Softwarewerkzeuge der Bioinformatik

Prof. Dr. Volkhard Helms
PD Dr. Michael Hutter, Kerstin Reuter, **Daria Gaidar**
Wintersemester 2017/2018

Saarland University
Department of Computational Biology

Project 3

01. Februar 2018

Tutoren: Lea Eckhard, Markus Dillmann

Pathway and Network Analysis of -Omics Data

This year we will focus on the pathway and network analysis of omics data. Continuing the practice of bringing to you the recent material to learn from the best, we will follow the course material from Canadian Bioinformatics Workshop that took place in June 2016.

Please follow this link to find throughout documented lectures and tutorials.

http://bioinformatics-ca.github.io/pathway_and_network_analysis_of_omics_data_2016/

Your task: Process the material under Day 1 to Day 3 and optionally the pre-workshop tutorials (they wont be graded).

What will be graded: A well documented report filled with your notes, results (in text and figure forms) and their discussion. In some instances (e.g. Module 4 Lab Practical) you will find a list of questions to answer and elaborate on. Please do your best.

How to:

- Work in the group of 3. The default group assignment is as stated earlier, see the file following the link. Cases of migration shall be self managed and well labeled with the names of the authors on the submission files
  https://www-cbi.cs.uni-saarland.de/wp-content/uploads/Softwarewerkzeuge_WS_17-18/teilnehmer_projekte_ws1718.pdf
- Submit one report from the group
- Submissions are accepted in electronic (daria.gaidar(at)bionformatik.uni-saarland.de) and/or printed versions
- Language on the reports' text is german or english
- Deadline is Thursday, 01.02.2018, at 10:30 in the morning.

Module 1: Introduction to Pathway and Network Analysis

Lecture

Module 2: Finding Over-Represented Pathways in Gene Lists

Lecture
Lab practical: Enrichment-Based Analysis - Performing ORA

(a) Learn how to use GSEA to explore the results

(b) Use g:Profiler to perform gene-set enrichment analysis

Module 3: Network Visualization and Analysis with Cytoscape

Lecture Part 1 and 2
Lab practical: Cytoscape Demo, Enrichment Map

(a) Create an EnrichmentMap from GSEA results and navigate through the network
During this exercise, you will learn how to create an EnrichmentMap from gene-set enrichment results. The enrichment tool chosen for this exercise is GSEA but an enrichment map can be created from various gene-set tools using the generic format or the more specific g:Profiler or BiNGO interface.
Post analysis (add drug target gene-sets to the network):
As second part of the exercise, you will learn how to expand the network by adding an extra layer of information.
Autoannotate:
A last optional exercise guides you toward the creation of automatically generated cluster labels to the network.

(b) Create an EnrichmentMap from g:Profiler results

(c) Integrated assignment: g:Profiler/EnrichmentMap

Module 4: More Depth on Pathway and Network Analysis

Lecture Part 1 and 2
Lab practical: De Novo Subnetwork Clustering Analysis in Reactome
This exercise will provide you with an opportunity to perform and network analysis using the Reactome Functional Interaction (FI) and the ReactomeFIViz app.

(a) Analyze somatic mutation data to identify biology that contributes to ovarian cancer.

Module 5: Gene Function Prediction

Lecture
Lab practical: GeneMANIA (web version)
Create GeneMANIA networks starting from a single gene to predict its function or starting from a gene list. Explore and understand the main output features of GeneMANIA such as the network composition or the enriched functions.

(a) Imagine that you are interested in exploring the function of the human GRN gene: GRN returned as the strongest hit from your omics experiment but not many information about this gene is available in functional databases. Use GeneMANIA to identify its predicted function as well as potential interaction partners.

(b) Here the task gets group specific. For the reporting please substitute the 30_prostate_cancer_genes.txt file with the text file containing the followong gene sets to be found on the http://www.cbioportal.org/. See the drop down menu lower on the page. For the report omit Step 12.

(1) General: Invasion and Metastasis (27 genes)

(2) General: Ras-Raf-MEK-Erk/JNK signaling (26 genes)
(3) General: PI3K-AKT-mTOR signaling (17 genes)
(4) General: Survival / cell death regulation signaling (23 genes)
(5) General: Notch signaling (55 genes)
(6) General: Cell Cycle Control (34 genes)
(7) Ovarian Cancer: Oncogenes associated with epithelial ovarian cancer (17 genes)
(8) Ovarian Cancer: Putative tumor-supressor genes in epithelial ovarian cancer (16 genes)
(9) Glioblastoma: RTK/Ras/PI3K/AKT signaling (17 genes)
(10) Prostate Cancer: Down-regulated by androgen (19 genes)

(c) Also group specific. Create the mixe_gene_list.txt from your gene list from previous exercise plus one set of genes above and one below in the drop down list found on the http://www.cbioportal.org/. Group 1, plese take the list of your choice as a substitute for the list below. Choose your favourite gene for the analysis where in the provided example it is PDPK1.

(d) Integrated assignment: g:Profiler/EnrichmentMap
All OPTIONAL question are obligatory to do.
The first integrated assignment showed how transcriptomics and pathway analysis is used for trying to understand how a cancerous tissue has evolved from normal tissue. Another common strategy it to sequence the DNA of cancerous tissue to find out the mutations. MutSig is a mutation signal processing tool created by Broad Institute. It estimates the significance of the gene mutation rate based on abundances of the mutations, clustering of the mutations in hotspots and conservation of the mutated positions.

Module 6: Regulatory Network Analysis

Lecture
Lab practical: iRegulon
Import a Cytoscape network and apply iRegulon on all the selected nodes. Explore and understand the main output features of iRegulon such as the Transcription target view. Learn how to display predicted targets of a specific transcription factor by creating its metatargetome.

(a) Detect regulons from co-expressed genes
In this exercise, we will continue to use the genes from the prostate cancer list from the GeneMANIA assignment. iRegulon requires a network from the start, and we will use the GeneMANIA network that we previously saved for this purpose. Using iRegulon, we will look for transcription factors (TFs) that may regulate a set of genes in this network.

(b) Create a metatargetome using iRegulon and merge 2 networks in Cytoscape.
This exercise will teach you to use the metatargetome function of iRegulon. This function displays a list of potential targets for a specific TF. We will create the metatargetome of two TFs, that we found as potential coregulators of the prostate cancer genes (exercise 1): MTF1 and LARP4. We will then learn how to use Cytoscape to merge two networks and visualize nodes in common.

Have fun!