Basic Expression Analysis in Cytoscape-Human

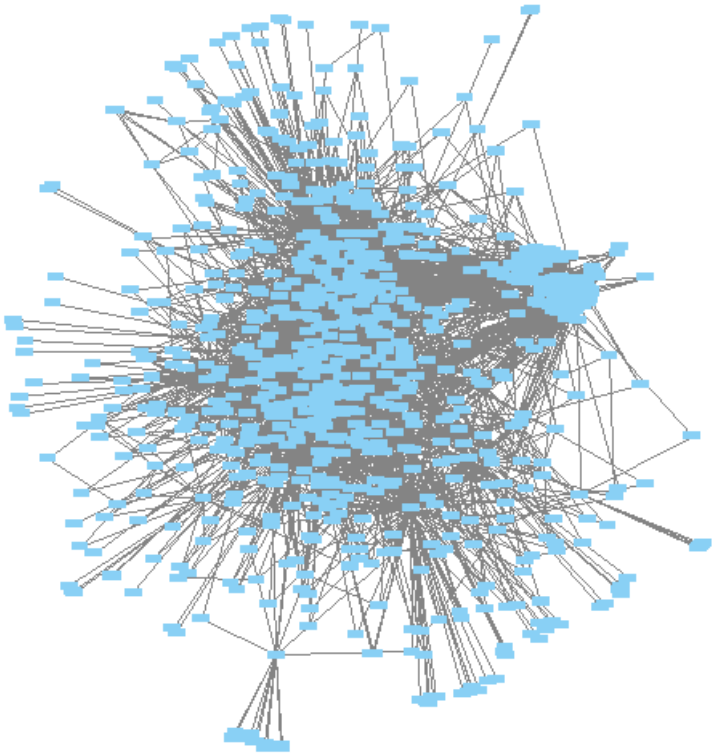
Tutorial Curators Kristina Hanspers, Alex Pico, Mike Smoot

Cytoscape is an open source software platform for integrating, visualizing, and analyzing measurement data in the context of networks. This tutorial will introduce you to:

- Combining data from two different sources: experimental data in the form of microarrray expression data and network data in the form of interaction data.
- Visualizing networks using expression data.
- Filtering networks based on expression data.

Loading Network

- Start Cytoscape and load the network **HumanInteractome_subset.sif**
- Apply the edge-weighted spring embedded layout to organize the layout of the nodes.
- The network should now look similar to this:



Loading expression data

- Using your favorite text editor, open the file **Pellegrini_et_al_Data.txt**. The first few lines of the file are as follows:

Pellegrini_et_al_Data.txt

```
1  Gene Symbol  Entrez id  Probeset  CREB kd control  p value  fold change  Sign  CREB binding
2  A1BG  1  229819_at  1286.5  1615.78  0.01  0.76  CR
3  A1CF  29974  220951_s_at  351.52  554.64  0.1  0.72  CR
4  A2BP1  54715  1553422_s_at  72.94  81.78  0.53  0.79  CR
```

