

Joint analysis of multi-omics data using R Finding drug targets in signal transduction networks using TRANSFAC®

Philip Stegmaier

geneXplain GmbH

ECCB 2018 - 17th European Conference on Computational Biology

Content

- Introductory
- Setup
- Hands-on

The geneXplain company

www.genexplain.com

Location: Wolfenbüttel, Germany

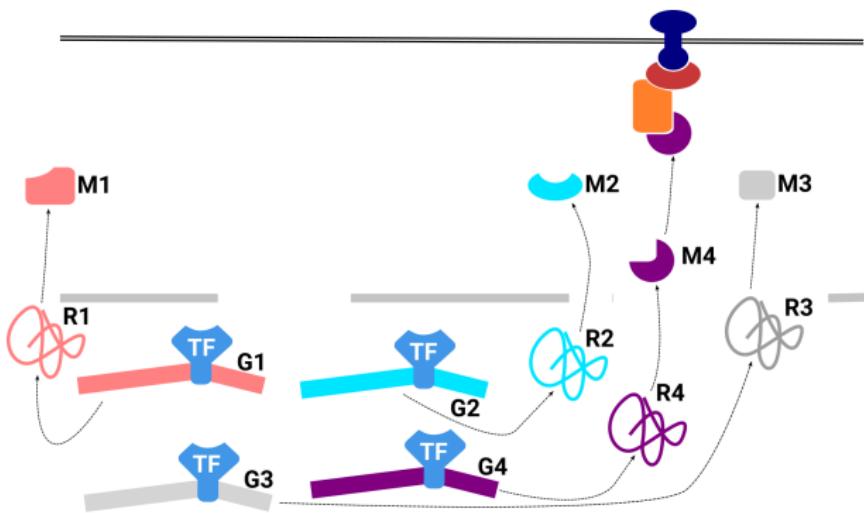


Software and databases for
bioinformatics, systems biology,
cheminformatics

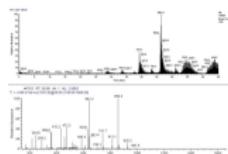
- geneXplain platform
- TRANSFAC®, TRANSPATH®,
HumanPSD™
- PASS, PharmaExpert, GUSAR

The screenshot shows the geneXplain website homepage. The header includes a 'Log in' button and a search bar. Below the header, there's a navigation menu with links to Home, Projects, Services, News, Research, Partners, About us, and Contact. The main content area features a large image of a fish swimming in water, with the text 'Upstream analysis' and a subtext: 'Discover unexpected causal relationships in your data with our unique upstream analysis.' At the bottom, there are two news cards: one titled 'Bioinformatics' with a small image of a brain and another titled 'News' with a date 'August 8, 2017'.

