Analysing Biological Networks with Cytoscape

Material adapted from a tutorial written by Alberto Calderone and from the Cytoscape user documentation at: http://www.cytoscape.org/documentation_users.html

Training session:

You will find basic commands here:
http://opentutorials.cgl.ucsf.edu/index.php/Tutorial:Introduction_to_Cytoscape_3

Loading Networks:
https://cytoscape.github.io/cytoscape-tutorials/protocols/modules/loading-network-data/#/

Loading Omics data:
https://cytoscape.github.io/cytoscape-tutorials/protocols/modules/loading-omics-data/#/

***

Do the Functional Enrichment analysis only if you had time left after completing the tutorial below.
Functional Enrichment analysis:
https://cytoscape.github.io/cytoscape-tutorials/protocols/modules/functional-enrichment/#/

***

This training session is focused on protein-protein interaction networks and on how to analyse them applying notions from Graph Theory. We are going to retrieve interaction information from an online resource that integrates data from different external databases and we are going to use this data to present properties of interest, such as interaction reliability score.

In this session you will:

- Go to the "mentha" interaction browser and retrieve the interactions of the ERBB gene in Homo sapiens
- Add other interactors to the network
- Download the network
- Start Cytoscape and import the network
- Prepare the network
- Analyse the network

1. Go to the "mentha" database: http://mentha.uniroma2.it

2. Type"erbb"in the search field and select *Homo sapiens*, then click "Search"

3. In the "Results Page" add ERBB2, ERBB3, ERBB4 and EGFR (also known as ERBB1) to the "Protein Bag" and click "List". A list of all interactions involving one of the ERBB proteins will appear.

4. Export the entire list of interactions by clicking "Simple".

5. Now, depending on your browser save this page to a text file on your computer. For instance, in Firefox you need to click the menu icon on the right and select “Save Page”, give it a name you remember, e.g. “erbb- interactions.csv”

6. Open Cytoscape.

7. We will now import interactions from the file you saved in step 5. Go to “File...” in the top menu “File” → “Import” → “Network” and select the file you save earlier.

8. A pop-up window will open showing the content of the file. Click on the “Advanced Option” button and specify as “Delimiter” semicolon. Click OK. In the window showing the content of the file in tabular format, click on the Gene A header and set it to “Source node” (green circle). Click on the Gene B header and set it to “Target node” (orange target sign). Click on the “Score” header and set it to “edge attribute” and check that it is a floating point number. Finally, click “OK”.

9. Click  to layout your network in a more pleasant way (or go to View → Fit content).

10. Now we are going to filter only interactions with a score higher than 0.4.

11. Click on “Select” tab and add a new condition clicking the “+” sign and “Column Filter”.

12. Select “Edge: Score” and set the interval between 0.4 and 1.0 (Cytoscape may be localised in your own language, pay attention to what decimal separator you use, either full-stop or comma). The filter should be applied automatically and some edges in your network should now be coloured in red (Selected edges).

13. Create a new network using only selected nodes (File → New → Network → From selected nodes, selected edges) (If you are using a version of Cytoscape lower than 3.2, you first need to click “Select” → “Nodes” → “Nodes connected by selected edges”)

14. You can now use the “Tools” → “NetworkAnalyzer” → “Network Analysis” → “Analyze Network” to see network topological features (Your network is undirected).

15. Go to View → Show results panel

16. Fill the following table:

Clustering coefficient (a measure of the degree to which nodes in a graph tend to cluster together)	
Network diameter	
Characteristic path length	
Avg. number of neighbors	
Number of nodes	

Would you say it is a “Scale free” network according to the definition given during the “Introduction to Graph Theory” session? For instance, click on “Node Degree Distributions” and try ”Fit Power Law” (Scale free networks are said to have a negative exponent between 2 and 3).

From Wikipedia: *A **scale-free network** is a network whose degree distribution follows a power law, at least asymptotically. That is, the fraction P(k) of nodes in the network having k connections to other nodes goes for large values of k as*

$P(k) \sim k^{-\gamma}$

Where γ is a parameter whose value is typically in the range $2 < \gamma < 3$, although occasionally it may lie outside these bounds.

17. “Hubs” are nodes with very high degree (number of links). What nodes do you think are“hubs” in you network? Find “Hubs” in “Table Panel” sorting by degree.

18. Now we would like our visualisation to communicate some information such as “Degrees” and “Interaction Reliability Score”. To this end, we are going to assign to each node a size proportional to its degree and to each link a colour proportional to its reliability score. Go to “Style” tab.

19. Check the “Lock node width and height”.

20. Double-click on the rounded rectangle beside “Shape” and select “Ellipse”.

21. Click on the little triangle beside “Size”. Set “Column” to “Degree”. Set “Mapping Type” to “Continuous”. Double click on the blue chart and set as minimum 10 and maximum 60 by double clicking the empty rectangles you see on the chart. Close this window.

22. Now go to “Edge” tab at the bottom of the Style panel.

23. Set “Stroke Colour” on “Score” and “Continuous” mapping.

24. Double clicking “Current mapping” you can change colour gradients. Double click on the top white triangle to change its colour, do the same for the right pointing triangle and set both to red; you should obtain something similar to the following image.

25. Playing with this setting you can highlight properties without looking at the actual numeric results of “Network Analysis”. For instance, you can now see that EGFR had more reported partner interactors compared to other ERBB proteins and you can also see that, despite the fact that ERBB proteins can form homo and etero dimers trimers and teramers, they actually do not have high

scores for each of these combinations. You can play around with other sub setting or graphical features to highlight different aspects of this network. Try selecting only ERBB proteins and creating a new network.

Conclusions

Starting from a collection of proteins you now know how to collect protein interaction information and how to extract a network from them. You now know how to import such networks to Cytoscape and how to analyse their characteristics. Finally, you are able to assign to different properties different attributes, such as colours, to highlight specific characteristics, like best interaction partners.

Installing Apps

To install an App (e.g., the BiNGO gene set enrichment analysis App) directly from the App Store website:

1. Launch Cytoscape and keep it running.
2. Use a browser to load the BiNGO web page: http://apps.cytoscape.org/apps/bingo
3. Click on the 'Install' button.
4. A dialog should pop out showing the progress.
5. When installed, the button on the BiNGO web page will change to 'Installed'.

To install an App (e.g., the MCODE network module detection App) from within Cytoscape using the Cytoscape App Manager:

1. Launch Cytoscape – close the Welcome screen if it is still visible.
2. Go to menu Apps → App Manager
3. In the 'Search:' text box, type 'MCODE' (no quotes).
4. Select the MCODE App.
5. Click on the 'Install' button.
6. When installed, the 'Installed' button will become grey.