component analysis of log transformed data

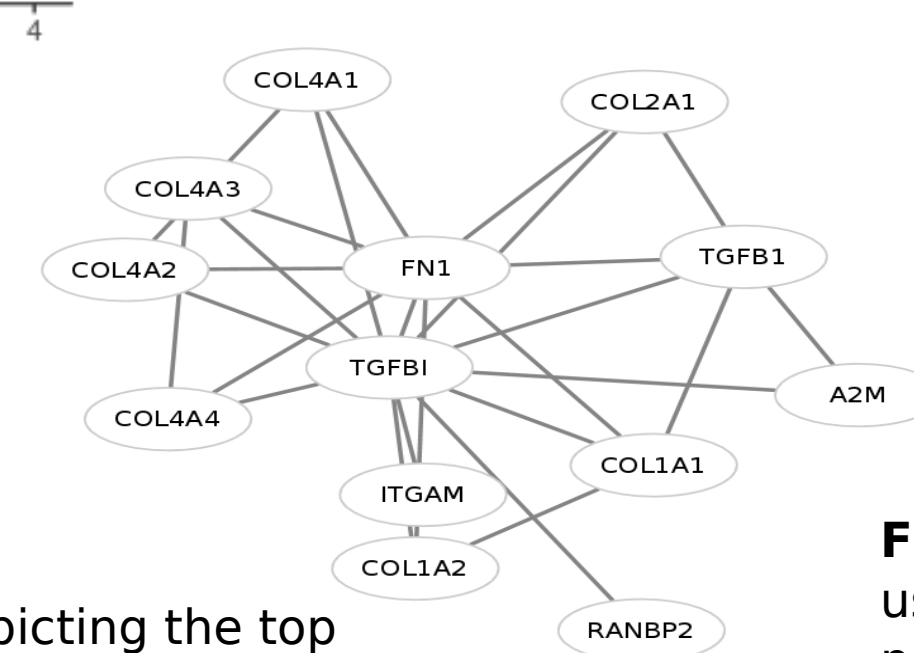


Figure2: Subnetwork depicting the top deregulated genes in Cirrhotic liver tissue