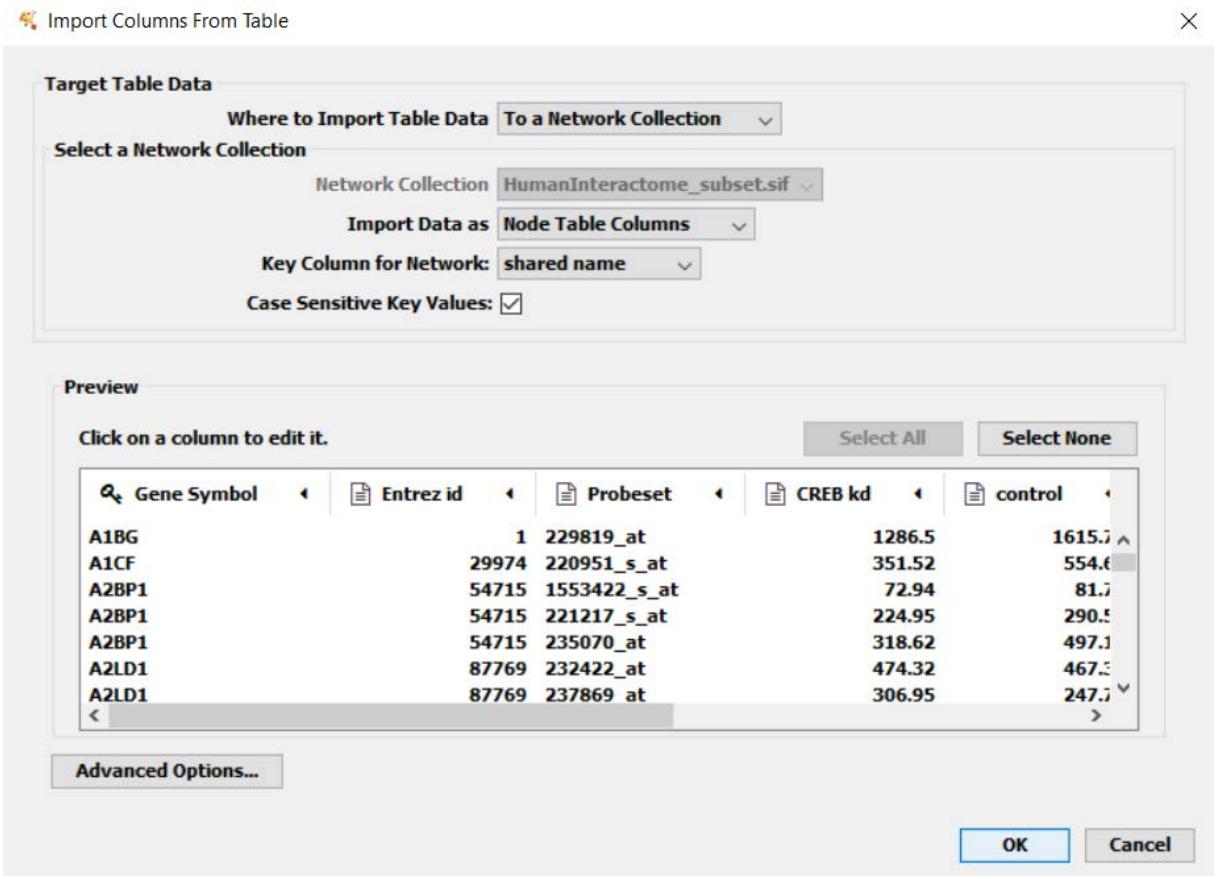
You should note the following information about the file:

- The first line consists of labels.

- All columns are separated by a single tab character.
- The first column contains node names, and must match the names of the nodes in your network exactly!
- The second column contains Entrez Gene IDs.
- The third column contains Affymetrix probe set IDs. This column is optional, and the data is not currently used by Cytoscape, but this column may be useful for analysis in other microarray analysis tools.
- The remaining columns contain experimental data; average expression for experimental and control groups, p-value and fold change for the comparison, and data on whether or not the gene binds CREB. See the manuscript [3] for details on the data generation.

NOTE: The expression data used in this example has been pre-processed to work with the interaction network used. The data is a composite of data files from Pellegrini et al, BMC Cancer, 2008 [3].

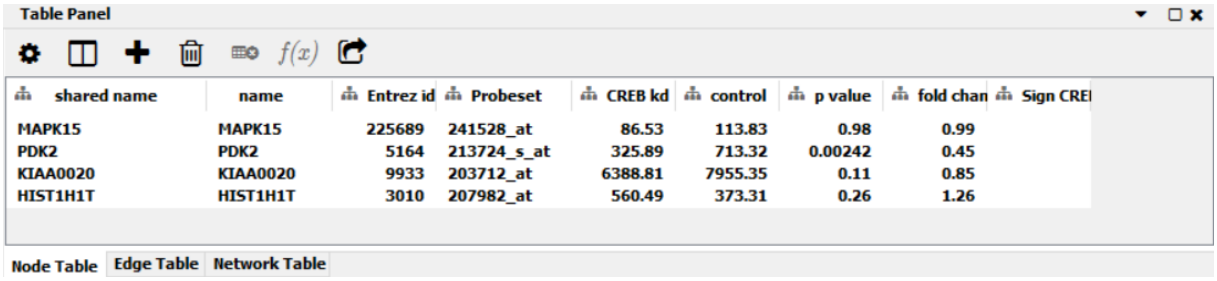
- Under the File menu, select Import > Table > File.
- Select the file **Pellegrini_et_al_Data.txt**.
- The preview should indicate that it is importing multiple columns of data, it should be using the first row of the input file as column names.
- Left-click the Gene Symbol column and indicate it as Key.
- The import window should look like the image below.
- Click the "OK" button to import the attribute data.



