Tutorial:Basic Expression Analysis in Cytoscape 3

Slideshow Basic Expression Analysis in Cytoscape 3 (30 min) ^[1]
 Handout Basic_Expression_Analysis_in_Cytoscape_3.pdf (10 pages) ^[2]
 Tutorial Sources Cytoscape Wiki (Natalie Yeung) ^[3]
 Tutorial Curators Kristina Hanspers, Alex Pico, Mike Smoot, Scooter Morris

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:

- Visualizing networks using expression data.
- Filtering networks based on expression data.
- Assessing expression data in the context of a biological network.

Loading Network

- Start Cytoscape and load the network galFiltered.sif using File → Import → Network->File....
- By default, the "Prefuse Force Directed Layout" is applied to organize the layout of the nodes. The default layout
 may be changed in the Layout → Settings menu.
- Cytoscape should now look similar to this:



Loading expression data

• Using your favorite text editor, open the file galExpData.csv. The first few lines of the file are as follows:

```
GENE, COMMON, gal1RGexp, gal4RGexp, gal80Rexp, gal1RGsig, gal4RGsig, gal80Rsig
YHR051W, COX6, -0.034, 0.111, -0.304, 3.75720e-01, 1.56240e-02, 7.91340e-06
YHR124W, NDT80, -0.090, 0.007, -0.348, 2.71460e-01, 9.64330e-01, 3.44760e-01
YKL181W, PRS1, -0.167, -0.233, 0.112, 6.27120e-03, 7.89400e-04, 1.44060e-01
```

You should note the following information about the file:

- 1. The first line consists of labels.
- 2. All columns are separated by a single comma character.
- 3. The first column contains node names, and must match the names of the nodes in your network exactly!
- 4. The second column contains common locus names. This column is optional, and the data is not currently used by Cytoscape, but including this column makes the format consistent with the output of many microarray analysis packages, and makes the file easier to read.
- 5. The remaining columns contain experimental data, two columns per experiment (one column represents the expression measurement and the second represents the significance value for that measurement), and one line per node. In this case, there are three expression results per node.
- Under the File menu, select Import → Table → File....
- Select the file galExpData.csv.
- The default settings are usually correct, although you may need to change the Key column to indicate which column will be used to match with the network key column.
- Click the "OK" button to import the data.

get Table Data						
Select a Network Collection	Where to Import Tab	ole Data: To a Netw	vork Collection	0		
	Network Co	llection: galFiltere	d.sif 🗘			
	Import	Data as: Node Tab	ole Columns 📀			
	import bata as. Note tuble commis					
	Key Column for N	letwork: snared na	ime 🗘			
	Case Sensitive Key	Values: 🗸				
				Serect		
🦧 gene 🕢 🖹 common	◀ 📄 gal1RGexp ◀	🖹 gal4RGexp ◀	🖹 gal80Rexp 🖪	🖹 gal1RGsig 🛛 🖣	🖹 gal4RGsig	
α GENE	I		■ gal80Rexp	■ gal1RGsig 4 3.75720e-01	gal4RGsig 1.56240e-	
Q GENE ↓ E COMMON (HR051W COX6 (HR124W NDT80	al 4 a a a a a a a a a a	gal4RGexp (0.111 0.007 0.222	■ gal80Rexp 4 -0.304 -0.348	■ gal1RGsig 3.75720e-01 2.71460e-01	gal4RGsig 1.56240e- 9.64330e-	
Q GENE Image: COMMON (HR051W COX6 (HR124W NDT80 (KL181W PR51 (CR072W LIPF3	al () allRGexp (-0.034 -0.090 -0.167 0.245	■ gal4RGexp	■ gal80Rexp -0.304 -0.348 0.112 0.787	■ gal1RGsig	gal4RGsig 1.56240e- 9.64330e- 7.89400e- 7.51280e- 	
Qcone Image: Common (HR051W) COX6 (HR124W) NDT80 (KL181W) PR51 (GR072W) UPF3 (HL020C) OPI1	a () () () () () () () () () () () () ()	■ gal4RGexp 4 0.111 0.007 -0.233 -0.471 -0.015	■ gal80Rexp	gal1RGsig ◀ 3.75720e-01 2.71460e-01 6.27120e-03 4.10450e-04 1.40160e-04	■ gal4RGsig 1.56240e- 9.64330e- 7.89400e- 7.51780e- 7.51780e- 7.9120e-	
Qc GENE Image: COMMON YHR051W COX6 YHR124W NDT80 YKL181W PR51 YGR072W UPF3 YHL020C OPI1 YGR145W YGR145W	a () () () () () () () () () () () () ()	■ gal4RGexp 4 0.111 0.007 -0.233 -0.471 -0.015 -0.577	■ gal80Rexp 4 -0.304 -0.348 0.112 0.787 0.151 -0.088	gal1RGsig (3.75720e-01 2.71460e-01 6.27120e-03 4.10450e-04 1.40160e-04 5.37920e-03	gal4RGsig 1.56240e- 9.64330e- 7.89400e- 7.51780e- 7.19120e- 8.27330e-	
Qc GENE COX6 YHR051W COX6 YHR124W NDT80 YKL181W PR51 YGR07ZW UPF3 YHL020C OPI1 YGR145W YGR145W YGR145C YGL041C	a a second secon	gal4RGexp 4 0.111 0.007 -0.233 -0.471 -0.015 -0.577 -0.866 -0.866	■ gal80Rexp 4 -0.304 -0.348 0.112 0.787 0.151 -0.088 0.103	gal1RGsig ◀ 3.75720e-01 2.71460e-01 6.27120e-03 4.10450e-04 1.40160e-04 5.37920e-03 4.46050e-04 4.46050e-04	gal4RGsig 1.56240e- 9.64330e- 7.89400e- 7.51780e- 7.19120e- 8.27330e- 4.50790e-	

