Network Biology I. – Introduction to Cytoscape

Dávid Fazekas & Balázs Bohár

November 12, 2018

1. **Setting up**
   Download the practice materials from the Dropbox folder, and open a New Session in Cytoscape.

   SignaLink.cys

2. Importing networks from files:

   2.1. Import the *SignaLink.cys* file as a network

   File → Import → Network → File...

3. **The Cytoscape User Interface**
   The three main UI elements we are going to work with: the Control Panel on the left, the Table Panel on the bottom, and the Network View in the middle.

   3.1. The control panel is used to select and manipulate the networks, styles, filters, plugins you want to use.
   The table panel in the bottom is used to access the data directly, you can view and work with nodes, networks from here.
   The network view is where the network visualization actually happens, you can select nodes/edges by clicking on them, or dragging a selection rectangle (ctrl+click).

   3.2. Try to select nodes using the table panel (right click → Select nodes from selected rows)

   3.3. What happens if you click on a node in the network view?

   3.4. We can manipulate our data directly using the table panel. Add a new column of the "string" or "integer" type.

   Click on the "+" icon → New Single Column → String

   3.5. Fill out the column or a selection of rows with the same input.

   Right click → Apply to entire column / Apply to selected nodes

   3.6. Cytoscape has built in functions to further manipulate your data. Have a look and select an applicable function based on the descriptions.

   Click on the f(x) icon

4. We can get some metrics about our network using the **NetworkAnalyzer** option.

   Tools → Network Analyzer → Network Analysis → Analyze Network

   4.1. Run the **NetworkAnalyzer**

   4.2. The analysis returns a nice summary page and adds quite a lot of new columns to our dataset.

   4.3. Average path length: average path length is a concept in network topology that is defined as the average number of steps along the shortest paths for all possible pairs of network nodes. It can be interpreted as a measure of the efficiency of information or mass transport on a network.

4.4. Clustering coefficient:In graph theory, a clustering coefficient is a measure of the degree to which nodes in a graph tend to cluster together.

4.5. Betweenness centrality: betweenness centrality is a measure of centrality in a graph based on shortest paths. *How central is my node? How many of the shortest paths go through it?*

4.6. Closeness: closeness centrality (or closeness) of a node is a measure of centrality in a network, calculated as the sum of the length of the shortest paths between the node and all other nodes in the graph.

4.7. Eccentricity: Reciprocal value of "the longest shortest path", a measure of centrality. High value means all nodes are close, low means at least one is far away.

4.8. Degree: the degree of the node shows the amount of connections it has.

5. **Filtering data**
We can get subsets of our data by filtering it using multiple properties.

5.1. Click on the **Select** menu in the control panel on the left side of the screen, and click on the "+" symbol afterwards.

5.2. There are multiple kinds of filters we can use. The column filter allows us to select data based on the imported values in our columns.
The degree filter allows the selection of nodes based on the amount of connections they have.
The topology filter lets you get to certain subgraphs by only selecting ones that have n amount of neighbours within a set distance.

5.3. **Task:** Select all **MAPK** proteins that have a degree > 5.

Column filter → column "name" → input "MAPK" into the query box → add the degree filter with the "+" icon

5.4. We can save filters and re-use them later even if we're working with completely different datasets.

Click on drop-down menu → Create new filter

6. **Subgraphs**

6.1. We can save the selected nodes as subgraphs, smaller networks. They will appear as a new network under our current one, and we can work on them separately.

6.2. Save the **MAPK** proteins as a new subnetwork.

File → New → Network → From selected nodes, all edges

7. **Styles**

7.1. One of the strongest points of Cytoscape are its easy to use and versatile visualization options. We can use preset styles for our network or create our own.

Click on the Style tab in the Control Panel

7.2. On the Style panel you can see three columns on the left side for each property: Default, Mapping and Bypass.

7.2.1. **Default**: sets the default value for all entities, a global setting

7.2.2. **Mapping**: discrete mapping lets you set a visual encoding for all values separately,
continuous mapping allows you to set an encoding (e.g. colour) along a function (i.e. from white to black, from small to large)
passthrough mapping is mostly used at names, it uses the value in the selected column as output

7.2.3. **Bypass**: bypass overwrites the former two and allows you to manipulate / highlight values manually

7.3. **Customizing styles**

7.3.1. Make sure the *Node* tab is selected on the bottom of the Control Panel.

7.3.2. Change the default colour of the nodes.

7.3.3. Change the default size of the nodes.

7.3.4. Try changing the size according to the degree using **continuous mapping.**

7.3.5. Colour part of the network manually using **bypass**.

7.3.6. Select the **Edge** tab on the bottom, and change the edge thickness based on one of the values we got from **NetworkAnalyzer**.

7.3.7. Select the **Network** tab on the bottom, and change the background colour

7.3.8. Save the style you created. These can be re-used any time.

Click on the drop-down menu→ Create new style

7.4. Create a clear style for the nodes and edges.

Delete the existing mapping settings from the Style

8. How you visualize a network is really context dependent, it is hard to define best practices. Because of this we are going to do the opposite. In the next 20 minutes try to create the *ugliest* layouts you possibly can. This is a good way to find out what works for you and what does not.

Useful links:

1. Up-to-date Cytoscape user manual: `http://manual.cytoscape.org/en/stable/`
2. Tutorials: `https://github.com/cytoscape/cytoscape-tutorials/wiki`
3. Cytoscape main webpage (downloads, apps): `http://www.cytoscape.org/`
4. Tamas' e-mail address: Tamas.Korcsmaros@earlham.ac.uk
5. David's e-mail address: David.Fazekas@earlham.ac.uk
6. Marton's e-mail address: Marton.Olbei@earlham.ac.uk