Now we will use the Node Table to browse through the expression data, as follows.

- Select a node on the Cytoscape canvas by clicking on it.

- In the Node Table, you should see your node listed with their expression (fold change) and significance values (p value), as shown.

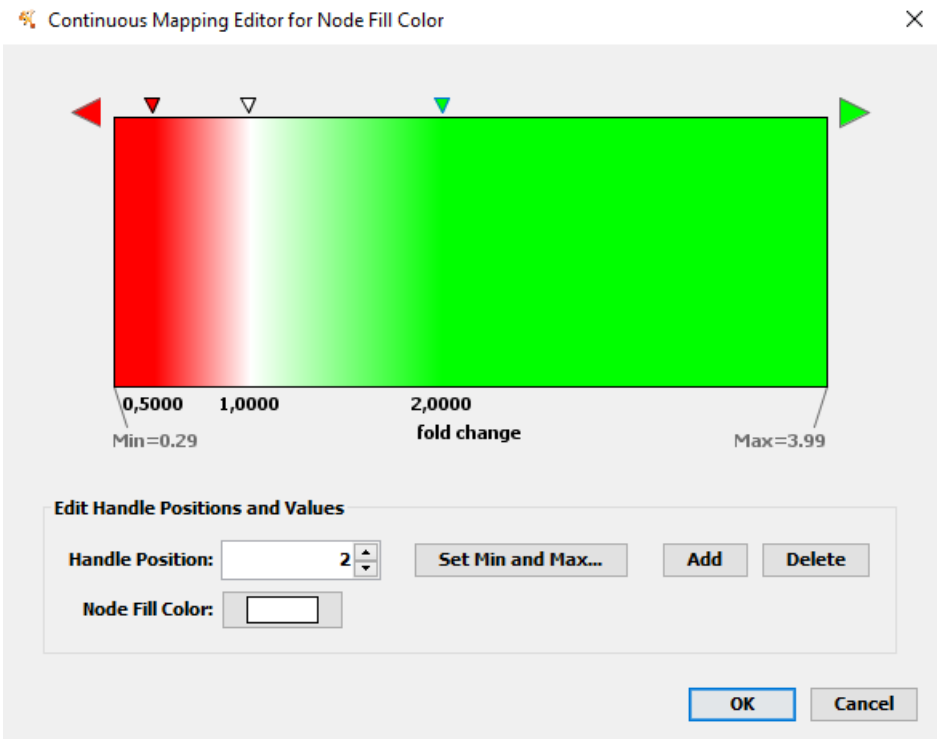


Visualizing Expression Data

Probably the most common use of expression data in Cytoscape is to set the visual attributes of the nodes in a network according to expression data. This creates a powerful visualization, portraying functional relation and experimental response at the same time. Here, we will walk through the steps for doing this.