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 can limit the columns shown by going to the Show Columns button \square , and select the attributes gal1RGexp, gal4RGexp, and gal80Rexp by clicking on the checkbox.
- Under the Node Table, you should see your node listed with their expression values, as shown.

Table Panel									□ ¥
* 🗆 +	t to f	(x)							
shared name	name	A COMMON	gal1RGex 🚢	gal4RGex 🚢 g	gal80Rexp	📥 gal1RGsig 🗄	gal4RGsig	📥 gal80Rsig	
YLR362W	YLR362W	STE11	0.144	0.037	0.374	0.023963	0.50139	0.00108	
				Node Table	Edge	e Table Net	work Tabl	e	

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.

Label the Nodes

- Open the Styles by selecting its tab in the "Control Panel" (the leftmost panel).
- Use the "Common" name attribute to give the nodes useful names.
- Zoom in on the network so that node labels are visible.
- Click the second column of the "Label" row in the Styles panel.
- This should produce a drop-down panel with "Column" and "Mapping Type"
- Change the "Column" to **COMMON** by clicking on the field to the right of the "Column" label. This should bring up a list of columns. Select **COMMON**.
- Verify that the node labels on the network have changed to their common names.
- By default, the "Mapping Type" is **Passthrough Mapping**, which is what we want to use. Other options are **Discrete Mapping** and **Continuous Mapping**.

Color the nodes

Define the node color of this visual style:

- Click on the middle square (Map.) next to the Fill Color row in the Styles panel.
- Click the "-- select value --" cell in the Column section.
- This will produce a drop-down menu of available column names. Select "gal80Rexp".
- 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 black to white color gradient.
- Click on the color gradient to change the colors. This will pop-up a gradient editing dialog.
- We're going to build a basic blue-white-yellow gradient for our expression values.
- Drag the left-most, black inverted triangle handle along the top of the gradient. Drag it to an value of approx. -1.2. Double-click on the handle and set a color in the blue range.
- Drag the white inverted triangle handle to approx 0.5. You can type the value in the Handle Position section to be more precise.
- Add a new handle by clicking Add, and drag that handle to 2.5. Set the color to yellow.
- Finally, set the Maximum Color by double-clicking on the white, left-pointing triangle. Set it to gree.
- You can also change the color of each handle by double-clicking or using the Node Fill Color selector button in the Handle Settings section.
- This should produce a Blue-White-Yellow Color gradient like the image below, with min and max extremes colored black and green, respectively.

• Click 'OK' to save the gradient adjustment dialog and verify that the nodes in the network reflect the new coloring scheme.

	Continuous Ma	apping Editor for Node Fil	ll Color
< _▼		∇	—
-1.2028	0.4	4996	2.5018
Min=-1.373		galookexp	Double-click handles to edit bounda
Edit Handle Positions	and Values		
Handle Position:	2.502 🗘	Set Min and Max	Add Delete
Node Fill Color:			
			Cancel OK

Set the default node color

Note that the default node color of pale blue falls within this spectrum. A useful trick is to choose a color outside this spectrum to distinguish nodes with no defined expression value and those with slight repression.