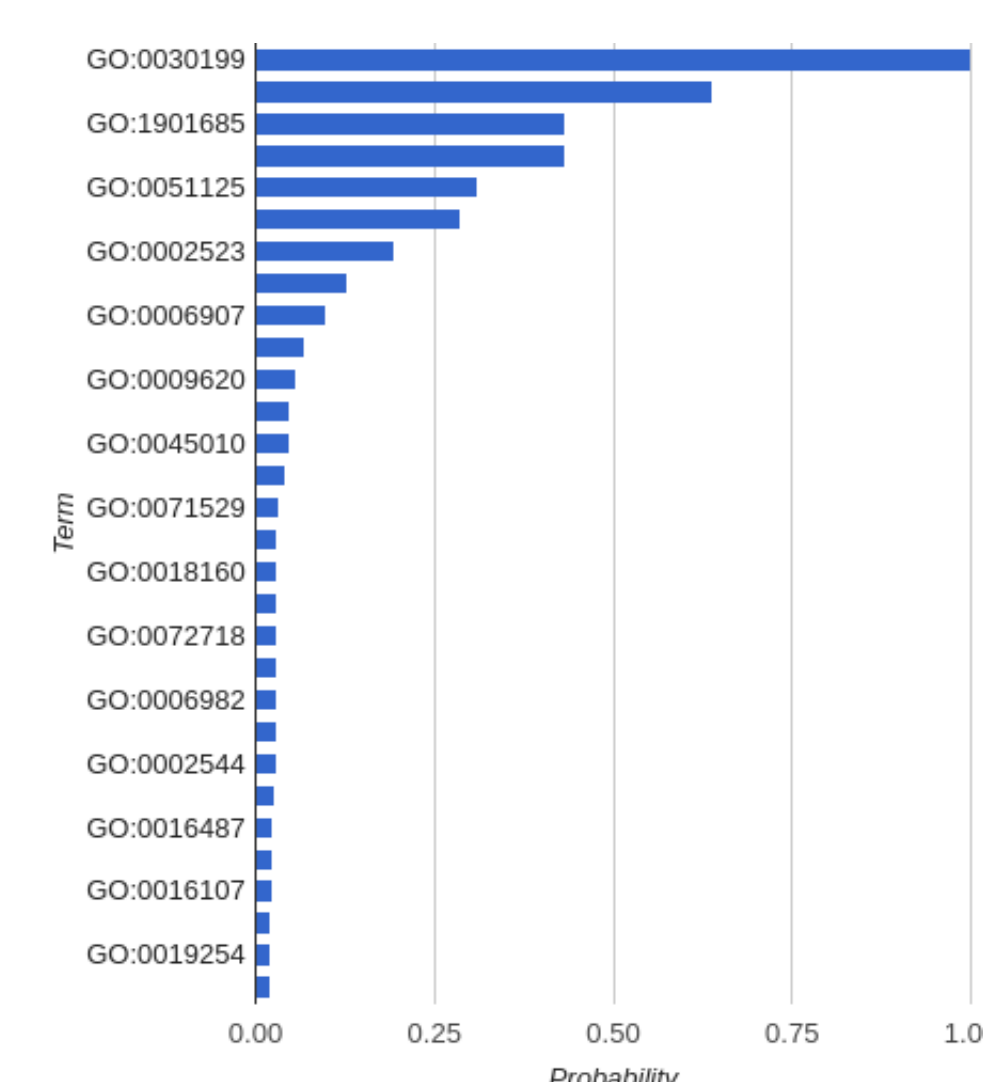
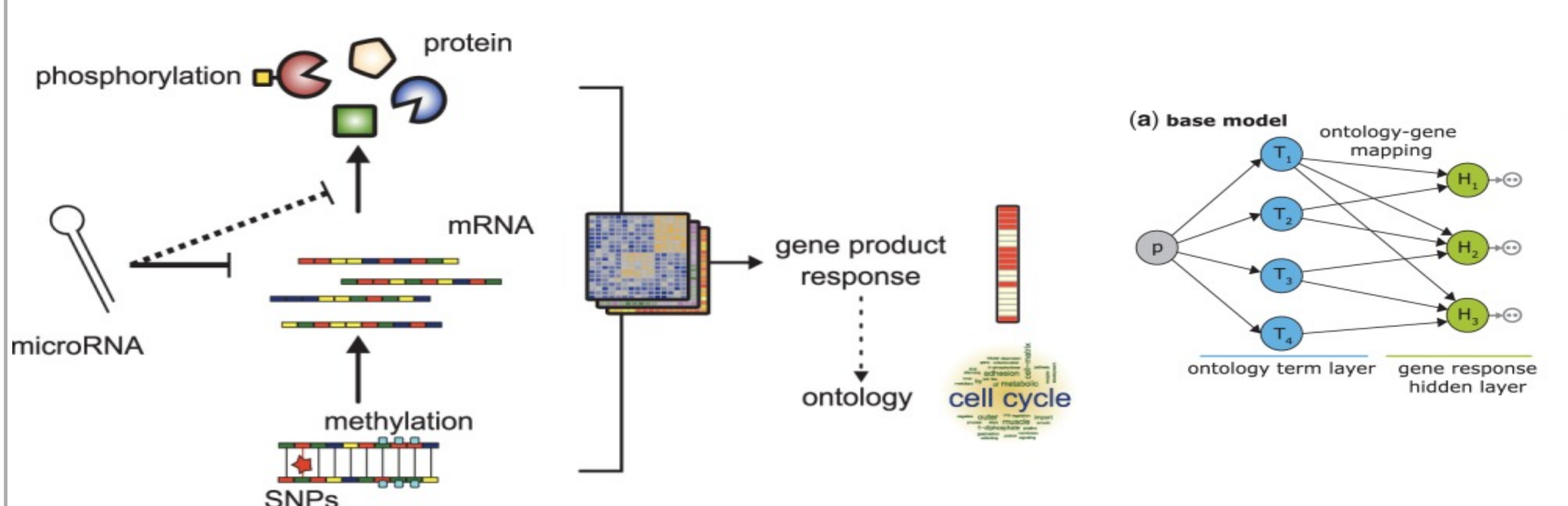


Figure3: Gene Ontology enrichment using in-house tool RAMONA. X-axis depicts probability of the term being enriched by chance