Set the node color

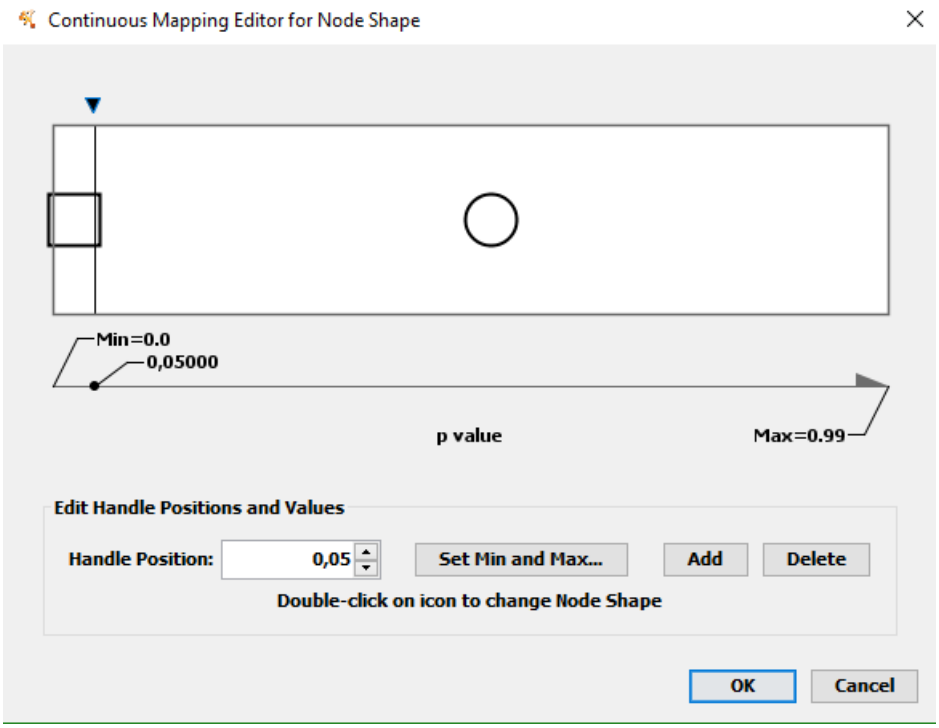
- Click the Fill Color row in the Style tab of the Control Panel.
- Click the "-- select value --" cell in the Column section.
- This will produce a drop-down menu of available attribute names. Select "fold change".
- Click the "-- select value --" cell in the Mapping Type section.
- This will produce a drop-down menu of available mapping types. Select "Continuous Mapping".
- This action will produce a basic white to red color gradient.
- Double-Click on the color gradient to change the colors. This will pop-up a gradient editing dialog.
- Double-Click on the left-most triangle to change the below color. Choose a bright red color.
- Click on the second triangle and slide it towards the center of the scale so that its value is close to 0,5 and select a bright red color
- Repeat for the right-most triangle to change the above color. Choose a bright green color.
- Click on the third triangle and slide it towards the center of the scale so that its value is close to 1 and select a white color.
- Click the Add button to add a new white triangle to the scale.
- Double-Click this new triangle and select a bright green color. Slide it so that its value is close to 2.
- This should produce a full Red-White-Green Color gradient like the image below.
- Close the gradient adjustment dialog and verify that the nodes in the network reflect the new coloring scheme.



Set the node shape

We imported both a fold change value and a p-value for the comparison between CREB kd and control cells. We can use the p-values to change the shape of the nodes so that measurements we have confidence in appear as squares while potentially bad measurements appear as circles.

- Click the Shape row in the Style tab of the Control Panel.
- Click the "-- select value --" cell in the Column section.
- This will produce a drop-down menu of available attribute names. Select "p value".
- Click the "-- select value --" cell in the Mapping Type section.
- This will produce a drop-down menu of available mapping types. Select "Continuous Mapping".
- This will create an empty icon in the "Current Mapping" row of the Shape section. Double-Click on this icon.
- This action will pop-up a continuous shape selection dialog.



- Click the Add button.
- This action will split the range of values with a slider down the middle with a node shape icon to either side of the slider.
- Double-Click on the left node icon (a circle).
- This will pop-up a node shape selection dialog.
- Choose the Rectangle shape and click the Apply button.
- The continuous shape selection dialog should now show both a square and a circle node shape icon.
- Click on the black triangle and move the slider to the left, to slightly lower that 0.05, our threshold for significance.
- Close the continuous shape selection dialog and verify that some nodes have a square shape and some nodes have a circular shape.

