

TP - Introduction to Network Analysis with Cytoscape

Save your work regularly!!!

Part 1

Create a protein-protein interaction network for a protein of interest

Import data from PSICQUIC

File → Import → Network → Public Databases

Write the name of the protein "HSPB1" → search

The databases are parsed, and the number of interactions found in each database for the protein of interest is shown. The choice of the database to use will depend on the type of data you want to retrieve: Each database contain different (but overlapping) data, and each have a different "philosophy".

Widely used databases are Reactome-FI, Intact and BioGrid, for instance.

Create 2 networks: one from the Intact database, and another one from Reactome-FIs (Merge → No).

Task 1: How many nodes and interactions are retrieved from each database for the HSPB1 protein (= what are the sizes of the networks)?

Cytoscape Panels

There 3 different cytoscape panels in the screen: Control panel (with the set of networks, the controls for the visualisation, the apps tabs ...), Networks panel (where the network are displayed) and Table panel (with all edge and node annotations).

The network created from the reactome-FI database form a connected component: all the nodes are linked to each others (i.e., no isolated nodes). This is not the case in the Intact network, which in addition contains nodes colored differently.

Task 2: Identify what happened with the Intact network? To this goal, use the node annotations given in the Table Panel below the networks

Manipulate the networks

Visually:

- Try all the buttons zoom in/out, expand on your 2 networks
- Try different algorithms for "Layout"

- Try rotate/align/scale (Layout => Layout Tools)
- Try rotate/align/scale for a selection of nodes

- In the control panel → Style, modify the default value of node size, and the default colors of nodes and edges

- In the control panel → Style, change the names of the labels to "Human readable labels"



- Choose a satisfying representation for each network (they might be different for the 2 networks), based on all the above tuning

Task 3: Export the 2 network pictures with File => Export => Network to Image.

Again, save your work regularly!!!

Adding/filtering data

Filtering data from the Intact Network

The Intact network contains many interactions for the HSPB1 protein, many of which might not be relevant. To remove node/edges, you can directly click and delete them, but you can also create automatic filters.

Create a filter for the network to remove non-human data

Control Panel, tab "Filter"

New filter "+" → Column filter → Choose column → Node taxonomy name → doesn't contain "human"

This will select on the screen all the nodes that are not annotated as "human" in the Table Panel. These can be removed with Edit → Delete selected nodes and edges.

Task 4: how many nodes/edges remain after this filtering in the Intact network?

Adding data to the Reactome-FI network

The same process that has been used to create the networks around the protein of interest can be re-applied with the interactors identified in this first step, in order to extend the network. This leads to the addition of the "interactors of the interactors" and "the interactions between the interactors".

Select all the lines in the Table Panel of the Reactome-FI network, copy (ctrl-C). Go to File → Import → Network → Public Databases, and paste the list of proteins → Search

This will give many more results. Select only "reactome-FIs" and import.

Create a degree/topology filter for this new large reactome-FI network

In the Filter tab, create a new Topology Filter, and select the nodes that have 2 or more neighbors (ie, that have at least 2 neighbors within distance 1).

Create a new network containing only the selected nodes, and their interactions. For that, File → New Network → From selected nodes, all edges.

Part 2

Using Cytoscape to integrate biological information

Different possibilities exist to integrate information using networks as scaffolds: integration of different types of biological interactions in a network, integration of expression data and networks (jActiveModule).

Integration of different types of biological interactions: Networks of Networks

- Load the 3 different networks available in Github in Cytoscape (PPI.gr, Pathways.gr, Complexes.gr (in GitHub)) (Import Network from File, take care at the network collection, column header and assignment of interaction sources and targets). Do not ask for the creation of a visualization for the Pathway and
- What is the intersection between the 3 networks ? Between PPI and Complexes ? Use Tools -> Merge -> Intersection
- Compute some statistics for the merged PPI_Complexes network, with Tools=> Analyze Network
- Select the largest connected component from this network and create a new network from it (File => new network from selected nodes, all edges)
- Run an partitioning algorithm (with App -> Clustermaker -> MCL Cluster (default))
- See clusters with Clustermaker -> clusterMaker Visualization -> create new network form clusters, or use the "Style" table, node color mapping to color clusters (Discrete mapping => right click, Mapping value generator)

Integration of Expression and Networks (jActive Module)

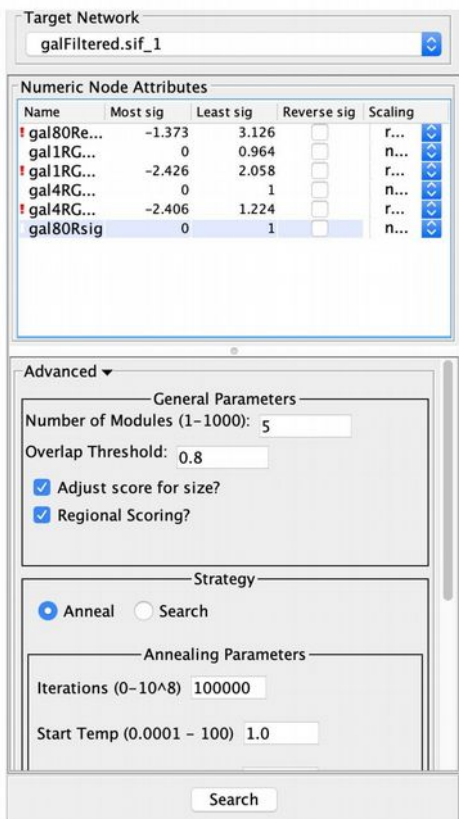
This tutorial is inspired from http://opentutorials.cgl.ucsf.edu/index.php/Tutorial:JActiveModules_3

- Install the jActiveModules App either from within Cytoscape 3 Apps → App Manager or directly from <http://apps.cytoscape.org/>
- Open the galFiltered.sif network (in GitHub) via File → Import → Network from File...
- Some of the network edges are directed ('pd') other are not ('pp') (see in the Edge Table), draw the directed edges with an arrow. To do so, Style => EdgeStyle => TargetArrowShape mapping according to the column "interactions", select "pd".
- Load the expression data matrix galExpData.csv (in GitHub) using the File → Import → Table from File... This file contains expression results for three sets of expression analysis from perturbation of

three transcription factors involved in the yeast galactose utilization pathway. This file also contains a necessary ingredient for jActiveModules: p-values indicating the significance of each expression value

- Color nodes according to gal8oRGexp
 - Node fill color
 - Continuous mapping

- Go to Apps → jActiveModules... The jActiveModules interface will appear as a tab in the Control Panel, select the correct target network
- Select gal8oRsig, Expand the Advanced panel, and run the algorithm with Anneal option. Notice that the Number of Modules is set to 5. This means that five putative hits will be returned even if only one good one is found



- When the results are ready you will see 5 new subnetworks as well as a network illustrating the search results
- Explore the 2 subnetworks with the highest scores and check their functions for instance with the BINGO plugin