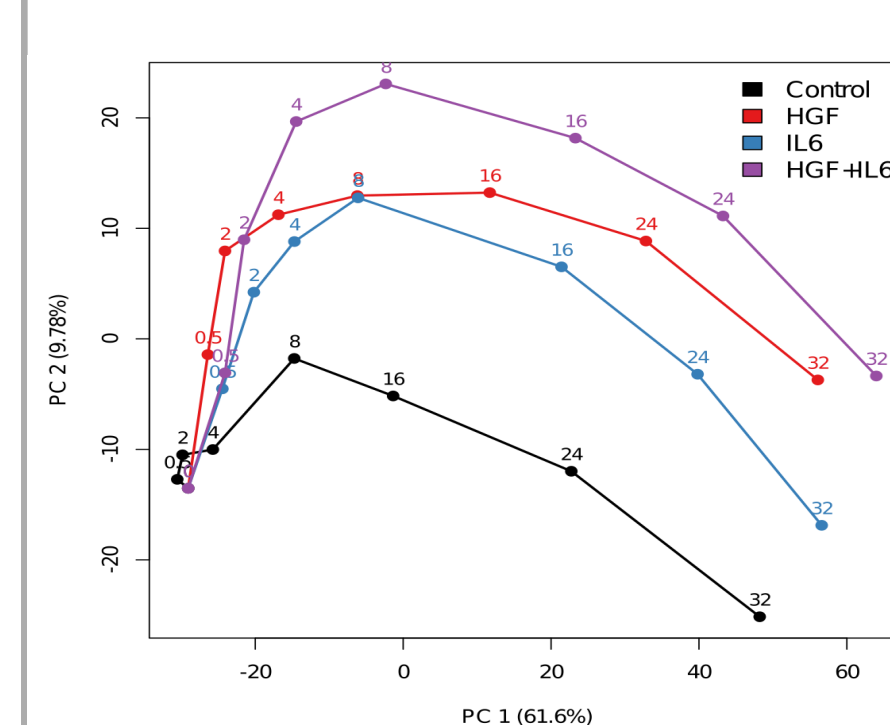
Multi-Omics Gene Ontology Enrichment



Bayesian enrichment method for the inference of overrepresented biological processes among given gene sets. Overrepresentation is quantified by interpretable term probabilities. It is able to handle data from various molecular levels, while in parallel coping with redundancies arising from gene set overlaps and related multiple testing problems

Saas et. al NAR (2013)