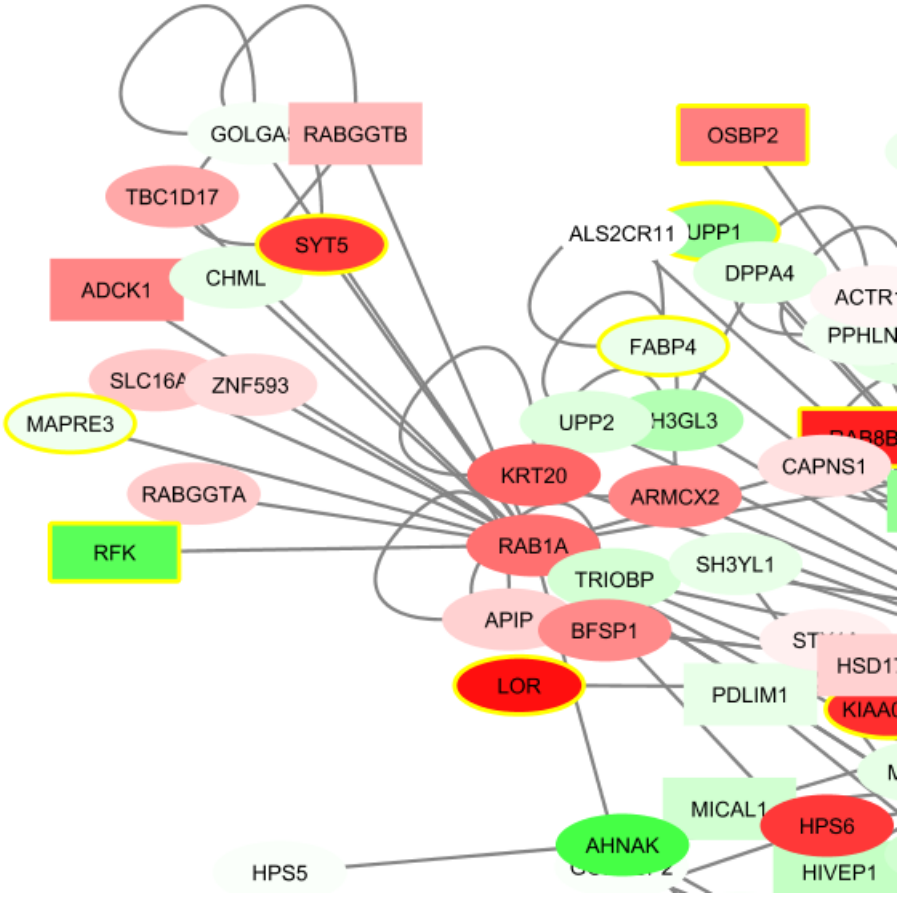
Set the node border

We can use the node border color and style to reflect whether a node has been found to be significantly bound by CREB. This data is already available in the dataset as an attribute.

- Click the Border Paint row in the Style tab of the Control Panel.
- Click the "-- select value --" cell in the Column section.
- This will produce a drop-down menu of available attribute names. Select "Sign CREB binding".
- Click the "-- select value --" cell in the Mapping Type section.
- This will produce a drop-down menu of available mapping types. Select "Discrete Mapping".
- This will create a new row for the value "yes", which is the only value available for this attribute. Click on the empty cell to the right of "yes", and then click on the square icon that appears.
- A color chooser will appear. Pick a color that will stand out against the color scheme, for example a bright yellow. The relevant nodes in the network will be outlined in yellow.

Next, we want to also change the border thickness, since the thin yellow border can be hard to see.

- Click the Border Width row in the Style tab of the Control Panel
- Click the "-- select value --" cell in the Column section.
- This will produce a drop-down menu of available attribute names. Select "Sign CREB binding".
- Click the "-- select value --" cell in the Mapping Type section.
- This will produce a drop-down menu of available mapping types. Select "Discrete Mapping".
- This will create a new row for the value "yes", which is the only value available for this attribute.
- Click in the box specifying the width and type 5.
- Zooming in to part of the network, it now looks like this:



Data analysis features

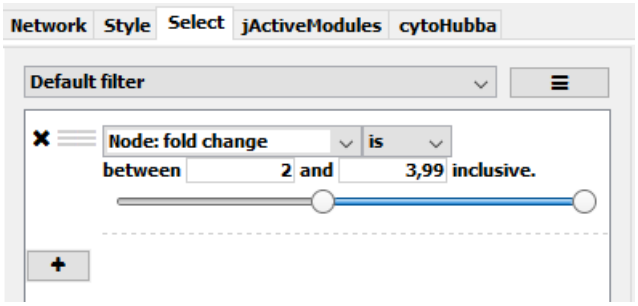
This section presents a few examples of features in Cytoscape that can be used to further analyze the network and associated data.

First, here is some background on your data. The data is from an experiment in a human myeloid leukemia cell line. The cAMP Response Element Binding Protein, CREB, was knocked down by shRNA and the expression profile of knockdown cells was compared to that of control cells from the same cell line.