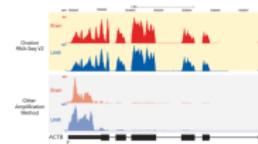
Finding causal molecular targets



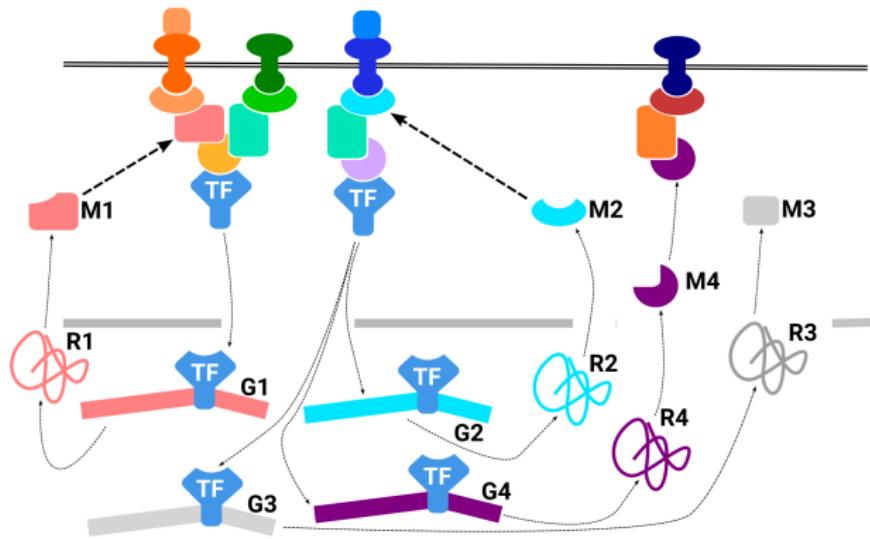
Proteomics



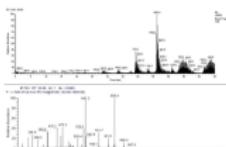
Transcriptomics (Epi-)Genomics



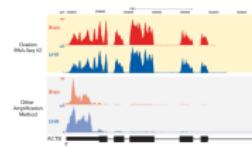
Finding causal molecular targets



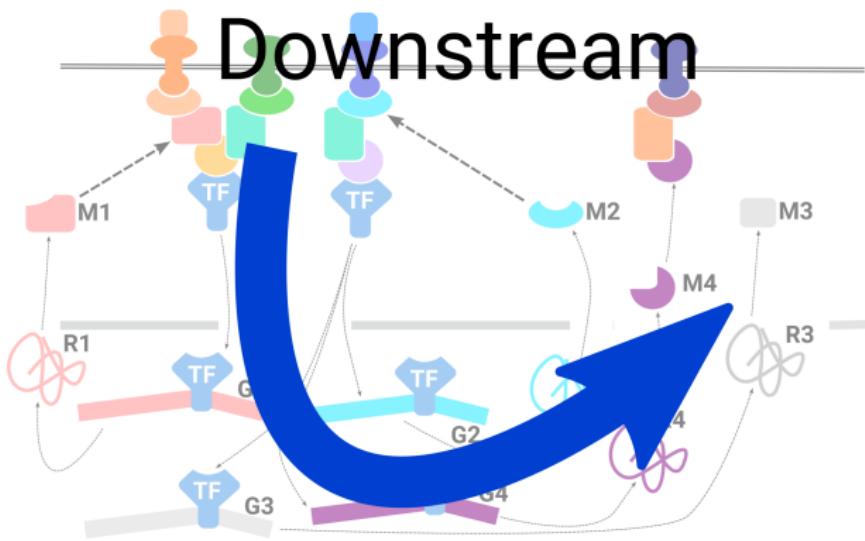
Proteomics



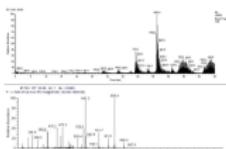
Transcriptomics (Epi-)Genomics



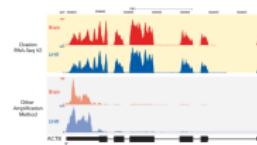
Finding causal molecular targets



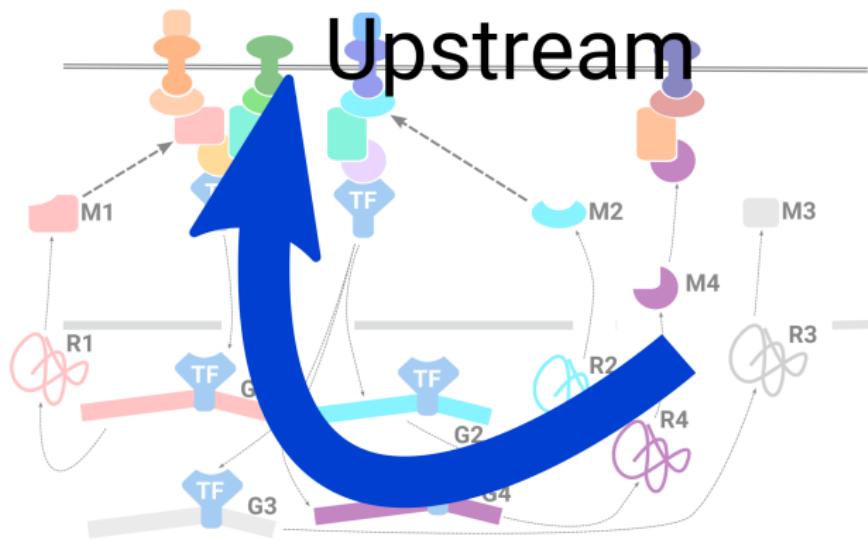
Proteomics



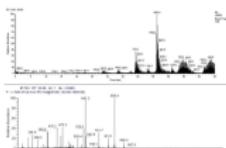
Transcriptomics (Epi-)Genomics



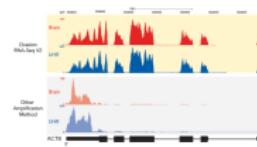
Finding causal molecular targets



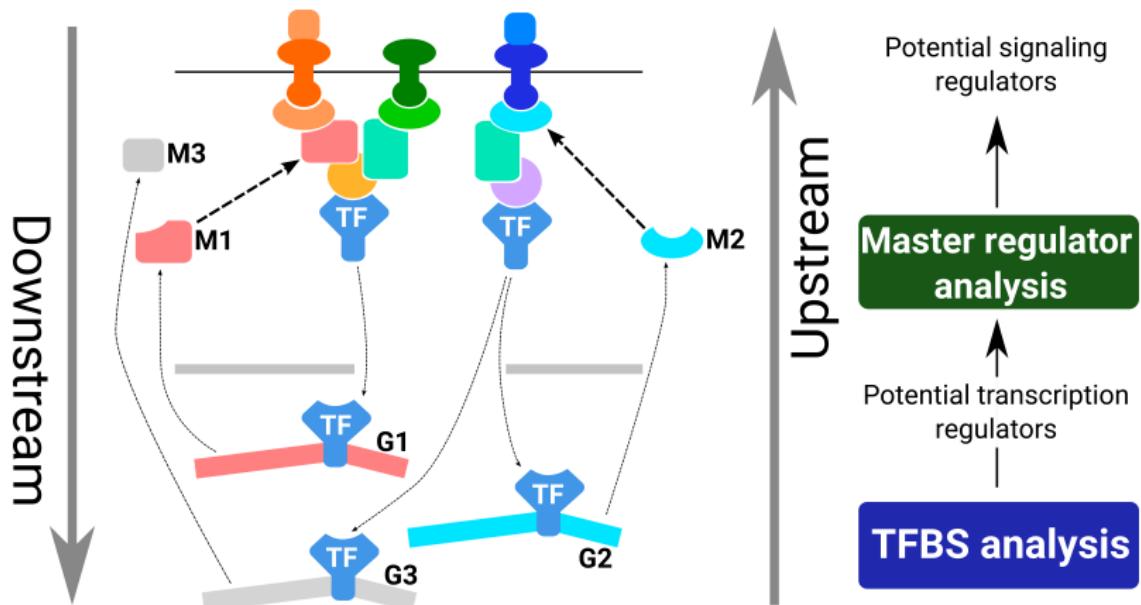
Proteomics



Transcriptomics (Epi-)Genomics



Upstream analysis



Prerequisites

- Internet connection
- R environment
- devtools package to install geneXplainR package
- geneXplainR package
- Advantageous: ability to run R script non-interactively

Installation of geneXplainR package

The screenshot shows a GitHub page for the geneXplainR package. The URL is https://github.com/genexplain/geneXplainR. The main heading is 'Installation'. Below it, a section titled 'From github.com' provides instructions for installing the package using devtools. It includes a code block with R code:

```
library(devtools)
install_github("genexplain/geneXplainR")
```

Tutorial overview

R scripts

- Tutorial consists of five scripts
- Script 0: Introduction to the R API
- Script 1: Limma analysis of RNA-seq counts (optional)
- Script 2 - 4: Drug discovery workflow

Tutorial overview

Archive with data and scripts

- Get access info for tutorial participants
- Get tutorial archive from platform
- Unpack tutorial archive
- Edit parameters in data/access.R within tutorial folder
- Start with running script 0

Tutorial overview

Example data for this session

Example data set

Nat Immunol. 2017 October ; 18(10): 1104–1116. doi:10.1038/ni.3818.

