Joint analysis multi-omics data to find drug targets in signal transduction networks using TRANSFAC® and R

Tutorial details

Date: September 9, 2018
Time: 13:30 – 17:00 (half day tutorial)
Venue: TBA

Tutors

- Philip Stegmaier, geneXplain GmbH, Germany

Summary

This tutorial will first provide an introduction to a selection of bioinformatics and systems biology tools for the analysis of gene regulatory mechanisms and to reveal drug targets. We will introduce tools such as: Match, MEME, ChIPMunk, PathFinder, and the databases: TRANSFAC, TRANSPATH, REACTOME, KEGG, HumanPSD as well as many other algorithms and tools that can be combined in task-oriented pipelines using web platforms such as BioUML (biouml.org; genexplain.com) and Galaxy (usegalaxy.org) as well as using R and Java APIs in your own applications. We will also explain the steps researchers should follow to get their tools running inside such platforms and how to build your specific pipelines that will use your own tools together with rich collection of other powerful tools in order to lead you quickly to the discovery. In the second part we will have a practical session where attendees will face examples of multi-omics data in different disease/biological conditions and will be building analysis pipelines for drug target discovery.

Transcription regulation is of central importance for nearly all processes in a living system, and erroneous transcription control is causative for numerous diseases. To enable a systems approach to transcription, we still have to struggle with the very first step that is to infer underlying wiring diagrams. Empirical information about the interaction of regulators (transcription factors) and the regulated target genes, obtained by either conventional or high-throughput methods, has been collected in the TRANSFAC database for over 30 years, and statistical models inferred from this information have been included as positional weight matrices (PWMs) and made available for the prediction of regulatory sites as well. New extensions include syntax (relevant combinations) and semantics (regulated processes) of regulatory sites. Extended annotation of gene-disease associations is available in the Human Proteome Survey Database (HumanPSD), connected with signaling pathways that control the activity of TFs (TRANSPATH database). All this carefully curated information can be used in full power to analyze disease related multi-omics data using the BioUML/geneXplain platform, which helps to decipher the molecular mechanisms of disease often on very early stages of its progression. First of all, differentially expressed genes revealed by microarray or RNA-seq analysis are combined with genomic (SNP) and/or epigenomic (ChIP-seq/ATAC-seq, DNA methylation) assays to find disease-related enhancers. Next, genetic algorithms reveal TFs synergistically acting in those enhancers. Finally, topology analysis of signal transduction networks upstream of transcription factors identifies master-regulators of the disease progression, which proposed as perspective therapeutic targets.

The goal is that users attending the tutorial acquire basic knowledge in:

- Concepts of molecular mechanisms of gene regulation. Regulatory code.
- How to write your own R script to analyze multi-omics data
- Practical experience with drug target discovery.

Target audience

The course addresses researchers in biology, bioinformatics, biochemistry and medicine with interest in gene regulation and related topics.

Prerequisites

- Laptop with installed and running Internet browser (Chrome, Firefox, Opera, Safari), is to be used during the tutorial
- Basic biochemistry
- Elementary computing skills
- Background knowledge in bioinformatics is not absolutely necessary, but may be instrumental

Schedule

Time	Subject
13:30	Welcome
13:30 – 14:30	Introduction to drug discovery workflow and tutorial (RNA-, ATAC-, ChIP-seq data, BioUML/geneXplain, TRANSFAC®, Reactome, TRANSPATH®, HumanPSD™)
14:30 – 14:45	Preparation for practical session, setup of workspaces, package installation
14:45 – 15:00	Coffee break
15:00 – 17:00	Practical session: RNA-seq differential expression analysis, analysis of TF binding sites, pathways and networks, drug discovery workflow