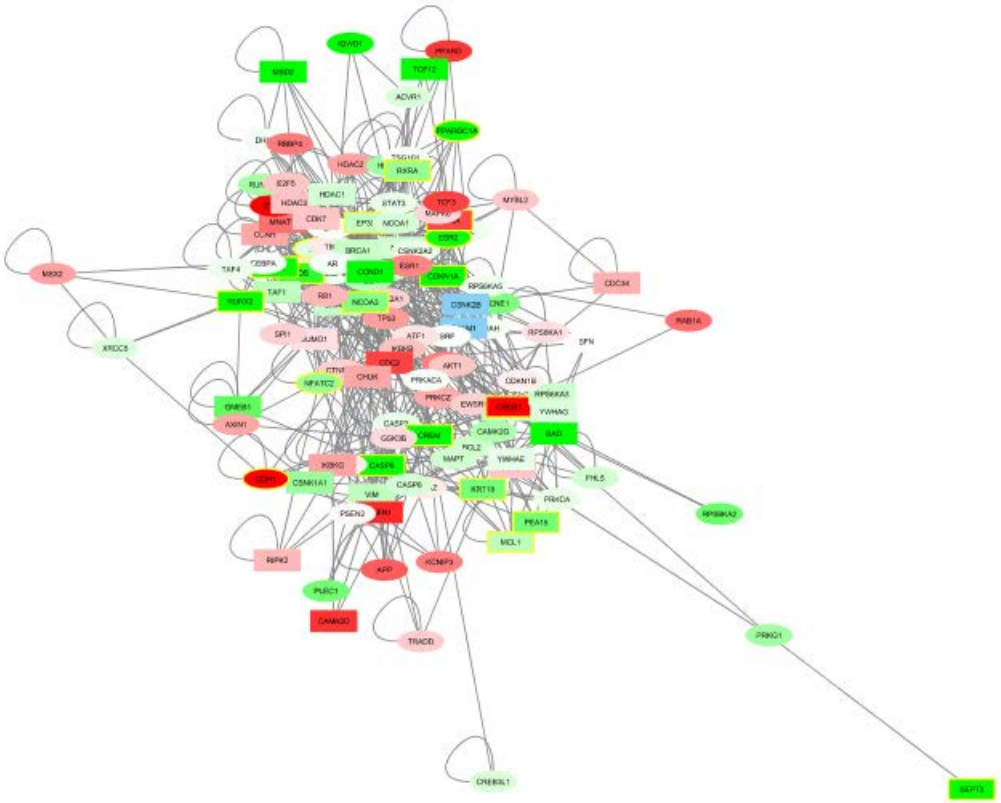
Filter Nodes

It is possible to filter any network in Cytoscape based on either node or edge attributes. Here, we filter the network based on high and low fold change between the two groups.

- Click the Select tab in the Control Panel.
- Click the “Add new condition” sign and select “Column Filter”.
- Choose the "Node: fold change" column.
- Set the low bound to 2.