PMC

Type I IFNs and TNF cooperatively reprogram the macrophage epigenome to promote inflammatory activation

Sung Ho Park¹, Kyuho Kang¹, Eugenia Giannopoulou^{1,3}, Yu Qiao¹, Keunsoo Kang⁴, Geonho Kim¹, Kyung-Hyun Park-Min¹, and Lionel B. Ivashkiv^{1,2}

Experiment types

RNA-seq: gene expression changes

ATAC-seq: open chromatin in different conditions

ChIP-seq:

IRF1 locations with or without IFN/TNF treatment

p65 locations with or without TNF treatment

H2BUb - H2B ubiquitination, different conditions

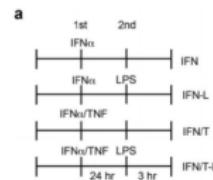
H3K4me3 - H3 Lys4 tri-methylation, different conditions

H3K27ac - H3 Lys27 acetylation, ~

H4ac - H4 acetylation, ~

The screenshot shows the NCBI GEO Accession Display page for Series GSE100383. The page includes a header with the NCBI logo and links for Home, Search, Site Map, Gene Expression Omnibus, Publications, FAQ, and Help. Below the header, there's a search bar with 'Scope: Self', 'Format: HTML', 'Amount: Quick', and 'GEO accession: gse100383'. The main content area displays experimental details: Status (Public on Jun 23, 2017), Title (Type I IFNs and TNF Cooperatively Reprogram Epigenomic Landscape of Human Macrophages to Promote Inflammatory Activation), Organism (Homo sapiens), and Experiment type (Expression profiling by array, Genome binding/occupancy profiling by high throughput sequencing).

Experiment design



Contrasts used in this tutorial:

IFNa versus no treatment

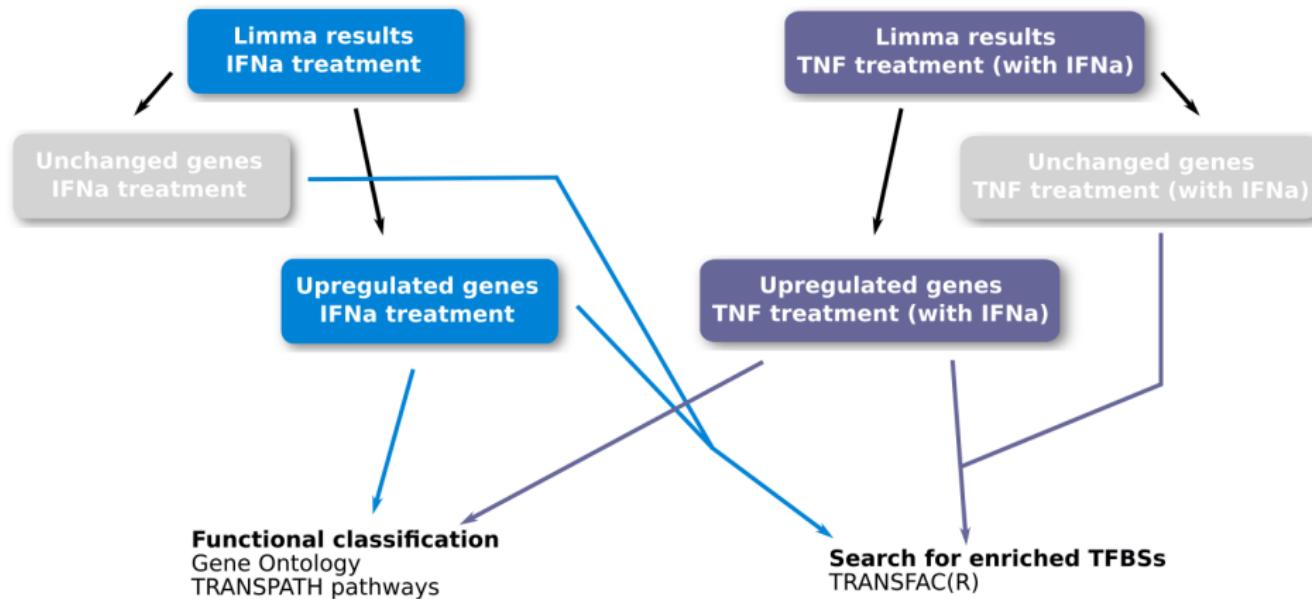
Events triggered by IFNa treatment

IFNa/TNF versus IFN treatment

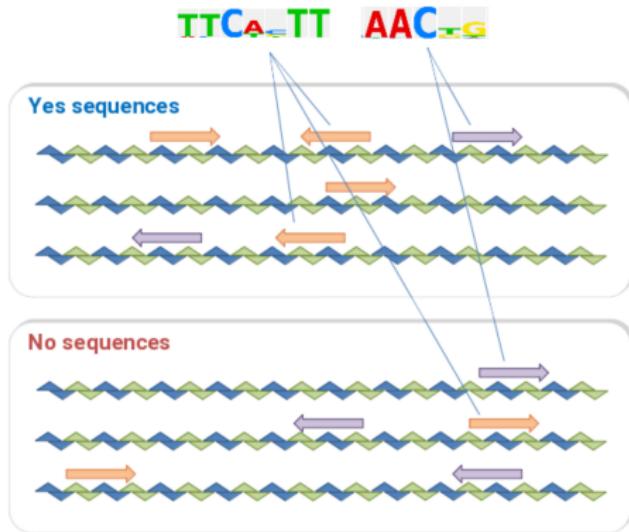
Examining TNF-related events in the context of IFNa treatment

Workflow overview

Workflow overview Script 02



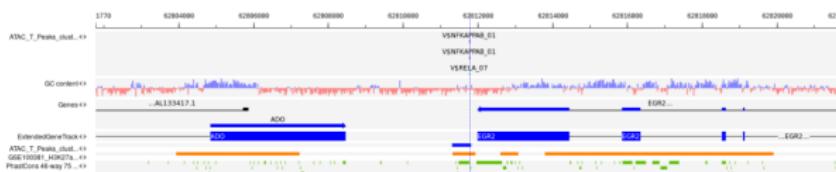
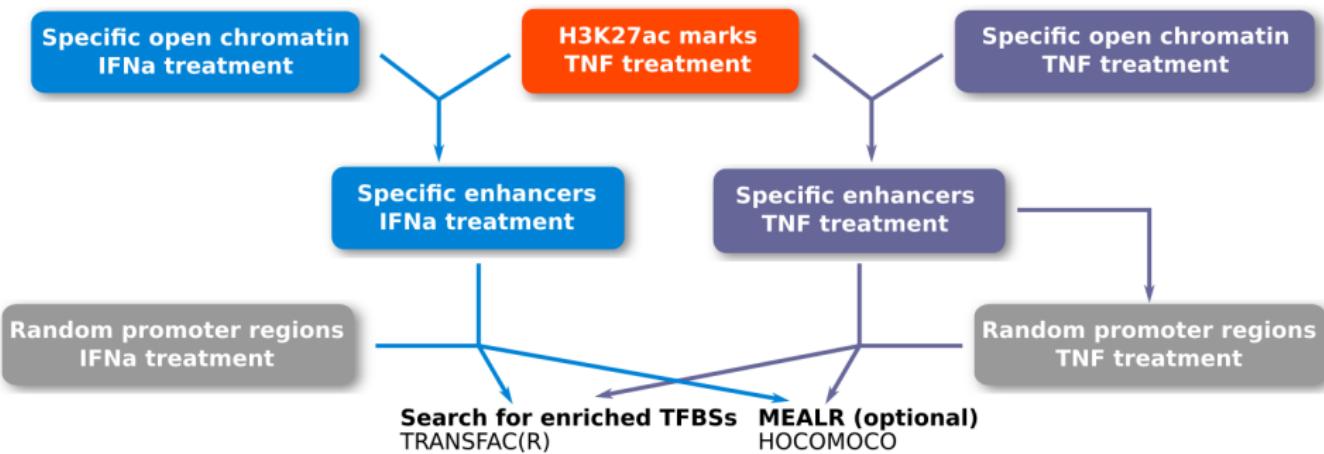
Transcription factor binding site (TFBS) analysis