Progressive analysis

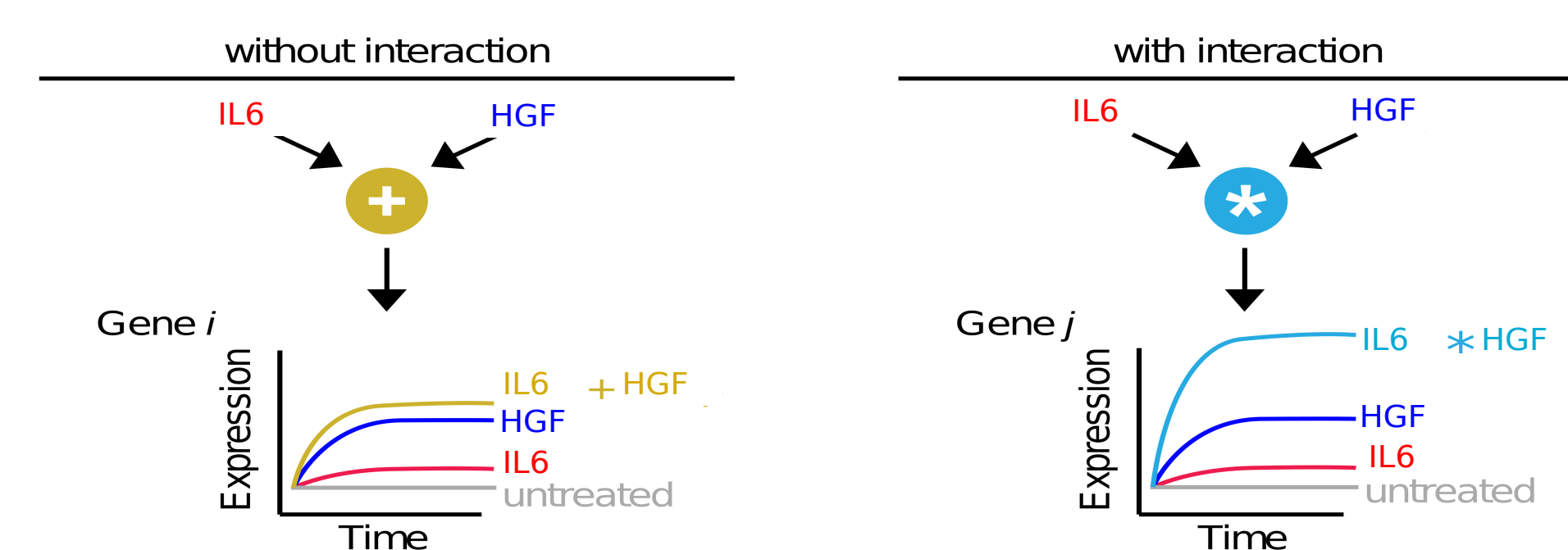


Sebastian Vlaic, Hauke Busch, Melanie Boerries, Lorenza D'Allesandro, Ursula Klingmüller, ...

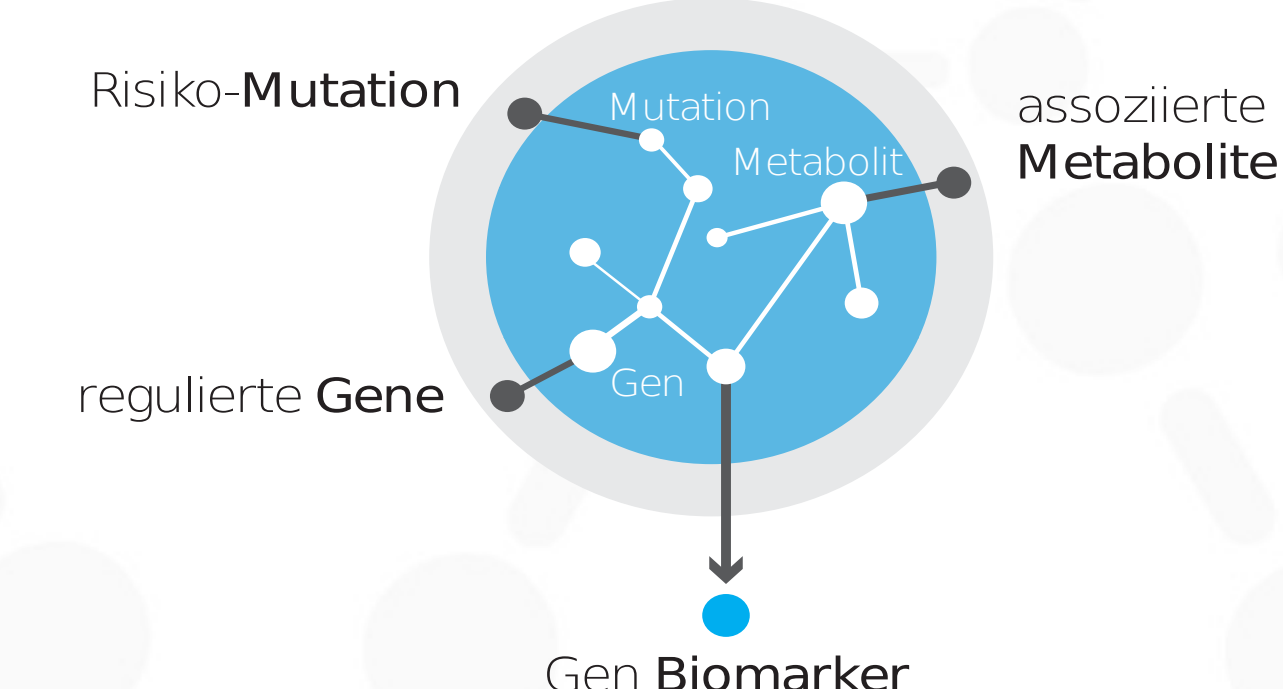
Generalized additive models with model selection