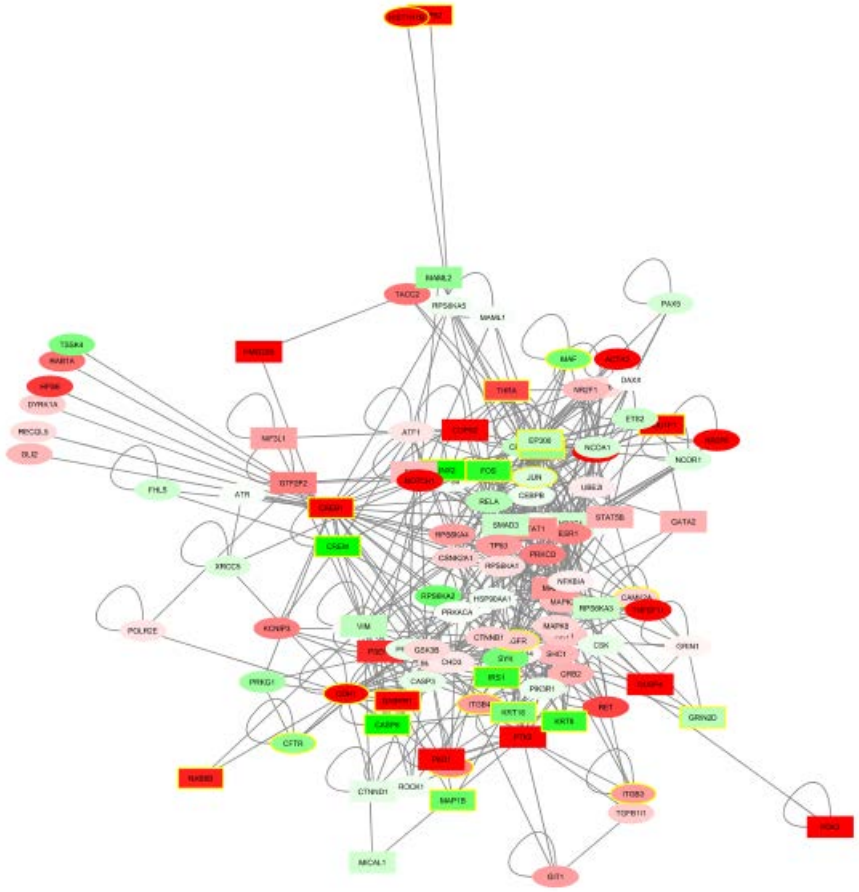


- Now, expand the selection to first neighbors of selected nodes by clicking Select, Nodes, First Neighbors of Selected Nodes, Undirected.
- Create a new network by clicking File, New, Network, From Selected Nodes, All Edges.
- Apply an edge-weighted spring embedded layout to the new network.
- Navigate to the Network tab in the Control Panel. Rename the new network by right-clicking on it and selecting Rename network. Type in "upregulated".
- The new network should now look like this:



- Repeat the Filter for down-regulated genes with fold change under 0.5. Name the second subnetwork "downregulated".

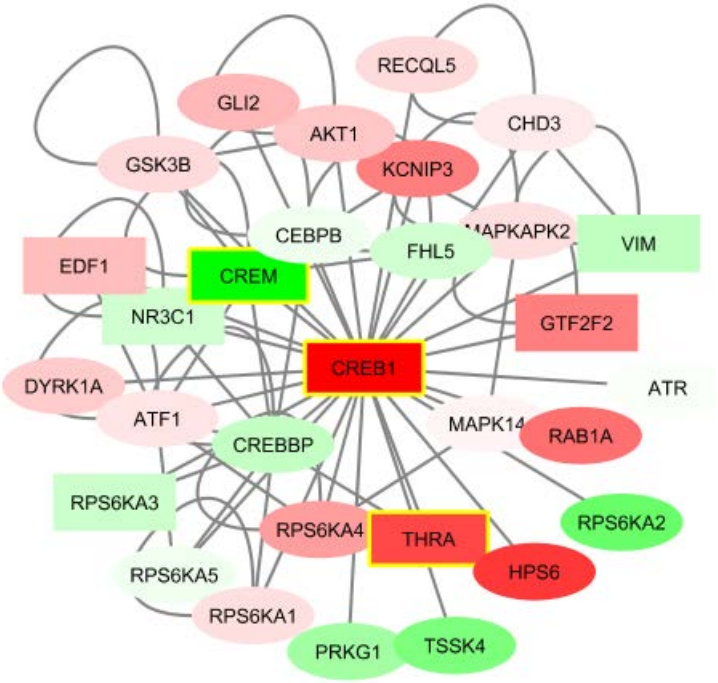
- The "downregulated" network should look like this:



Search for a node

We will now search for the CREB1 (CREB) node in the network.

- In the search field, type in "name:CREB1" and Enter. In the Node table, you will see that there is one node named CREB1, which is an alias name for the CREB transcription factor. Select this node from the list and click Enter.
- The CREB node will be highlighted in the network.
- To make it easier to explore the interactions immediately surrounding CREB, we can create a network based on the first degree neighbors of CREB by clicking Select, Nodes, First Neighbors of Selected Nodes, Undirected.
- A set of nodes should now be highlighted. Click File, New, Network, From Selected Nodes, All Edges.
- A new network will be produced.
- Clean up the network by applying a prefuse force directed layout.
- The network should now look like this:



- By examining the network, we can see that CREB interacts with both up- and down-regulated nodes. CREB is known to act both as a transcriptional repressor and transcriptional activator. If you search for CREB in the two subnetworks representing all up- and down-regulated nodes respectively, you will see that CREB is present in both networks.