- Motif enrichment analysis
- Searches for TFBSs significantly enriched in sequences of interest
- Koschmann, J., Stegmaier, P. et al., Microarrays, 2015
- Tests and reports every motif of a library
- MEALR searches for a sparse logistic regression model
- TFBS combinations as sequence classifier
- Retains a smaller selection of motifs from a library

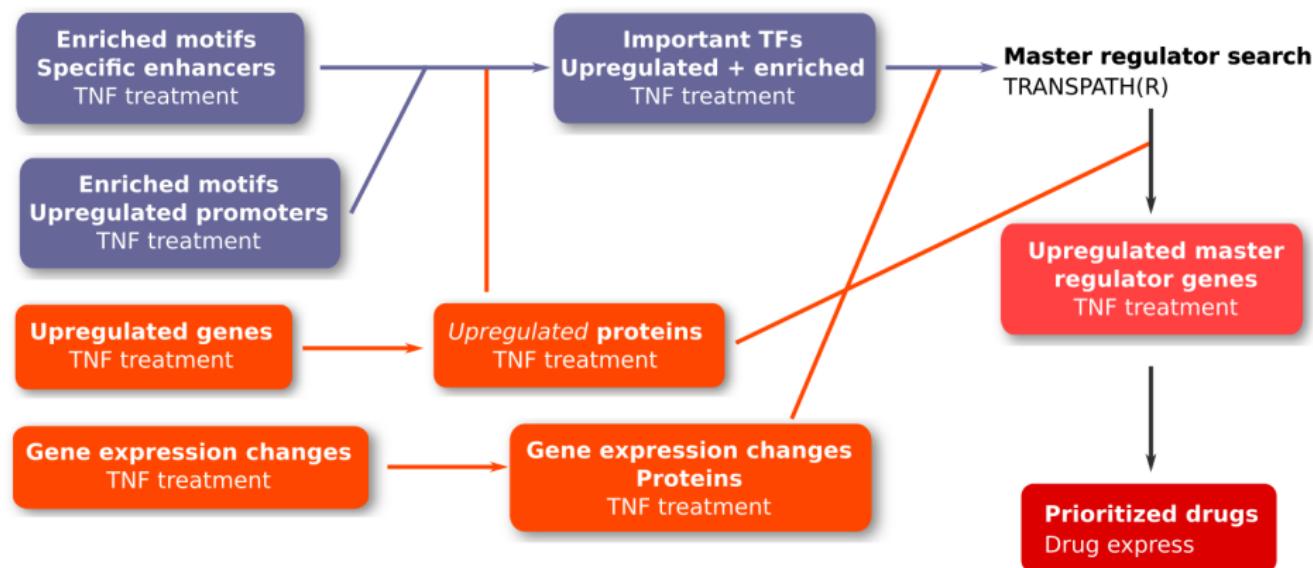
Workflow overview

Workflow overview Script 03



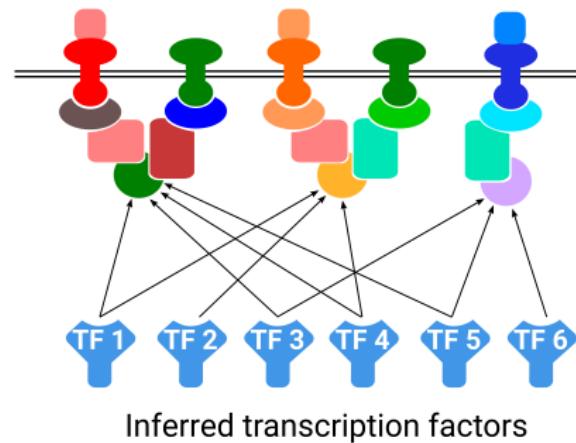
Workflow overview

Workflow overview Script 04



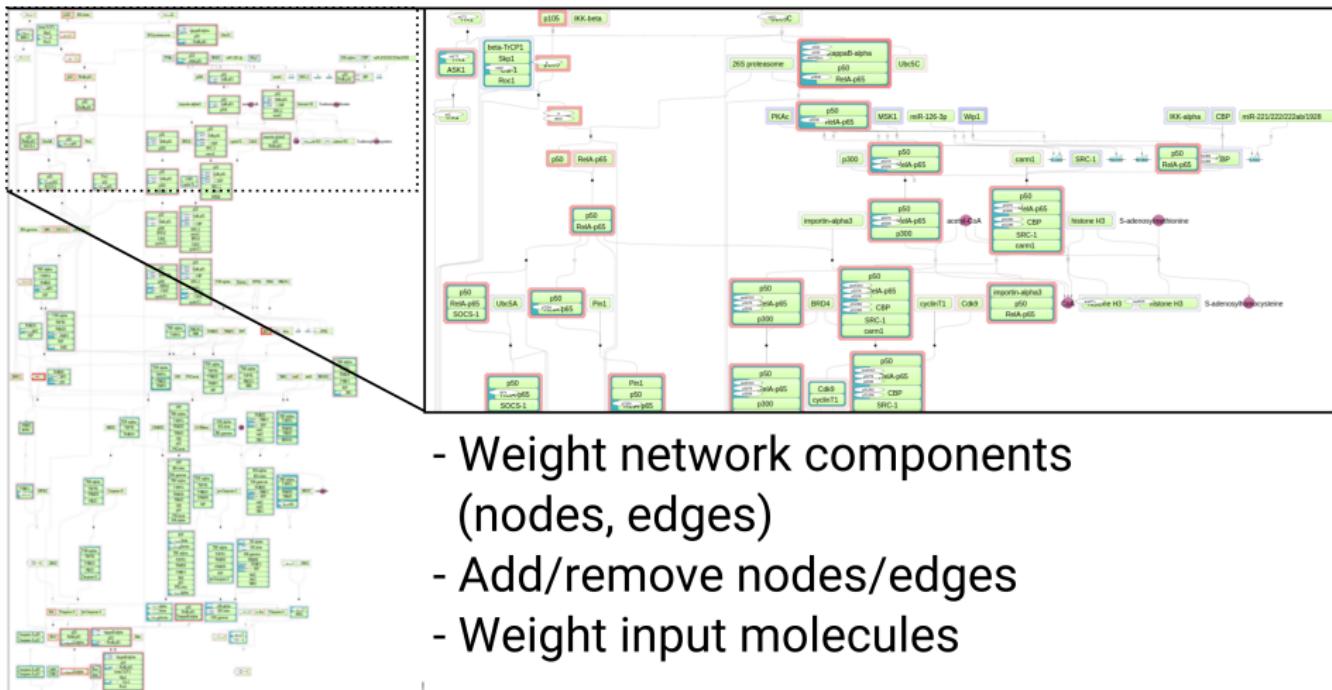
Master regulator analysis

- Constructs pathways (single regulators or effectors, multiple input molecules)
- Uses knowledge about pathways, cascades, molecular network(s)
- Stegmaier et al., PLoS One, 2011
- Traverses molecular network to find best common upstream regulator
- Takes into account network topology
- Downstream: effector analysis
- Capable of integrating additional information, e.g. experimental



Master regulator analysis

Interfaces for data integration



Acknowledgements

Management and staff of geneXplain

CEO Edgar Wingender

CSO Alexander Kel

Product manager Jeannette Koschmann

Database engineer Mathias Krull

Database editor Volker Matys

Colleagues around the world

BioUML developers
in Novosibirsk, RU,

Testers and curators
in Bangalore, IN

Current grant projects

- COLOSSUS (EU Horizon 2020)
- ExiTox II (BMBF)
- GlioTrain (Horizon 2020)
- miRCol (BMBF/DLR)
- miRNA DisEASY (EU Horizon2020)
- MyPathSem (BMBF i:DSEM)
- Optogenetrapy (EU Horizon2020)
- OxidoCurin (Eurostars)