- Click the Def. (leftmost) square next to "Fill Color" and choose a dark gray color.
- Zoom out on the network view to verify that a few nodes have been colored gray.

Set the Node Shape

We imported both expression measurement values and significance values for those measurements. We can use the significance 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 Map. cell next to the Shape row in the Style Panel.
- Click the "-- select value --" cell next to Column.
- This will produce a drop-down menu of available column names. Select "gal80Rsig".
- Click the "-- select value --" cell next to Mapping Type.
- 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. Click on this icon.
- This action will pop-up a continuous shape selection dialog.

🔴 🔵 🔵 Cor	tinuous Mapping Editor for Node Shap	9
•		
	\sim	
	\bigcirc	
/ Min=1.7403E-20		
0.05008		
	gal80Rsig	Max=0.9999999
Edit Handle Positions and Values		
Handle Position:	0 🗘 Set Min and Max	Add Delete
D	ouble-click on icon to change Node Shape	
		Cancel OK

- 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.

The network should now look like this:



A biological analysis scenario

This section presents one scenario of how expression data can be combined with network data to tell a biological story.

First, here is some background on your data. You are working with yeast, and the genes Gal1, Gal4, and Gal80 are all yeast transcription factors. Your expression experiments all involve some perturbation of these transcription factor genes. Gal1, Gal4, and Gal80 are also represented in your interaction network, where they are labeled according to yeast locus tags: Gal1 corresponds to YBR020W, Gal4 to YPL248C, and Gal80 to YML051W.

Filter Interactions

Your network contains a combination of protein-protein (pp) and protein-DNA (pd) interactions. Here, we shall filter out the protein-protein interactions to focus on the protein-DNA interactions.

- Click the Select tab in the Control Panel.
- Click the + icon + and select Column Filter.
- Select the drop-down list and choose Edge: interaction.
- This action will create a text search box entry in the filter.
- Type the letters "pp" into the text search box. This indicates that we're searching for all edge interaction attributes that match the string "pp".
- For reasonably sized networks, the filter will automatically be applied. For larger networks, you may need to click the **Apply** button at the bottom of the Select panel.

	-	*	S	.	F	Ð	Q
Control Par	nel	• •					⊐ ×
		Network	Style	Selec			
Defau	lt filter				<u></u>	•	_
×	Edge: ir	nteraction	٢	conta	lins	٢	
	pp						
+							
							0
🔽 App	ly when	filter char	iges				
Appl	y						\oslash
Selected () nodes ar	nd 251 edge	s in 80ms				
		FI	ter Ch	nain			

- You should now see many edges in the network selected (i.e. colored red).
- Since we're only interested in the protein-DNA edges, we can delete the protein-protein edges we've just selected.
- Select the menu Edit → Delete Selected Nodes and Edge. You should now see many unconnected nodes in the network.
- Select the menu Layout → Cytoscape Layouts → Prefuse Force-Directed Layout to clean up the network visualization.
- The largest component of the final filtered and cleaned up network should look like this:



Observe the Network

Notice that three bright green (highly induced) nodes are in the same region of the graph. Zoom into the graph to see more details.

- Notice that there are two nodes that interact with all three green nodes: GAL4 (YPL248C) and GAL11 (YOL051W).
- Select these two nodes and their immediate neighbors by selecting the menu Select → Nodes → First Neighbors of Selected Nodes → Undirected.
- It is sometimes useful to create a new network from selected nodes. Do this by selecting the menu File → New → Network → From Selected Nodes, All Edges.
- With some layout and zooming, this new network should appear similar to the one shown:



Exploring Nodes

- Right click on the node GAL4.
- Select the menu External Links → Sequences and Proteins → Entrez Gene.
- This action will pop-up a browser window and search the Entrez Gene database for the term "YPL248C", the id of the node.
- In the results in the browser the first entry should be labeled GAL4. Click on this entry.
- The description of GAL4 tells us that it is repressed by GAL80.
- Our data show precisely this:
 - Both nodes (GAL4 and GAL11) show fairly small changes in expression, and neither change is statistically significant: they are rendered as light-colored circles. These slight changes in expression suggest that the critical change affecting the black nodes might be somewhere else in the network, and not either of these nodes.
 - GAL4 interacts with GAL80 (YML051W), which shows a significant level of repression: it is depicted as a red square.
 - Note that while GAL80 shows evidence of significant repression, most nodes interacting with GAL4 show significant levels of induction: they are rendered as green squares.
 - GAL11 is a general transcription co-factor with many interactions.
- Putting all of this together, we see that the transcriptional activation activity of Gal4 is repressed by Gal80. So, repression of Gal80 increases the transcriptional activation activity of Gal4. Even though the expression of Gal4 itself did not change much, the Gal4 transcripts were much more likely to be active transcription factors when Gal80 was repressed. This explains why there is so much up-regulation in the vicinity of Gal4.