1. Time-sensitive
2. Appropriate integration of co-stimulation effects

$$y \sim s(t) + \beta_1 * [IL6] + \beta_2 * [HGF] + \beta_3 * [IL6 * HGF]$$



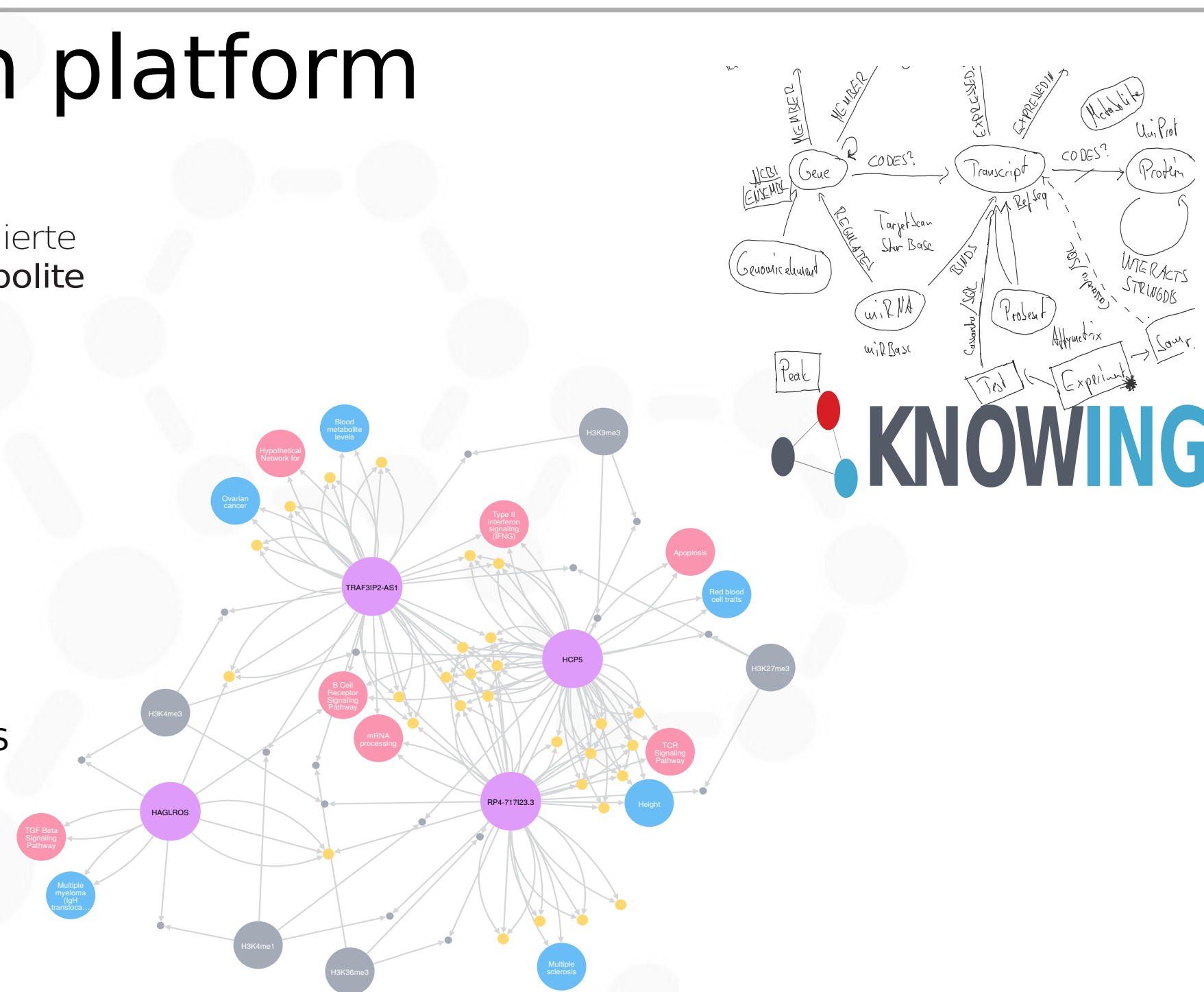
Output integration platform



Use case: Integration of heterogeneous/ multi-OMICS derived research results on LncRNA function

- Holistic graph-database analysis to identify common mechanisms of action

Preusse et. al RNA Biology (2015)



Acknowledgements

Helmholtz Zentrum Munich

Ivan Kondofersky
Marin Preusse
Fabian Theis
Theis Lab and ICB

KHI, Jena

Sebastian Vlaic



Nikola Müller
Group Leader, Computational cell maps
Institute of Computational Biology
Helmholtz Zentrum Munich

DKFZ

Claire Barbieux
Martin Böhm
Nao Iwamoto
Lorenza D'Allesandro
Marcel Schilling
Ursula Klingmüller
Klingmüller Lab

DKFZ, Uni Freiburg

Hauke Busch
Melanie Börries



Fabian Theis
Head, Institute of Computational Biology
Helmholtz Zentrum Munich

funded by



Federal Ministry
of Education
and Research