Fun with Charts

In addition to coloring the nodes, Cytoscape also provides the ability to draw charts and graphs on each node. For example, suppose we wanted to display a bar chart showing all of the expression values on each of our nodes?

- To reset things a little, remove the mapping for **Fill Color** by doing a Right-Click over the **Fill Color** row and selecting **Edit** → **Remove Mappings from Selected Visual Properties**
- Now change the default value to a lighter shade of grey so we can see our chart.
- Near the top of the panel, select **Properties** and choose **Paint → Custom Paint 1 → Image/Chart 1**. This will add a new row in our list of Node Visual Properties called **Image/Chart 1**.
- Select the Def. (leftmost) cell in the Image/Chart 1 row to bring up the Graphics dialog.
- Select the Charts tab.
- Move the three columns containing the expression data (gal1RGexp, gal4RGexp, gal80Rexp) from "Available Columns:" to "Selected Columns:" by selecting the rows and clicking the right arrow. This indicates that we're going to use the data from these three columns to create our chart.
- Now select "Heat Strips" for the type of bar chart.

	Graphics	
	Images Charts Gradients	
	○ Grouped ○ Stacked ○ Heat Strips ○ Up-Down	
Available Columns:	Selected Columns:	A V
1. gal1RGsig 2. gal4RGsig 3. gal80Rsig	> 1. gallRGexp 2. gal4RCexp > 3. gal80Rexp	
Data	«	
🗸 🗹 Network-Wide Axis Rar	ige	
🖸 🗹 Automatic Range Min	: -2.426 Max : 3.126	3
đ		
Remove Graphics	Cancel	Apply

- Click on "Options" if you want to add labels to the graphs, change the default coloring, etc.
- Click **Apply** to see the resulting charts.



References

- [1] http://opentutorials.rbvi.ucsf.edu/index.php?title=Tutorial:Basic_Expression_Analysis_in_Cytoscape_3&ce_slide=true& ce_style=cytoscape
- $\cite{thm:line(1)} \cite{thm:line(2)} tttp://opentutorials.rbvi.ucsf.edu/index.php/File:Basic_Expression_Analysis_in_Cytoscape_3.pdf$
- $[3] \ http://cytoscape.wodaklab.org/wiki/Presentations/04_Expression_Data$

Article Sources and Contributors

Tutorial:Basic Expression Analysis in Cytoscape 3 Source: http://opentutorials.cgl.ucsf.edu/index.php?oldid=3096 Contributors: ScooterMorris

Image Sources, Licenses and Contributors

File:GalFilteredLoaded3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:GalFilteredLoaded3.png License: unknown Contributors: ScooterMorris
File:LoadGalExp3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Select3.png License: unknown Contributors: ScooterMorris
File:galbrowse3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Select3.png License: unknown Contributors: ScooterMorris
File:galbrowse3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Select3.png License: unknown Contributors: ScooterMorris
File:Node_color_gradient3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Node_color_gradient3.png License: unknown Contributors: ScooterMorris
File:Node_shape_editor3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Select3.png License: unknown Contributors: ScooterMorris
File:Galfiltered-visualization3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Galfiltered-visualization3.png License: unknown Contributors: ScooterMorris
File:Select_panel.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Select_panel.png License: unknown Contributors: ScooterMorris
File:Select_panel.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Select_panel.png License: unknown Contributors: ScooterMorris
File:Final_network3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Final_network3.png License: unknown Contributors: ScooterMorris
File:Final_subnetwork3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Final_network3.png License: unknown Contributors: ScooterMorris
File:Final_subnetwork3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Final_subnetwork3.png License: unknown Contributors: ScooterMorris
File:Final_subnetwork3.png Source: http://opentutorials.cgl.ucsf.edu/index.php?title=File:Final_subnetwork3

License

Attribution-Noncommercial-Share Alike 3.0 Unported http://creativecommons.org/licenses/by-nc-sa/3.0/